Indian Journal of Natural Sciences www.tnsroindia.org.in ©IJONS

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Anti-Colorectal Cancer Properties of Hill Banana (cv.Virupakshi AAB) fruits: An *in vitro* assay

M.Praveena[2], M. Surya Prabha[2] , I Ravi[1]* and M Mayil Vaganan[1]

[1]Principal Scientist, ICAR-National Research Center for Banana, Thayanur Post, Trichirappalli, TamilNadu, India.
[2]M.Sc.Student, Dr. N. G. P. Arts and Science College, Coimbatore – 48, TamilNadu,India.

Received: 25 Feb 2018 Revised: 15 Mar 2018 Accepted: 30 Mar 2018

***Address for correspondence**
Dr.I Ravi
Principal Scientist, Crop Production Section**,**
ICAR-National Research Center for Banana,
Thayanur Post, Trichy,TamilNadu,India.
E.mail: iravi24@gmail.com; Ravi.I@icar.gov.in

ABSTRACT

Banana fruits are known for nutrients, antioxidants and other phytonutrients which provides immense disease protection through providing many nutrients, dietary fibers and vitamins.The ripened banana fruits are having higher anti-cancer properties than unripened fruits. Colorectal cancer is the fourth most common cancer in the world and in India, it is the fifth most common cancer. The *in vitro* assay of hill banana (Virupakshi) fruit juice revealed that, it has higher potential to inhibit growth the HT-29 cells and causes mortality at very low concentration of fruit juice. We have illustrated that the banana fruit juice has inhibited the proliferation of the colon cancer cell line HT-29 at low concentration through *in vitro* assay.

Keywords: Hill Banana Fruit; Anti-Colorectal Cancer; MTT Assay ; ETBr Assay

INTRODUCTION

In India, banana fruits are called as poor man's apple and “kalpatharu" which means a "virtuous plant. As a whole plant, roots, stem and flowers of banana plants have been widely used in traditional system of medicine for various ailments [1, 2]. Bananas are loaded with antioxidant compounds, thereby helping to reduce premature aging of the body’s cells. Alpha linolenic acid, one of these compounds, demonstrates significant antioxidant, anti-inflammatory and anti-cancer activity**.** Alpha linolenic acid also boosts immune function and enhances blood circulation. foods contain not only nutrients but also large amounts of compounds called phytochemicals. About 10,000 types of phytochemicals are considered to be present in nature. As banana ripening process progresses, its antioxidant levels rises. The well ripened banana fruit has better anti-cancer property and the dark spots on ripe yellow bananas

produce a substance called Tumor Necrosis Factor (TNF) that destroys cancerous tumors [3]. In terms of cancer fighting potential, the ripened bananas are eight times as effective as in their un-ripened state of green banana [3]. Colorectal cancer is the fourth most common cancer in the world with 1.3 million new cases each year and a 5-year prevalence rate of 3.2 million [4, 5]. There were an estimated 693,333 deaths due to colorectal cancer (CRC) in 2012. In India, it is the fifth most common cancer following breast, cervix/uteri, lip/oral cavity, and lung cancer [6]. In India, it is predicted to rise approximately by 80% in 2035, with an incidence of 114,986 new cases and a mortality of 87,502 [3].

There are many types of cancer treatments, it includes surgery, chemotherapy, radiation therapy, hormonal therapy or combined therapy. However, these kinds of therapies known to have different side-effects. Hence, identification and development of new chemotherapeutic agents from plants "phytochemicals" have gained significant recognition in the field of cancer therapy and become a major area of experimental cancer research [7].Recently, scientists all over the world are concentrating on the herbal medicines to fight against cancer. By understanding the complex synergistic interaction of various constituents of anticancer herbs, new novel herbal anticancer agents can be discovered and designed to attack the cancerous cells without affecting normal cells of the body [8]. With this background, the objective of the work is to analyze the anti-colon cancer properties of hill banana fruits through *in vitro* assay.

MATERIALS AND METHODS

The experiment was carried out ICAR-National Research Center for Banana, Trichy. The hill banana fruits (Virupakshi) are purchased from the local market and allowed to ripen fully at 19 - 21^{0}C. After full ripening the pulp was homogenized and pectinase enzyme (5 ml / Kg of pulp) was added and incubated for two hours. Then the clear banana juice was extracted and subjected to MTT and ETBr AO (Ethidium Bromide Acridine Orange dye). All this operation was done under sterile to ensure free from contamination. The hill banana juice (BJ) was assayed for anti-colorectal cancer activity through *in vitro* MTT and EtBr assays. These assays were done at Trichy Research Institute of Biotechnology (P) Ltd (TRI Biotech), Trichy. The procedure is described briefly as follows.

Cell culture

HT - 29 cells (Human Colon Carcinoma) cell line were cultured in liquid medium (DMEM) supplemented 10% Fetal Bovine Serum (FBS), 100 u/ml penicillin and 100 µg/ml streptomycin, and maintained under an atmosphere of 5% CO_2 at 37^{o}C.

MTT Assay

The BJ Sample was tested for *in vitro* cytotoxicity, using HT - 29 cells (Human colon carcinoma) cells by 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay [9,10,11]. Briefly, the cultured HT - 29 cells (Human colon carcinoma) cells were harvested by trypsinization, pooled in a 15 ml tube. Then, the cells were plated at a density of 1×10^5 cells/ml cells/well (200 µL) into 96-well tissue culture plate in DMEM medium containing 10 % FBS and 1% antibiotic solution for 24-48 hour at 37°C. The wells were washed with sterile PBS and treated with various concentrations of the BJ sample in a serum free DMEM medium. Each BJ sample was replicated three times and the cells were incubated at 37°C in a humidified 5% CO2 incubator for 24 h. After the incubation period, MTT (20 µL of 5 mg/ml) was added into each well and the cells incubated for another 2-4 h until purple precipitates were clearly visible under an inverted microscope. Finally, the medium together with MTT (220 µL) were aspirated off the wells and washed with 1X PBS (200 µl). Furthermore, to dissolve formazan crystals, DMSO (100 µL) was added and the plate was shaken for 5 min. The absorbance for each well was measured at 570 nm using a micro plate reader

(Thermo Fisher Scientific, USA) and the percentage cell viability and IC50 value was calculated using GraphPad Prism 6.0 software (USA).

ETBr and AO Staining

This protocol was followed on line with [16,17].Briefly, 5 x 10^5 cells/ml of HT-29 (human colon carcinoma) cells were seeded into the 24 well tissue culture plate and treated with 92.09 µg/ml of BJ sample in a serum free DMEM medium. The plate was incubated at 37°C in 5% Co2 incubator for 24 hours. After incubation, 50 µl of 1 mg/ml acridine orange (AO) and ethidium bromide (ETBr) were added to the wells and mixed gently. Finally, the plate was centrifuged at 800 rpm for 2 minutes and evaluated immediately within an hour and examined at least 100 cells by fluorescence microscope using a fluorescent filter.

RESULTS AND DISSCUSSION

The intensity of dark blue colour developed by the MTT reaction is directly proportional to the number of viable cancer cells. In this test, we used different concentrations (Table 1) of samples to evaluate its effect on the viability of cancer cells. In the control reaction without any sample didn't showed any mortality on the cell population (Sl. No.1). In the reaction with the lowest sample concentration of 20µg / ml kept most of the cancer cells viable without showing any notable effect on cancer cells. In the reaction with the highest sample concentration of 200 µg/ml showed nearly 55% of mortality on cancer cells. This trend is visible in the dose response curve (Fig.1). The anticancer activity of the samples increases as the concentration of the sample increases. When the cells are photographed under the microscope the formazan crystals formed in the viable cell. The number of cells are more in HT-29 untreated cells than hill banana juice (BJ) treated samples (Fig. 3).

The cell viability reaction is directly proportional to the number of viable cancer cells. In this test, we used different concentrations (Table 2) of samples to evaluate its effect on the viability of cancer cells. In the control reaction without any sample didn't showed any mortality on the cell population. In the reaction with the lowest sample concentration of 20 µg/ml kept most of the cancer cells viable without showing any notable effect on cancer cells. In the reaction with the highest sample concentration of 200 µg/ml showed nearly 55% of mortality on cancer cells. The anticancer activity of the samples increases as the concentration of the sample increases is also well depicted in the dose response curve (Fig2).

The half maximal inhibitory concentration (IC50) is a measure of the effectiveness of a substance in inhibiting a specific biological or biochemical function. This quantitative measure indicates how much of a banana juice substance (inhibitor) is needed to inhibit a given biological process. In this analysis the hill banana juice (BJ) recorded IC50 as 92.05 µg/ml (Table. 3). The Fig 4a (Control) showed, the normal tumor cells (control). The Fig 4b exhibited early & late apoptotic cells, and necrotic cells under fluorescent microscopy (Fig 4b). Early-stage apoptotic cells were marked by a crescent-shaped or granular yellow-green acridine orange nuclear staining. Late-stage apoptotic cells were marked with concentrated and asymmetrically localized orange nuclear ethidium bromide staining. Necrotic cells increased in volume and showed uneven, orange-red fluorescence at their periphery.

Many fruits such as guava, banana, papaya, orange, lemon, apple, litchi possess proven medicinal activities as whole fruit, seeds, leaves, and as peels and many of them are reported to have anticancer potential such as lemon, orange, papaya, guava [10]. The well ripened banana fruit has better anti-cancer property and the dark spots on ripe yellow bananas produce a substance called Tumor Necrosis Factor (TNF) that destroys cancerous tumors [3]. The anticancer activity was carried out by MTT assay in banana, guava, orange and papaya showed good anticancer activity with IC_{50} values 31.7, 27, 95.5, and 18.5 µg /ml, respectively [17, 18]. The similarity in our present study results and the previous study data on banana fruits anti cancer effect in human cancer cells (Ht-29) confirms the efficiency of the

extract in inhibition the cancer cells growth. Further more study and careful experiments are needed to be done in the future to explore more pathways and mechanisms induced by this fruit extract.

CONCLUSION

Nowadays, finding a balance in cell proliferation expected to be a master key in homeostatic maintenance by inhibiting uncontrolled cells proliferation. Through arresting the cell cycle and induction of apoptosis in progressed cancer cells. An increasing number of evidences focused on the importance of phytochemicals and their effect in the metastasis of different cancers including colon cancer. At this study, we have illustrated that the banana juice extract inhibited the proliferation of the colon cancer cell line HT-29. We have investigated a cell death mechanism which inhibits the cell cancer growth.

ACKNOWLEDGMENTS

I would like to take this opportunity to express my heartfelt thanks to Director of ICAR – National Research Centre, Trichy, TN, India for extending all facilities to carry out the work.

REFERENCES

1. Goel RK, Gupta S, Shankar S , Sanyal AK. "Anti-ulcerogenic effect of banana powder (*Musa sapientum* var. paradisiaca) and its effect on mucosal resistance". J. Ethnopharmacol, 1986, 18 (1), 33-44.
2. Sachdewa A, Khemani LD . Effect of Hibiscus rosasinensis Linn.ethanol flower extract on blood glucose and lipid profile in streptozotocin induced diabetes in rats". J Ethnopharmacol, 2003, 89 (1), 61-66.
3. Iwasawa H and Yamazaki M . Differences in Biological Response Modifier-like Activities According to the Strain and Maturity of Bananas. Food Sci. Technol. Res., 2009,15 (3), 275 – 282,
4. Ferlay J, Dikshit R, Elser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F (2013) GLOBOCAN 2012 v1.0, Cancer incidence and mortality worldwide: IARC Cancer Base No. 11 [internet]. Lyon: International Agency for Research on Cancer. Available from: http://globocan.iarc.fr/Pages/burden_sel.aspx.
5. Bray F, Ren JS, Masuyer E, Ferlay J . Estimates of global cancer prevalence for 27 sites in the adult population in 2008. Int J Cancer. 2013, 132(5):1133–1145.
6. Center MM, Jemal A, Smith RA, Ward E . Worldwide variations in colorectal cancer. CA Cancer J Clin. 2009, 59:366–378.
7. Gordaliza M. Natural products as leads to anticancer drugs. *Clin. Transl. Oncol*. 2007, 9: 767–776.
8. Sampath Kumar KP, Debjit B, Duraivel S, Umadevi M. "Traditional and Medicinal Uses of Banana", Journal of Pharmacognosy and Phytochemistry. 2012, 1 (3): 57 -62.
9. Mosmann, T. "Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays". Journal of Immunological Methods.1983, ***65*** *(1–2): 55–63*
10. Chen KC, Peng CC, Chiu WT, Cheng YT, Huang GT, Hsieh CL,. Action mechanism and signal pathways of *Psidium guajava* L. aqueous extract in killing prostate cancer LNCaP cells**.** Nutr. Cancer. 2010;62:260–70.
11. Berridge MV, Tan AS. Characterisation of the cellular reduction of 3-(4,5-dimethylthiazol-2yl)-2,5-diphenyltetrazolium bromide (MTT): Subcellular localization, substrate dependence, and involvement of mitochondrial electron transport in MTT reduction. Archives Biochem Biophys 1993, 303:474-482
12. Lemaire, C., Andréau, K., Souvannavong, V., & Adam, A.. Inhibition of caspase activity induces a switch from apoptosis to necrosis. FEBS Letters.1998, 425(2), 266–270
13. Ligresti, A., Bisogno, T., Matias, I., De Petrocellis, L., Cascio, M. G., Cosenza, V.. . Possible endocannabinoid control of colorectal cancer growth. Gastroenterology. 2003, 125, 677–687.

14. Manosroi J, Dhumtanom P, Manosroi A. Anti-proliferative activity of essential oil extracted from Thai medicinal plants on KB and P388 cell lines. *Cancer Lett*. 2006, 235:114–20.
15. Baskic D, Popovic S, Pistic P, Arsenijevic N. Analysis of cycloheximide-induced apoptosis in human leukocytes: Fluorescence microscopy using V/propidium iodide versus acridine orange/ethidium bromide. Cell Biology International.2006, 30:924-32.
16. Brousseau P, Payette Y, Tryphonas H, Blakley B, Flipo D, Fournier M.. Assessment of cell viability: Determination of cell viability and cell concentration with ethidium bromide and acridine orange. In: Beudet M, Kouassi E, Lapierre P, Voccia I, editors. Manual of immunological methods. Boca Raton (FL): CRC Press LLC. 1999, 28-29.
17. Sun J, Chu YF, Wu X, Liu RH. Antioxidant and antiproliferative activities of common fruits. *J Agric Food Chem*. 2002 , 50:7449–54.
18. Dahham, S, Taleb A, Mohamad T, Yasser ,AM, Shah AM. Antioxidant activities and anticancer screening of extracts from banana fruit (*Musa sapientum*). Academic Journal of Cancer Research. 2015, 8. 28-34.

Table: 1. The OD values of the cell concentration at 570 nm

S. No.	Hill Banana Juice Concentration (µg/ml)	OD at 570 nm Mean
1.	Control	0.467
2.	200 µg/ml	0.177
3.	180 µg/ml	0.213
4.	160 µg/ml	0.240
5.	140 µg/ml	0.293
6.	120 µg/ml	0.300
7.	100 µg/ml	0.318
8.	80 µg/ml	0.337
9.	60 µg/ml	0.372
10.	40 µg/ml	0.417
11.	20 µg/ml	0.466

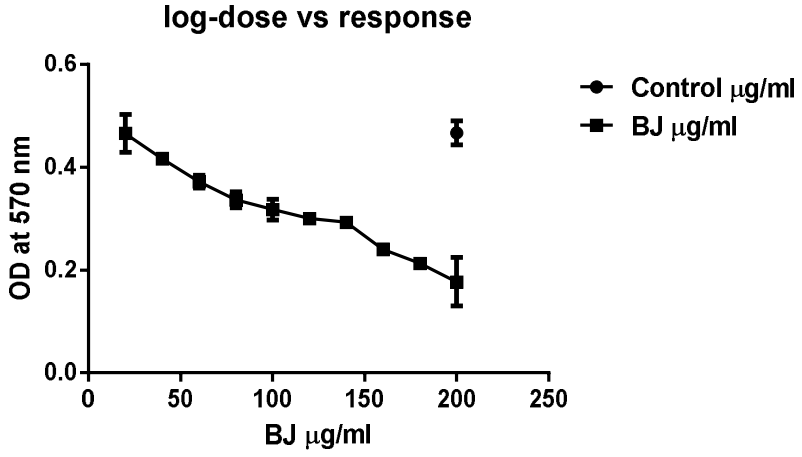


Fig.1: The dose response curve of bill banana juice against HT—29 cells

Table 2: Cell (HT-29) Viability (%) against the Hill banana juice treatment

S. No.	Hill Banana Juice Concentration (µg/ml)	Mean Value (%)
1.	Control	100
2.	200 µg/ml	37.89
3.	180 µg/ml	45.60
4.	160 µg/ml	51.45
5.	140 µg/ml	62.66
6.	120 µg/ml	64.30
7.	100 µg/ml	68.01
8.	80 µg/ml	72.16
9.	60 µg/ml	79.65
10.	40 µg/ml	89.21
11.	20 µg/ml	99.78

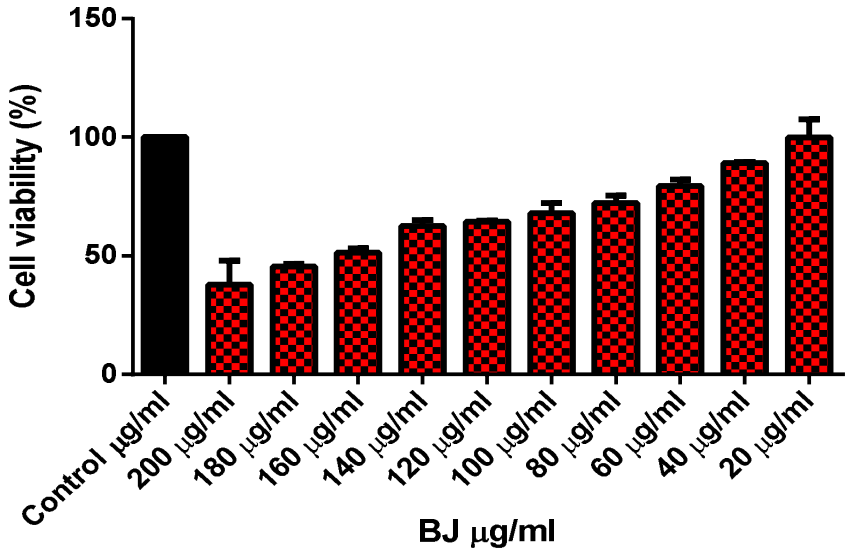
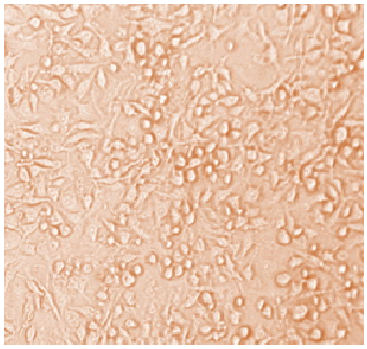


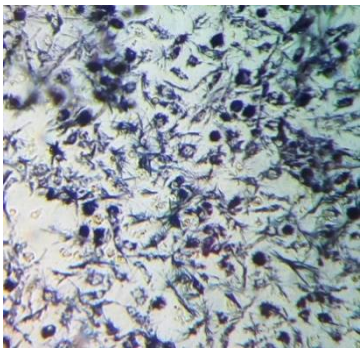
Fig-2: The cell (HT-29) viability percentage versus the hill Banana Juice (BJ)

Table 3: IC50 Value of tested BJ Sample: 92.09 µg/ml

log(inhibitor) vs. normalized response -- Variable slope	
Best-fit values	
LogIC50	1.964
HillSlope	-2.225
IC50	92.09
Std. Error	
LogIC50	0.01962
HillSlope	0.2370
95% Confidence Intervals	
LogIC50	1.924 to 2.004
HillSlope	-2.710 to -1.740
IC50	83.95 to 101.0
Goodness of Fit	
Degrees of Freedom	28
R square	0.9032
Absolute Sum of Squares	2630
Sy.x	9.691
Number of points	
Analyzed	30

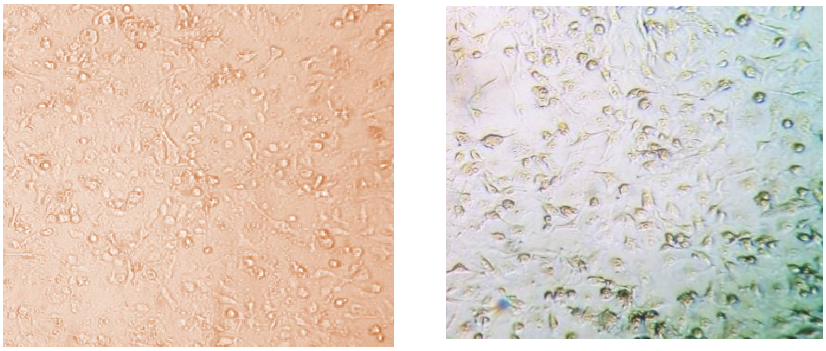


Before MTT treatment



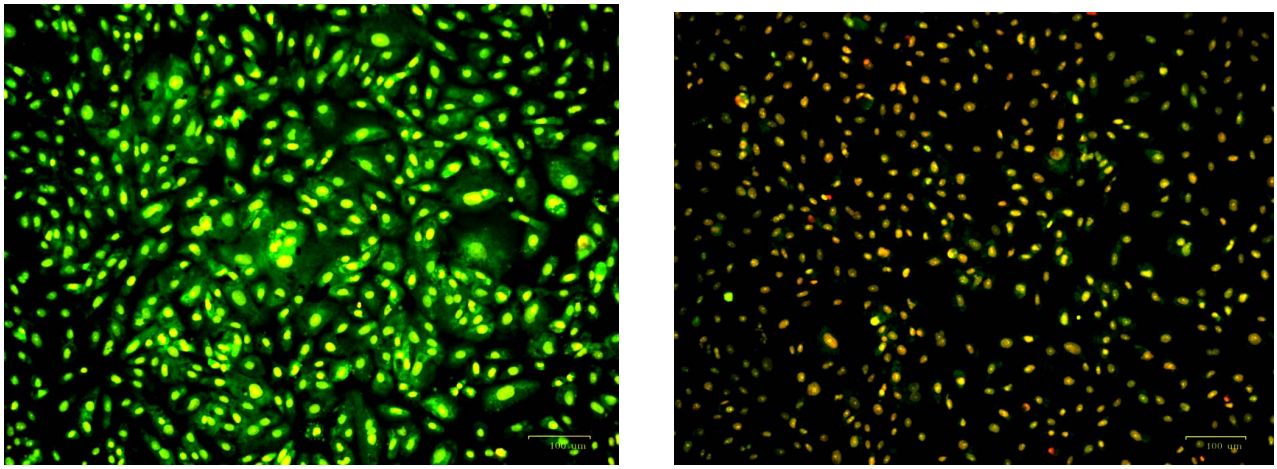
After MTT treatment

1. Control Cells (HT- 29)



BJ (200 μg/ml Treated Cells (HT-29)

Fig 3. Formation of formazan crystals in control cells and BJ Sample treated cells



4a. Control **4 b. Treated with 92.09 μg/ml of BJ sample**

Fig.4.ETBr assay with Hill Banana Juice (BJ)

Indian Journal of Natural Sciences www.tnsroindia.org.in ©IJONS

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Heuristic Study on Image Segmentation Algorithm in Image Processing

A.Rhagini[1*] and Mohana Priya.A[2]

[1]Assistant Professor, Department of Information Technology, M.Kumarasamy College of Engineering, Karur,TamilNadu,India
[2]Assistant Professor, Department of Computer Science and Engineering,M. Kumarasamy College of Engineering, Karur, TamilNadu,India.

Received: 10 Feb 2018 Revised: 21 Feb 2018 Accepted: 23 Mar 2018

[*]**Address for correspondence**
A.Rhagini
AssistantProfessor,
Department of Information Technology,
M.Kumarasamy College of Engineering,Karur,TamilNadu,India
E.mail: rhagini@gmail.com, mohanapriyaa.cse@mkce.ac.in

ABSTRACT

Image segmentation is one of the important method to divide the images into different multiple segments. It is used to convert image smooth and more effective. Main motive of image segmentation is to concentrate on particular part of image and make them more meaningful and powerful. This each method is used to improve the size, color, pixel, quality of images. There are many segmentation methods available. Each and every method has their unique activities. This survey is used to study the few existing segmentation algorithms and to know the advantage and disadvantage of those algorithms. Finally conclude that some hybrid techniques should be used to provide the most effective segmentation and to produce final result of high quality image.

Keywords: Image segmentation, high quality image, hybrid techniques.

INTRODUCTION

Image segmentation is one of the multiple problems available. It is the important process to analysis of image and image understanding. It is the process of dividing the images into different sub images of homogeneity. Segmenting the image is to cluster pixels into several region to individual surfaces, objects or every parts of images.(S. Thilagamani and N. Shanthi 2011).Image segmentation is very useful in many areas. It is very much efficiently used in medical area for diagnosis the cells, tissues, bones and each and every parts of the body. In research area it is used in image processing field. Image is differentiated into foreground image and background image using the image segmentation, whereas foreground is based on region of interest and background is the remaining image. There are three approaches for image segmentation available they are Threshold, Edge, and Region based (Muhammad Waseem Khan 2014).

IMAGE PROCESSING

Image processing is a method to improve the quality of sample images which was taken in many several applications. Image processing is used in various applications some of them are:

- Textiles
- Medical Study
- Military
- Research area
- Graphic arts
- Film industry

A..Methods of Image Processing:

There are two methods available in image processing. They are

1.Analog Image Processing

By using the electrical means we will alter the images in Analog Image Processing. The most common example is television image. The brightness of the image will be improved by varying the amplitude in the television signal which is at voltage level. The brightness and contrast on a TV is set to adjust the reference of video signal and amplitude, which results in the brightness, darkness and alteration of displayed image.(K.M.M. Rao)

2.Digital Image Processing

Digitizer is used for converting the image to digital and then the image will be processed.In digital system the two-dimensional picture will be processed by the method of Digital image processing. Bits represented in an array of real numbers are the digital image. Repeatability, versatility and original data precision will be preserved are the main advantage of digital image processing.(K.M.M. Rao)

Image processing has many techniques they are:

- Image representation
- Image preprocessing
- Image analysis
- Image reconstruction
- Image data compression
- Image enhancement
- Image restoration

ARCHITECTURE

The overview of image segmentation architecture in static frame will be having the following several steps. They are

Input the image
Preprocessing of image
Forming super pixel of image
Segmenting objects
Final Object output

Input the image: First the image from which the object is sent as a input to the framework.

Preprocessing of image: Image is send into the framework for preprocessing.
Forming super pixel of image: Several pixels of images are grouped to form a super pixels
Segmenting objects: Input image is segmented into several sub images.
Final Object output: In static frame, final objects are segmented and obtained.

LITERATURE REVIEW OF IMAGE SEGMENTATION TECHNIQUES

A. Edge Based Image Segmentation

The spectral method is used for segmentation of new image using edge and region based details. In this algorithm initially the noise from the image is reduced by using bilateral filter. Multi-class Normalized cut method has been used for group based region grouping and region merging is also used. It is found that the proposed method have some outperformed other methods and produce good results(Muhammad Waseem Khan 2014).

B. Fuzzy Theory Based Image Segmentation

This algorithm uses morphological opening and closing operations on the final output result. It solves the problem of watershed algorithm. It is used to save the information of image and increase the speed also. Fuzzy object model is applied on output image. It is evaluated on the basis of Fast Positive Volume Fraction(FPVF) and Fast Negative Volume Fraction(FNVF).The new Fuzzy Rule based image segmentation technique is much useful to segment even the thin segment images. The new method of image segmentation using Fuzzy Rule based system and Graph Cuts. In this new algorithm first images are segmented with gray scale, color and texture of images by using Graph Cuts. By using fuzzy rule, weights are assigned to all the features of image. It extract the features of image, then find the constant by using fuzzy rules, find the weight average of constants to calculate the similarity matrix using Normalized Graph Cut method(Muhammad Waseem Khan 2014).

C. Artificial Neural Network (ANN) Based Image

It based on texture features and neural network to separate targeted image from the background. Here micro-CT images dataset are used. De-noising filters are used to eliminate the noise from the pre-processing method then further back propagation Neural Network is crated and finally it modifies the weight of the number of network and save the output. Further new algorithm is compared with the Thresholding method and Region Growing method. This are used to improve the speed and accuracy of segmentation. Neural network algorithm is used to segment the image based on the color of the images. Image Texture Classification technique is based on the Artificial Neural Networks (ANN).In this algorithm, first the image is captured and it is pre-processed, then feature extraction is done by ANN classifier is used for texture classification. Trained ANN combines the input pixels into two clusters which finally gives a result (Muhammad Waseem Khan 2014).

CONCLUSION

This survey on segmentation gives a clear idea about the various segmentation techniques. Each individual algorithm has its own usage in research area. This algorithm is used to reduce the complexities in better way. The need for image segmentation is recognized to produce a improved and effective result. So it is necessary to have a hybrid algorithm which combines the features of more than three algorithms together to produce a effective result.

REFERENCES

1. S. Thilagamani and N. Shanthi," A Survey on Image Segmentation Through Clustering", International Journal of Research and Reviews in Information Sciences, Vol. 1, No. 1, March 2011.
2. Muhammad Waseem Khan "A Survey: Image Segmentation Techniques" International Journal of Future Computer and Communication, Vol. 3, No. 2, April 2014.
3. Waseem Khan" Image Segmentation Techniques: A Survey" Journal of Image and Graphics Vol. 1, No. 4, December 2013.
4. S.Thilagamani and Dr. N.Shanthi" Innovative Methodology for Segmenting the Object from a Static Frame" International Journal of Engineering and Innovative Technology (IJEIT),Volume 2, Issue 8, February 2013.
5. S. Thilagamani and N. Shanthi" Object Recognition Based on Image Segmentation and Clustering" Journal of Computer Science 7

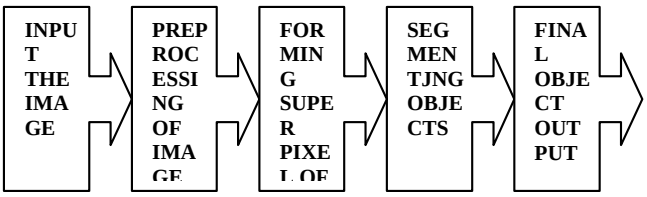


Fig. 1: Architecture of Image Segmentation in Static Frame

Indian Journal of Natural Sciences *www.tnsroindia.org.in ©IJONS*

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Classification of Sentimental Reviews using MPSO Algorithm

Sharmila D[1]* and N.Selvaganesh[2]

[1]Assistant Professor, M.Kumarasamy College of Engineering,Karur,TamilNadu,India
[2]Assistant Professor, PSNA College of Engineering and Technology, Dindigul, TamilNadu,India

Received: 13 Feb 2018 Revised: 25 Feb 2018 Accepted: 23 Mar 2018

***Address for correspondence**
Sharmila D
Assistant Professor,
M.Kumarasamy College of Engineering,
Karur,TamilNadu,India
E.mail: sharmi.draj@gmail.com / selva492@gmail.com

ABSTRACT

The number of reviews and blogs for movies increase every day with the increase in number of social networking sites, which can be positive, negative or neutral. These unstructured reviews are considered as an important source of analyzing the movie and they are classified or clustered to provide useful information. They are classified with the help of supervised machine-learning algorithms like SVM, Naïve Bayes classification, etc. In this paper, the reviews in IMDB are classified using MPSO algorithm and the accuracy is calculated using precision, recall and f-measure. The accuracy of classification increased from 84% (existing system) to 86.38% in the proposed system.

Keywords: Sentimental reviews, MPSO algorithm, IMDB dataset.

INTRODUCTION

Opinion mining (sentimental analysis) analyzes people views and their emotions towards products, organizations, and their associated attributes. Now-a-days, people gather information about a product from various sources like reviews, comments and blogs in social media. Machine learning is used by experts to extract useful information from people opinion, (Liu, et al). Sentiment analysis can be done in document (or) sentence (or) aspect level (Feldman, et al). The option of the document is classified as positive, negative or neutral by the document level analysis. The expression of opinion of a sentence is classified as positive, negative or neutral by the sentence level analysis. Expressions of sentiments in document and the aspect to which it refers are given more importance in aspect level sentimental classification. Supervised and unsupervised learning are the two major machine learning techniques used for document level sentimental analysis. In supervised learning, the labeled datasets are trained to obtain a useful output which for decision making (Gautam, et al). In Unsupervised learning, clustering algorithms are used to solve the problems in unlabeled data (Hastie, et al). The proposed paper deals with supervised learning. Pre-processing is done by removing the stop words and digits from the movie reviews for further analysis, as they are

mostly in text and unstructured format. These reviews are changed into vector of numerical matrices which are given as input to MPSO algorithm for classification.

RELATED WORKS

In Document level arrangement, the part of supposition characterization in view of order examine, with positive and negative estimations as considered (Pang, et al). Experiments were conducted using Naïve-bayes, Scalable Vector machine and ME machine learning algorithms. Individual and combination of n-grams like unigram, bigram etc were used along with a framework of bag of words, to classify the reviews. In that experiment, SVM algorithm classified more reviews correctly, where the NB algorithm produced very low accuracy of classification. More derived features were included in the model and labeled data were tested to estimate relative influence. Turney's approach was used to generate a new labeled set of corpus from the old model (Turney, et al). SVM algorithm was utilized by Mullen and Collier for characterizing the audits, in which esteems were allocated to choose words and afterward consolidated to shape a model for grouping of surveys.

The syntactic relation among words was considered as a basic of document level analysis, by Matsumoto et.al (Matsumoto, et al).In that approach, the succession of continuous words and reliance sub tree were separated from sentences and utilized as highlights for SVM calculation. The unigram, bigram n-grams, word subsequence and reliance sub tree were separated frame each sentence in the IMDb and extremity datasets (Pang, et al). The preparation and testing datasets were given independently in IMDb dataset, though in Polarity dataset a 10-overlay cross approval procedure was utilized for arrangement since no different information was assigned for testing or preparing.

Liu and Chen proposed a multi-name grouping on opinion order (Liu, et al) where eleven multilevel arrangement techniques were thought about utilizing datasets of two online journals and their execution was assessed utilizing eight measurements. Three diverse opinion word references were utilized for multi-level characterization. The multi-mark order process was performed in two stages to be specific issue change and calculation adaption (Zhang, et al). In issue change, the issue was changed into different single name issues. The framework gains from these changed single name information amid preparing stage and in the testing stage, forecast at single labels are made by learned classifiers at that point makes an interpretation of it to numerous names. In calculation adaption, the information is changed according to the prerequisite of the calculation.

Luo et.al, suggested that content information can be changed over into low measurement passionate space, for which they utilized little estimated commented on words with unmistakable and clear importance. Ekman Paul's examination was utilized to group the words into six fundamental classifications in particular astonishment, outrage, fear, disturb, bliss and tragic. Two diverse methodologies were utilized for doling out weight to words with passionate labels. The aggregate weight of passionate labels were figured, in view of which grouping was wear which gave a decent outcome for stock message board. It can be connected in any dataset or space.

From these studies, the following aspects are taken into consideration for further research. First is, unigram approach was used for classification by most of the authors which provide comparatively better result, but failed in some cases (Pang, et al). When unigram approach is used to analyze the sentence "The item is not good", provides the view of sentence as neutral since it has a positive word (good) and a negative word (not). When bigram approach was used to analyze the same sentence, it gives the view of sentence as negative as it has the phrase 'not good', which is correct. The result of classification improves when a higher level of n-gram is considered. Secondly, POS tags (Part-Of-Speech) were used for classification by a number of authors. But , the POS tag of a word changes as per the context in which the word is used. For example the word 'ticket' can be tagged as 'noun' in the sentence 'I booked a ticket' and it can be tagged as 'verb' in the sentence 'ticket booking is easy'. The entire word can be used as a parameter for classification, instead of using the POS tag of that word. Third, data is represented as a matrix of

numbers in most of the machine learning algorithms, though the sentiment data is in text format. So the data should be converted to number matrix in order to give it as input. Most commonly used method to convert text to number matrix is to use TF or TF-IDF. In the proposed approach, TF-IDF is used for conversion by representing a text file as rows of the matrix and each word in the file is represented in the column of the matrix.

METHODOLOGY

In the proposed framework, Modified Particle Swarm Optimization calculation was utilized for the arrangement of surveys by client. Arrangement of estimations is of two kinds: double opinion characterization and multi class conclusion order. In double grouping write, each report d in D, where D = {d1, ..., dn} is named positive or negative. In multi-class opinion investigation, a record'd' is named solid positive, positive negative and solid negative utilizing the grouping name 'C'. Binary classification method was used for classification by some authors. The dataset is a collection of reviews in text format where as numerical matrices are being used by machine learning algorithms for classification.

The modules in the proposed system are:

1. Pre-processing.
2. Vectorization
3. Classification using MPSO algorithm

Pre-processing

In this module, Noise removal (removal of mark-up tags, non-words, digits and normalizing the white space) is done. After which each document is represented as a long string, with one space between each token. This process is done for all the documents in the dataset.

Vectorization

The features used in the proposed system are as listed in the table 1:

Average Sentence Length(ASL) of a document d is calculated by dividing the total number of words in d divided by the total number of sentences in it. It is used to improve the accuracy of classification.

$$ASL = \frac{Total\ number\ of\ words\ in\ a\ document\ X}{Total\ number\ of\ sentences\ in\ the\ document\ X} \quad (1)$$

*Construction of vector:*After the extraction of these features, a feature vector is constructed for each document.. for example, now the feature vector (Φ) is calculated for feature-category "unigam+bigram" word god (Halvani, et al).

$$\Phi = \frac{No.of\ occurences\ of\ "god"\ in\ T}{Length(T)-n+1} \quad (2)$$

Where T is the document considered and n =3, the same procedure is repeated for all features.
After the construction of vector, the formula below is used to calculate the Manhattan distance:

$$Dist\ (X,Y) = \sum_{i=1}^{n} |xi - yi| \quad (3)$$

Where X and Y are documents. The result of Manhattan distance is converted into the similarity score using the formula (Halvani, et al) in equation (4):

$$sim(X, Y) = 1 / (1+dist(X,Y)) \quad (4)$$

Classification using MPSO algorithm

Before starting classification of reviews, parameter optimization is done to select the parameters with best values and high accuracies compared to the other parameters in each feature (Induja, et al). The above steps are repeated for each feature. The accuracies of all possible parameter of each features is obtained and then for each feature, the parameter value that leads to maximum accuracy is obtained and stored as the model M.

In any PSO algorithms, the fitness function plays a vital role in classification. The fitness function is derived for the accuracy and then particle position and velocity are initiated after which the updated positions and velocity of each particle is calculated (Lakshmi, et al).The general PSO algorithm consists of a cognitive component. In MPSO algorithm, the cognitive component is divided into two components. They are good experience component and bad experience component. Using the first component, the particle has memory of its previously visited best position. The bad experience segment encourages the particle to have memory about its already visited worst positions. For the computation of speed of the molecule, the terrible experience segment is likewise thought about (Deepa, et al).

$$V_{i+1} = \omega * V_i + c_1 * R1 (p_{best} - S_i) + c_2 * R_2(S_i - p_{worsti}) + c_3 * R3(g_{best\,i} - S_i) \quad (5)$$

The position update equation is same as general PSO algorithm:

$$Si+1 = Si + Vi+1 \quad (6)$$

Where ω, c_1, c_2, c_3 are the inertia weight and acceleration coefficients. c_1, c_2, accelerates the particle towards its best position, while c_3 accelerates the particle away from its worst position; p_{worst} is the worst position of the particle. R1, R2, R3 are uniformly distributed random numbers in the range (0, 1).

The steps in Modified PSO algorithm (Deepa, et al) are:

Step 1 Select the acceleration coefficients c_1, c_2, and c_3 random variables R1, R2, R3, number of particles and generations to start optimal solution searching.

Step 2 The velocity and position of each particle are initialized.

Step 3 The particle's individual best position and velocity value is chosen for each generation is selected.

Step 4 Select the particle's global best value, i.e particle near the target among all the particles, is obtained by comparing all the individual best values.

Step 5 Select the particle's individual worst value, i.e. particle too away from the target.

Step 6 Update the particle individual best p_{best}, global best g_{best}, particle worst P_{worst} in the velocity equation and obtain new velocity.

Step 7 Update the new velocity in Eq (3.7) and obtain the position of the particle.

Step 8 The steps are repeated for pre defined number of generations (or) till the expected accuracy is achieved.

In the proposed method, the number of particles and generations is fixed as twenty and 10 to 100. Each particle is a combination of randomly selected features represented as 1 if the feature is present and 0 if it is not present in the particle. For example 101010100 is a particle which consists of the features punctuation n-gram, n frequent tokes, Token k-suffixes and Token k-suffix n-grams. The values of ω, $c_{1, c2, c3,}$ R1, R2, R3can be fixed between 0 to 2.

The initial position and velocity of each particle are assigned randomly in vector form, for example 001010100. For the first generation, both the p_{best} and also $p_{worst}$of each particle is considered. The p_{best} is the value of the particle which has the highest fitness value among all the particles is considered as the g_{best} of the generation. The updated velocity and position of each particle are obtained by substituting these values in the equations 3.5 and 3.6(Deepa, et al).

For the second generation, the initial position and the updated position value of each particle are considered, the position with highest fitness value is considered as p_{best} and position with lowest fitness value is considered as the p_{worst} of the particle. If the g_{best} value of second generation is higher than the g_{best} of first generation, then the g_{best} of the first generation is used to obtain the new position and velocity of each particle, else the new g_{best} will be used. The same procedure is repeated till the expected accuracy is achieved.

RESULTS AND DISCUSSION

Dataset used: Internet Movie Database (IMDb 2011) is used for the proposed method of text classification. The dataset has 12000 positive and negative reviews for testing and training the classification model. It also has 40000 unlabeled reviews.

The results obtained using proposed method of classification is compared with that of the existing system which follows n-gram classification approach and they are listed in the table 2. Machine learning algorithms like Naïve Bayes, SVM method using n-gram were used early.

In the proposed paper, MPSO algorithm was used for classification using the punctuation n-grams, unigram, bigram, unigram + bigram, character n-grams, n % frequent tokens, token k-prefixes, token k-suffixes and Average Sentence Length (ASL). The results were measured using accuracy, F-measure, precision and recall. The proposed model performs better than the existing system of classification.

Precision: It is the ratio reviews classified as true positive to the sum of number of documents classified as true positive and true negative.

$$\text{Precision} = \frac{tp}{tp+fp} \quad (7)$$

Recall: It is the ratio of number of reviews classified as true positive to the sum of number of reviews classified as true positive and false negative and used to measure the completeness of the classification.

$$\text{Recall} = \frac{tp}{tp+fn} \quad (8)$$

F-Measure: It can be defined as the mean of equations 7 and 8, used to optimize the system as either precision or recall, which has more impact in the final result.

$$\text{F-measure} = \frac{2*Precision*Recall}{Precision+Recall} \quad (9)$$

Accuracy: It is the ratio of number of movie reviews classified correctly to the total number of reviews in the IMDb dataset used.

$$\text{Accuracy} = \frac{numberofcorrectlyclassifiedreviews}{Totalnumberofreviewsinthecorpus} \quad (10)$$

Twenty randomly generated particles with 10 generations were used in the MPSO process. Multiple trails were carried out and the best performance achieved was 86.38% in the seventh generation. The proposed method has a low time complexity compared to that of the existing system.

CONCLUSION AND FUTURE WORK

The objective of Sentiment classification is to classify texts according to the sentimental polarities of opinions it has. It will also be helpful in recommender systems and business intelligence applications where the user input and feedback can be summarized easily.

In the existing system, the text classification was done mainly using features based on n-grams, where the accuracy of single features where more effective compared to the combination of features. The execution time was long as many features had to be derived individually.In the proposed method; classification task is done using the model constructed using MPSO algorithm. Each particle is assigned with a random velocity and position which were updated in each generation. In this algorithm all individual particles will have the memory of their worst and best positions. Low time complexity is the main advantage of the proposed method. The time taken for constructing the classification model and task of verifying the reviews by the proposed method is low, compared to that of the existing system. Average Sentence Length, a lexical feature is added to improve the accuracy of the classification. The proposed classification model was tested IMDb 2011 dataset. The accuracy of classification in proposed model was improved in the seventh generation itself.

In future, works can be carried out to improve the accuracy of classification using more features along with the lexical and character features in the proposed model. Trails can be carried out in fixing constant values of inertia co-efficient and acceleration co-efficient instead of random values.

REFERENCES

1. Feldman, Ronen, Techniques and applications for sentiment analysis, Communications of the ACM 56.4, 82-89 (2013).
2. Gautam, Geetika and Divakar Yadav, Sentiment analysis of twitter data using machine learning approaches and semantic analysis, Contemporary computing (IC3), In the seventh international conference on. IEEE (2014).
3. Halvani O, Winter C, Pflug A, Authorship Verification for Different Languages, Genres and Topics, In Proceedings of DFRWS Europe, vol.16, pp. 33-43 (2016).
4. Hastie, Trevor, Robert Tibshirani, and Jerome Friedman, Unsupervised learning the elements of statistical learning, Springer, New York, NY, 485-585(2009).
5. Induja, M.S.Kamalesh and N.Senthilnathan, Fuel Cost Minimization for a thermal power plant using PSO Algorithm, In the special issue on Innovations in information Embedded and Communication Systems,Pp. 81 - 84 (2016).
7. Lakshmi, A. Jaya, and Vuyyuru Krishna Reddy, An Efficient resource Allocation Strategy based on IMPROVED PARTICLE SWARM OPTIMIZATION (IPSO), In the International Conference on Engineering and Technology Systems (ICET'16) Pp. 125 - 129 (2016).
8. Liu, Bing, Sentiment analysis and opinion mining, Synthesis lectures on human language technologies, 5.1, 1-167 (2012).
9. Liu, Shuhua Monica, and Jiun-Hung Chen, A multi-label classification based approach for sentiment classificatio Expert Systems with Applications 42.3: 1083-1093(2015).
10. Matsumoto, Shotaro, Hiroya Takamura, and Manabu Okumura, Sentiment classification using word sub-sequences and dependency sub-trees, Pacific-Asia Conference on Knowledge Discovery and Data Mining. Springer, Berlin, Heidelberg (2005).

11. Pang, Bo, Lillian Lee, and Shivakumar Vaithyanathan, Thumbs up? sentiment classification using machine learning techniques, In the Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10. Association for Computational Linguistics (2002).
12. Turney, Peter D, Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews, In the Proceedings of the 40th annual meeting on association for computational linguistics, Association for Computational Linguistics(2002).
13. S N Deepa, G.Sugumaran , Model Order Formulation of a Multivariable Discrete System Using a Modified Particle Swarm Optimization Approach, In Proceedings of Swarm and Evolutionary computation,vol.1, pp. 204-212 (2011).
14. Zhang, Dongwen, et al, Chinese comments sentiment classification based on word2vec and SVM perf, Expert Systems with Applications 42.4 1857-1863 (2015).

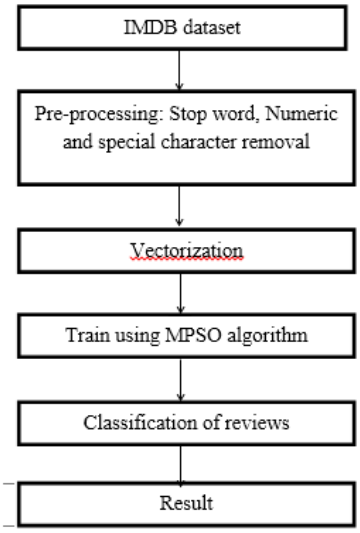


Figure 1: System Architecture of the proposed system.

Table 1: List of features used in the proposed method

Features	**Description**
punctuation n-grams	A sequence of n punctuation marks taken from training document after reduction to punctuation characters
Unigram	N -gram of size 1
Bigram	N -gram of size 2
Unigram + Bigram	Combination of unigram and bigram
character n-grams	A sequence of n consecutive characters in training document
n % frequent tokens	The n % most frequently occurring tokens in training document
token k-prefixes	The first k characters of a token
token k-suffixes	The last k characters of a token
Average Sentence Length	total no. of words in a document / total number of sentences

Table 2: Comparison of results obtained using the existing and proposed methods

Classification method	Features	Accuracy of classification
Ensemble model using MPSO algorithm	1,3,4,10	86.38%
n-gram machine learning approach	4,5	84%

Indian Journal of Natural Sciences www.tnsroindia.org.in ©IJONS

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

REVIEW ARTICLE

A Review on Hybrid Cloud as Big Data and Analytics Solutions

B.Padmini Devi[1]* and S.Kanimozhi[2]

[1]Associate Professor, Department of IT,M.Kumarasamy College of Engineering,Karur,TamilNadu,India
[2]Assistant Professor, Department of IT,M.Kumarasamy College of Engineering ,Karur, TamilNadu,India

Received: 12 Feb 2018 Revised: 23 Feb 2018 Accepted: 23 Mar 2018

*Address for correspondence
Dr.B.Padmini Devi
Associate Professor, Department of IT
M.Kumarasamy College of Engineering,
Karur,TamilNadu,India
E.mail: padminidevib.it@mkce.ac.in/kanimozhis.it@mkce.ac.in

ABSTRACT

In a emerging world, the IT corporate growing more and more towards technology. most of the business application belongs to cloud ,it has more number of service like Iaas, Paas,Saas,Testing as a service(Taas),Security as a service (SecaaS).cloud provide the less expensive service to all the user. Cloud has 3 types, such as private, public, and hybrid cloud, organization expecting enlarging storage capacity to store the data, and protecting data also. Because of this purpose, new idea has been furnished. Integration of hybrid cloud and big data will provide the analytical solution.

Keywords: Cloud.Hybrid cloud,Big data,analytics solution,loads, business application.

INTRODUCTION

IT corporate and business application,day by day rapid changes will occur.they mainly viewed on – or minimum level of consideration – either transferring existent load to cloud cloud, boost existent loads, or constructing new load and combining new load with previous load.Recurrently Many organization,need for security ,because of security issues in cloud cloud some organization doubting about shifting to the cloud. the new service has been launched in cloud as Security As A Service.(SecAAS)[7].since no problem with the cloud.Now a days, collection of data has been enormous.user and many business developer need of very large storage media.the variety of data collected and used by businesses developer.managing data must adopt to all level(local&global), they expecting to protecting data and service offered by cloud should be low. Favorably, business application for cloud technology is growing and it is to allow many businesses application to set up hybrid cloud, specifically the field is big data and analytics.

A hybrid cloud is a integration of private and public cloud.composite of the public and private cloud with committed refers to **"private cloud".** The public cloud and private are systematic so that they are operate separnately.but they exchange the information with each other through an encrypted system of connection on a private and/or public system of connection using technologies that assist the progress of ability to move the applications and data.A hybrid cloud grant organization to composite individual and/ or secret information from the private cloud plus applications

operating on the public cloud, while influencing the public cloud's computational resources and storage. For example,organization could create triable awareness by combining the data from system of Archieve (private cloud) with system of bond in a public cloud or by execute edge-analytics on the devices in the public cloud.More over ,hybrid cloud access scalability by granting organization to use public cloud resources for station where the private cloud having lack of computational power. Furthermore, containers could be used to increase the movability of loads between private cloud and public cloud. Finally, hybrid cloud is a model for universal allocation of applications and data, granting more excellent management of data domination and agreement.

Hybrid cloud

A hybrid cloud is association of the private cloud and more than one public cloud as exposed.it is shown in picture 1. It influences good cloud to be offer, and contributing the adaptability to data, place and services based on developer needs. Data will be placed and a admission based on analytical load provisions and consuming design within that hybrid cloud. Delivering data and analytics to the various personas whenever they expected or wanted. Approach for every fields of the hybrid cloud can be controlled, trained to maintain private data, security,and all other information[8] .

The digital conversion needs a new hybrid cloud — in this expose, new hybrid cloud will be a flexile by design,.it provides clients feel free to select, modify cloud information and services in which the client expected. The system grants cloud services, application to run fastly,and it can constructed by good information and vision free,cultivating good clarity, governance, security and integrated control, in all the place

As exposed in picture 2, the most of the system of Archieve usually stay on the private cloud, at the time of system of bond and system of Automation are mainly in the public cloud and the system of understanding overall condition of Hybrid.The hybrid cloud property is understanding and flexibleness.It will grants collaborated analytical load to be stored and data whenever it generate the many function as per the business needs. Information secrecy, protections are managed.It will be driven constantly all over the system of the hybrid cloud

The hybrid cloud basic terminology is depend on the maximum client whoever need to upgrade from private cloud to the public cloud. Trust private cloud is very important because many business application will consistently need section of their infrastructure and data to survive back of company firewall because of rules and regulation of industry.This will makes flexile architecture by providing application many privilage to change and select their conditions ,applications and delivere the services more fastly.Generally user can see hybrid cloud as an global design solution. It is unmovement path. Objective of cloud is grant to boost loads to hybrid cloud from original private cloud it will combine private and public clouds. This approach can be used in industry; still it can assist to achieve the integration.

Hybrid cloud for big data and analytics

A hybrid cloud admit various characters to work along with analytics and data qualification .it provide easy platform to the client to perform and aid is to explain provisions whenever data skills and analytics stored at the hybrid cloud .Outcome of this is, analytics loads can run more powerful anywhere the data is reserved.The very important thing is to have hybrid as a choice, as locality must be one another first architectural decisions for any analytics project. Specifically organization have to examine the processing of analytical can be located correspondent to the data. Where the data should be stored Concurrently, regulatory and legal provisions also implicate the data could be located..

System is moving to hold many centers of force that will command whenever the user do processing. For example, construction of a data pool as one of part of a system of Insight.The data pool consists of data.This will be in the

private and the force of mid point will be on the private cloud .The data processing should be interior of the private cloud. Yet the system of Insight begins the creation of data on the public cloud then there could be a second force of mid point

Primary drivers needed for big data and analytics in the hybrid cloud

Integration: Organization want to combine data is to saved and managed Usually, these organization want to combined system of bond and/ or system of Automation (IOT) applications, like customer management system, devices, with system of Insight ,social media, and predicted analytics and real-time analytics hosted on public clouds.

Brokerage/management for load and resource optimization: multiple loads gets multiple provisions for resources, security storage and time to process [8]. Many business industry directed to hybrid cloud,as a result of, The company need to store the analytical data load whenever it meets the client expectation at higher levels. These organization need agility and less expense , at the same time enjoying the adaptability to transfer the analytical data loads from private to public cloud

Portability: One More option for hybrid cloud is to assure flexibility of analytical data loads To manage expenses and performance, organization need to transfer data loads to good platform that meets dynamic client request. This ability to perform organization to think over the workability of the new analytical data load on a particular hybrid cloud .

Compliance: A hybrid allows to delivering all applications, data loads overall. Initially the company / industry has to get agreement for all regular activities afterthat only company will display the cloud . the country of the particular company / industry is full reaposible for all the activities done by the company.

The key points for executing a hybrid cloud strategy include:

Cultural shift: The main point is to demanding and shifting from private to a hybrid cloud is implementing,contributing a service-oriented approach to providing analytics data and self-service abilities to expand private cloud to the public cloud.

Varying levels of hybrid sophistication: A hybrid cloud approach having various levels of styles: deep integration,(private cloud and public cloud), or simple, and constant SPoC(single point of connections) using a virtual private network (VPN), a secure gateway, and an API manager constructed to explore system of Archieve (private cloud) to system of bond(public cloud).

Primary Key factors are

1. Integration of System of bond(SOB)
2. System of Automation (SOAu) with system of Insights (SOI)
3. System of Archieves (SOA).

This is the primary case for expanding private data to the public cloud. Organization want the full ability and adaptability to construct personalized customer offers and response to market dynamically. it requires a new type of architecture, such as an enterprise data pool; that is, a collection of fit-for-purpose repositories that are well managed, governed, protected, connected by metadata and provide self-service access2. How a typical enterprise data pool may fit in the hybrid cloud strategy landscape, including different IT domains such as system of Archieve, system of bond, system of Automation, and system of Insight? It depends on the data gravity of the organization. If the data gravity stays inside of the private cloud, then the data pool will be defined inside of the private cloud. If the

data gravity is shared between the private cloud and public cloud cloud, then the data pool will be defined across both private cloud and public cloud cloud. If the data gravity stays inside of the public cloud, then the data pool will be defined on the public cloud.

Some of the key topics that need to be considered when planning a hybrid cloud strategy are:

Analytics Data — it provide the place to store the data based on the analytical load and the type of access.
Data movement and replication — it also provide the place to store the data to stay away from the data movement and replication.
Data preparation and integration — which will give the best cloud to do data preparation/integration
Data sovereignty and compliance — this will give the place to store the data based on data sovereignty provisions and it will also provide required compliance regulations for this process
Data governance and security — how to secure the data over all and whichever access control wants to be in fuction?
High availability and disaster recovery — are we want high availability and disaster recovery for business application in the cloud?
Network configuration and latency — need to alter the network configuration to comfort the latency provisions of an application in the cloud.

- **Portability** — application and associated data want to be compact to variety of .
- **Scalability** — application need ability to scale up and/or down
- **Resource orchestration** — it will define how much resource orchestration needed for all loads.

CONCLUSION

In this survey paper , a hybrid cloud is a integration of on –premises and local cloud joined with one or more dedicated cloud(s) and one or more public cloud(s).composite of the on-premises & local cloud with the dedicated cloud(s) refers to **"private cloud".** A hybrid cloud admit different characters to work with data and analytics capabilities where it makes the maximum sense for them to do so and this helps to describe the provisions whenever the data and analytics skills should be placed/available in the hybrid cloud.As a result, analytics loads could run more powerful wherever the data is stored. This must need the Primary Key factors are :-1. Integration of System of Bond(SOB) 2. System of Automation (SOAu) with system of Insights (SOI) ,3. System of Achieves (SOA).

REFERENCES

1. C. Arun and K. Prabu, Architecture Of Mobile Cloud Computing,Pak. J. Biotechnol. Vol. 14 (Special Issue II) Pp. 301-304 (2017)
2. Deepak Dagar,Amit Gupta, New paradigm in Conventional Software Testing: Cloud TestingJerry Gao 1,2, Xiaoying Bai2, and Wei - Tek Tsai2,3 2011, Cloud Testing - Issues, Challenges, Needs andPractice, Software Engineering : An International Journal (SEIJ), Vol. 1, No. 1, SEPTEMBER 2011
3. P. Harish, S.Vigneshwari, K.B.S. Ravi Teja,,Enhancing The Security Of Cloud Storage For Medical Data Retrieval Using Double Encryption With Data Anonymization,, Pak. J. Biotechnol. Vol. 14 (Special Issue II) Pp. 75- 78 (2017)
4. Padmaja, K. and R. Seshadri,,A Review On Cloud Computing Technologies And Security Issues, Pak. J. Biotechnol. Vol. 13 (special issue on Innovations in information Embedded and communication Systems) Pp. 503-507 (2016)
5. B.Padmini Devi, V.Dhanapandi, Delta Encoding For Deduplication Of Hybrid Clouds, in IJAER (International Journal of Applied Engineering Research), Vol. 10 No.1 (2015) pp. 766-772.

6. B.Padmini Devi,R1.Vinoth. Dr.P.Sudhakar,Dynamic Resource Allocation-Failure Tolerance Matching Efficiency for Clusters in Cloud ,in IJAER (International Journal of Applied Engineering Research) Vol. 10 No.1 (2015) pp. 38047-38051
7. B.Padmini Devi,Dr.P.Sudhakar, A Survey On TaaS & SECaaS IN CLOUD, in Journal of Advances in Chemistry Vol.12.No.15(2016) pp.4795-4798
8. B Padmini Devi, S Chitra, B Madhusudhanan , Improving Security in Portable Medical Devices and Mobile Health Care System Using Trust, Journal of Medical Imaging and Health Informatics 6 (8), 1955-1960, (2016)
9. Priyadarsini K Cloud Testing as a Service IJAEST)International Journal of advanced Engineering sciences and technologies: Vol No. 6, Issue No. 2, 173 –177 ISSN: 2230
10. SergiyVilkomir ,Cloud Testing: a state of the Art Review
11. V. Shrividhya and P. Manimegalai, Network Security In Cloud Computing, Pak. J. Biotechnol. Vol. 14 (Special Issue II) Pp. 220- 222 (2017)
12. Dr.P.Sudhakar, Mrs.B.Padmini Devi, A Survey on TaaS&SECaaS in Cloud, Journal of Advances in Chemistry,vol.12 No.15(2016),ISSN 2321-807X,PP:4795-4798
13. Research on testing tools available at http://research.cs.queensu.ca/~shepard/testing.dir
14. Vanitha,Katherine A, Alagarsamy K ,2012, Software Testing in Cloud Platform, International Journal of ComputerApplications (0975 – 8887)Volume 46 –No.6, May 2012
15. Vengattaraman T, Dhavachelvan P, Baskaran R, 2010 Model of Cloud Based Application Environment for Software Testing, (IJCSIS) International Journal of Computer Science and Information Security , Vol. 7, No. 3, 2010
16. Vinit B. Mohata , Dhananjay M.Dakhane , Ravindra L.Pardhi,2013 Cloud Based Testing: Need of Testing inCloud Platforms, International Journal of Application or Innovation in Engineering & Management (IJAIEM)Volume 2, Issue 3, March 2013
17. http://en.wikipedia.org/wiki/Cloud_computing

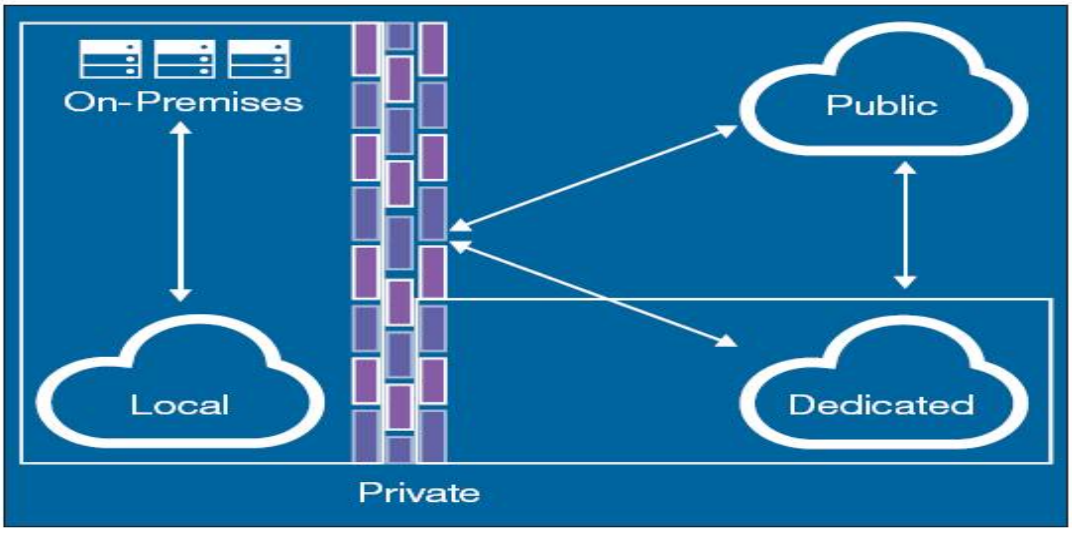


Figure 1: Hybrid cloud

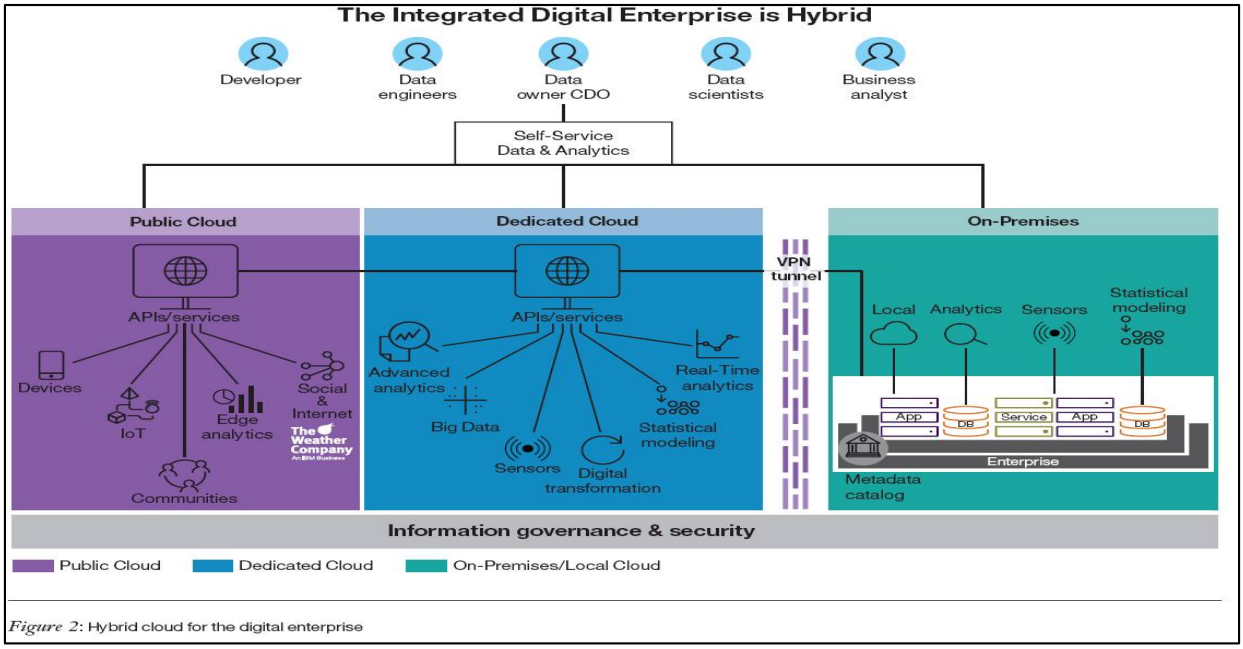


Figure 2: Hybrid cloud for the digital enterprise

Indian Journal of Natural Sciences www.tnsroindia.org.in ©IJONS

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

A Predicting Model for Student Placement Analysis Using Data Mining Algorithms

E.Balraj* and D.Maalini.

Assistant Professor, Department of Information Technology,M.Kumarasamy College of Engineering, Karur,TamilNadu,India.

Received: 13 Feb 2018 Revised: 24 Feb 2018 Accepted: 23 Mar 2018

***Address for correspondence**
E.Balraj
Assistant Professor ,
Department of Information Technology,
M.Kumarasamy College of Engineering,Karur,TamilNadu,India.
E mail:balraje.it@mkce.ac.in

ABSTRACT

Nowadays data mining is becoming a wider technology which is used in many fields. In the field of education, Educational mining is used to understand the different perspective in student data. The scope of this research work to identify the related attributes such academic information, curriculum details, skills of final year student and design a simple predictive model to describe the student placement prediction by using classification technique based on ID3 algorithm. This model is useful for the academic institutions, teachers to predict the students those who are not eligible for placement.

Keywords: Classification, ID3, Educational Mining, Decision Tree.,

INTRODUCTION

The main aim of this research work to classify the students for the placement procedure in academic institutions. For the student & academician, Campus Placement is one of the major roles of any academic institutions. Campus placement is process where all the recruitment industries meet in a single forum and identify the well qualified and talented students, before they are graduating from our academic institution. This paper proposes a simple predictive model to predict the student placement activities.

Data Mining

The technique which is used to extract the information from the huge amount of data is known as data mining. Extracting useful information from large amount of data is more complex when we are using standard approaches .Data Mining is approach which provides more number of data mining algorithms to identify the both relevant & irrelevant attributes from the large amount of data. Normally, it is used to identify the similarity patterns associated

with data. It supports all the sectors where the large amount of data is processed. It provides many tools to analyze the data based on the customer & business requirements. Educational mining is approach which is used to analyze the various information of student, courses. Normally the data were classified into two broad categories. Descriptive mining is a technique which is used to analyze the general properties of any kind of data. Predictive mining is a technique which is used to analyze the current data & predict the future based on predictive mining results. Predictive mining is widely used in the field of education to predict the student dropout, placement prediction (etc).

Classification

The important task of data mining is Classification. Classification technique is used to identify the relationship between output class & input class. A set of classes and attributes defined by a classification problem, where each class is associated with a class attribute and explained by predictor attributes. Classification techniques are classified into two phases. First phase contains data set which consist of class with label —known value of class label input is used to classify the model was given in the dataset that represents the relationship between predictor and class attribute values is built. Second Phase is the classification model which is used to classify with unknown class value(unknown examples) .

Literature review

Erkan Er(2012) proposed simple case study to analyze the student dropout rate for predicting students' performance levels which employs three machine learning algorithms: instance-based learning Classifier, Decision Tree and Naïve Bayes. His results are shown that instance-based learning algorithm gives higher accuracy then other 2 algorithms.

Irena et.al(2015) investigated detect students at risk of failing early in the semester for timely intervention. This paper collects the data from three different data sources PASTA, PIAZZA and assessment marks. It proposes a simple decision tree classifier to identify the student whether he will get pass (or) fail mark in their end semester examination. This classifier is used by the teachers and students to give remediation for the subject.

Mangasuli Sheetal et.al (2016) proposed the method for mining the student's performance based on attributes relevant to academic,co-curricular activities, communication skills to predict whether student will be hired by the companies or not. For that, author is used two algorithms called and "KNN" and "Fuzzy Logic ". The authors concluded that KNN is the best algorithm.KNN gives 97.33% acuuracy which is higher that Fuzzy Logic.

NeelamNaik et.al (2012) given model that states to performance classification of the students pertaining to the placement activity. The error rate of their result was 38.46% when classification were done using result prediction model and while for validation of placement prediction model by using classification tree was recorded as 45.38% respectively.

Algorithm

ID3 (Iterative Dichotomizer 3) algorithm is invented by J. Ross Quinlan in 1979. It is used for building the decision tree using information theory. It builds the decision tree from top down approach with no backtracking. Information Gain is used to select the best attribute for classification.

Algorithm:
Compute classification entropy.
For all attributes
Calculate information gain for each attribute

Select the attribute with highest Information gain
Remove the attribute
End

Entropy

The purpose of Entropy measure is to identify uncertainty about a source of message. It lies between 0 to 1.When entropy is 1 means dataset is homogenous. Entropy is calculated by formula:

$$E(S) = \sum_{j=1}^{c} -P_j \log_2 P_j$$

Where E(S) is the Entropy of S, Pj is the probability of S belonging to class j.

Information Gain

It measures the expected reduction in entropy. ID3 calculates the Gain of all attributes, and select the one with highest gain.

$$G(S,A) = E(S) - \sum_{VeVariables(A)} \frac{|S_V|}{|S|} E(S_V)$$

Where G(S,A) is Information gain ,Sv is the subset of S for which the attribute A has value v, values(A) is the set of all possible values for attribute A.

Evaluating Performance

To evaluate the performance of classification rule F-Measure, recall, precision methods are used.

True Positive = $\frac{TP}{N}$

False Negative = $\frac{FP}{N}$

Recall = $\frac{TP}{TP+FN}$

Precision = $\frac{TP}{TP+FP}$

F-measure = $\frac{2*\text{Recall}*\text{Precison}}{\text{Recall}+\text{Precision}}$

Dataset Information

The following attributes are used to present placement prediction model to improve the student placement percentage. There are 30 attributes are available. Out of that only 9 attributes are used to construct model. Some of the attributes are categorical variable which can help to improve the accuracy of our results.

Attributes	Variable Type	Values
10th Mark	Categorical Variable	Grade 1 (Above 80%) Grade 2 (70-79%) Grade 3(<70%)
12th Mark	Categorical Variable	Grade 1 (Above 80%) Grade 2 (70-79%) Grade 3(<70%)
Period of Study	Categorical Variable	Regular/Lateral

10th Medium	Categorical Variable	Tamil/English
12th Medium	Categorical Variable	Tamil/English
CGPA	Categorical Variable	Grade1 (Above 8) Grade2 (7.0 -7.99) Grade3 (<7.0)
Current Arrear	Numerical variable	Numerical value
History of Arrear	Numerical Variable	Numerical value
Placed	Categorical Variable	Yes/No

RESULTS AND DISCUSSION

The above dataset was used as input and verified the with three different algorithms ID3, CHAID, C5.0. Four Attributes have been taken for to construct decision tree such as Placed, CGPA, 10th Mark, 12th Mark.

CONCLUSION

The main purpose of this study is to predict the student placement information by using classification technique. For that, the datas are collected from the 250 students in the form of database. Results are validated against three algorithms such as ID3, CHAID, and C5.0. Among these algorithms, ID3 is given best accuracy with 96% when compare to remaining two algorithm. Here we conclude that ID3 is the best algorithm to predict the student placement information. This study will help the students, teachers, college management to identify the students who need special care about their placement activities.

REFERENCES

1. Anitha and Nivetha ,Review Extractor and Performance Analyzer by Opinion Mining Techniques, Pak. J. Biotechnol. Vol. 13 Pp. 428- 430 (2016).
2. Dr.S.Chitra,Mrs.G.Mohanaprabha ,Design and Development of an efficient hierarchical approach for multi label protein function predicition,Biomedical Research Special Issue.India PP-S370-379 (2017).
3. Erkan Er, Identifying At-Risk Students Using Machine Learning Techniques: A Case Study with IS 100, International Journal of Machine Learning and Computing, Vol. 2, No. 4PP.476-480 , August (2012) .
4. Mangasuli Sheetal , Prof. Savita Bakare, Prediction of Campus Placement Using Data Mining Algorithm-Fuzzy logic and K nearest neighbor, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 5, Issue 6, ,PP – 309-312 June (2016).
5. G.Mohanaprabha, E.Balraj, A h_m Ant miner using evolutionary algorithm,IJIRSET, Volume 3, Special Issue 3, PP 1687- 1692 March (2014).
6. Neelam Naik, Seema Purohit, Prediction of Final Result and Placement of Students using Classification Algorithm, International Journal of Computer Applications (0975 – 8887) Volume 56– No.12, PP35-40, October (2012).
7. D. Saravana Kumar, T. Jeevalatha,Performance Analysis of Undergraduate Students Placement Selection using Decision Tree Algorithms,International Journal of Computer Applications (0975 – 8887) Volume 108 – No 15, December (2014).
8. V.Ramesh, P.Parkavi, P.Yasodha,Performance Analysis of Data Mining Techniques for Placement Chance Prediction,International Journal of Scientific & Engineering Research Volume 2, Issue 8, ISSN 2229 August (2011).

9. Romero, C., Ventura, S., Eapejo, P.G. and Hervas, C,Data Mining Algorithms to Classify Students, In Proceedings of the 1st International Conference on Educational Data Mining, pp. 8-17 (2008) .
10. Vinothan Damodaran and Saravanan , Institution System Analysis By Using Similarity Based Clustering on Social Network Access, Pak. J. Biotechnol. Vol. 13 Pp. 1-4 (2016).
11. Irena Koprinska, Joshua Stretton and Kalina Yacef, Students at Risk: Detection and Remediation, Proceedings of the 8th International Conference on Educational Data Mining , PP – 512-515. (2015)

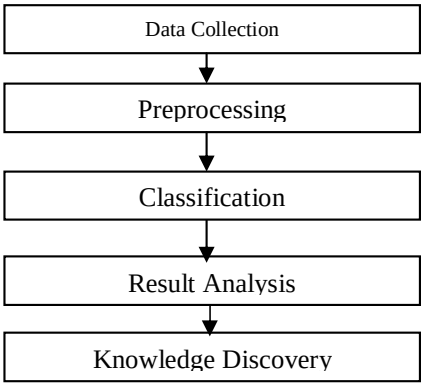


Figure 1 – Architecture of Work

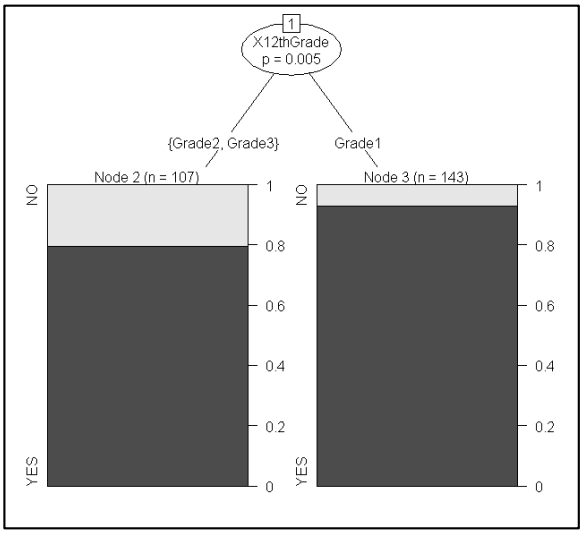


Figure 2 – Decision Tree for Placement Prediction Analysis

Table 1.Accuracy Results of Various Algorithms

Algorithm	No. of Students	No. of Students Results Wrongly Predicted	Accuracy Percentage
ID3	250	10	96%
CHAID	250	15	94%
C5.0	250	19	92.4%

Indian Journal of Natural Sciences www.tnsroindia.org.in ©IJONS

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Phytochemical Analysis of Iraqi *Lycium barbarum* L. by Liquid Chromatography Mass Spectrometry

Thukaa Zuhair Abdul-Jalil*[1], Ahmed A. Hussein[2] and Kawkab Yacoup Saour[3]

[1]Department of Pharmacognosy, College of Pharmacy, University of Baghdad, Iraq.
[2]Department of Pharmaceutics, Dean of College of Pharmacy, University of Baghdad, Iraq.
[3]Department of Pharmaceutical Chemistry, College of Pharmacy, University of Baghdad, Iraq.

Received: 12 Feb 2018 Revised: 20 Feb 2018 Accepted: 14 Mar 2018

***Address for correspondence**
Thukaa Zuhair Abdul-Jalil
Department of Pharmacognosy,
College of Pharmacy, University of Baghdad, Iraq.
E mail: aljas_dadid@yahoo.com

ABSTRACT

The genus Lycium has about 80 species, but only a limited number of species have been studied chemically, *Awsaj {Lycium barbarum* (Solanaceae family)} recently received much attention as one of the trendiest plants containing essential nutritional elements with a wide array of pharmacological activity. A sensitive and specific liquid chromatography mass spectrometry method developed for the first time in Iraq to identify nine bioactive compounds (two hydroxycinnamic acids, one hydroxycoumarin, two flavan-3-ols, three flavonols and flavonol glycosides and one tropane alkaloid) in the fruits, leaves, stems and roots of the wild Iraqi Awsaj plant. Each part of the Awsaj plant was extracted separately with 85% methanol using Soxhlet apparatus, fractionated three times sequentially with petroleum ether which discarded chloroform, ethyl acetate, and n-butanol solvents. The chromatographic analysis was performed using column C18 (15cm, 3.5*4.6µm) and gradient elution with methanol and water containing 0.1% acetic acid. Polarity switching mode was used in the optimization of multiple reaction monitoring condition. With the aid of reference standards complemented by accurate mass measurement data, Caffeic acid, p-coumaric acid, scopoletin, catechin, epicatechin, quercetin, rutin, quercitrin and atropine were identified in all Awsaj parts and distributed in the extraction fraction as follows: atropine and scopoletin in chloroform fraction F2; Caffeic acid, p-coumaric acid, scopoletin, catechin, epicatechin, quercetin, rutin, and quercitrin in ethyl acetate fraction F3; and catechin, epicatechin, quercetin, rutin, and quercitrin in n-butanol fraction F4. All of these compounds were identified from Awsaj for the first time in Iraq.

Keywords: Awsaj, liquid chromatography mass spectrometry (LC-MS), Lycium barbarum, Solanaceae

INTRODUCTION

The increasing evidence for a relationship between diet and health shows the importance of plant secondary metabolites and their effect on different physiological functions and health. This has generated new concepts in nutrition aiming at developing and promoting functional foods, known to be rich with such bioactive compounds [1]. Awsaj [Lycium barbarum L. (Solanaceae)], also known as Goji, can be considered one of those foods [2]. Awsaj has a long tradition of being a food and medicinal plant in China and various other Asian countries, since this plant is widely distributed in warm regions of the world, particularly in Mediterranean area, and southwest and central Asia [3]. At the beginning of the twenty-first century, Awsaj became increasingly popular in Europe and North America, due to its nutritional richness in various vitamins, minerals, antioxidants, and amino acids [4]. In parallel to all these nutritional benefits, it has many health-protective benefits such as age-related macular degeneration [5], antioxidant [6], antitumor activities [7], neuroprotective effects [8], male fertility facilitation [9], reduction of blood glucose and serum lipid in Alloxan induced diabetic models [10], as well as immunity enhancement [11].

This paper describes, for the first time in Iraq, the identification and characterization of bioactive compounds of four different parts of the Iraqi Awsaj plant: the fruits, leaves, stems, and roots (figure 1) using sensitive, accurate and specific methods coupling high performance liquid chromatography with diode array detector and electrospray ionization mass spectrometry.

MATERIALS AND METHODS

Plant material

Whole plants of Awsaj (*Lycium barbarum* L.), which grows as a wild plant in Iraq, were collected at the University of Baghdad in Al-Jadiryia District, Baghdad, then authenticated by the herbarium of the Department of Biology, College of Science at the University of Baghdad registered at BUH No. 50777.

Extraction procedure

The fruits, leaves, stems, and roots of the Awsaj plant were collected and cleaned to be free of extraneous materials, shade dried, and ground to a fine powder by using electric blender then weighted for extraction procedure. The extraction was carried out by using Soxhlet apparatus in which each part of Awsaj (fruits, leaves, stem and roots) was extracted separately with 85% methanol. The crude methanol extract was filtered by passing the extract through Whatmann No. 1 filter paper and then concentrated under reduced pressure at 40 C^{o} by using rotary evaporator. The residues were then suspended in distilled water (250 ml) and partitioned successively with petroleum ether (B.P. 60-80), chloroform Ethyl acetate and n-butanol (3*500 for each). Fractions were concentrated under reduced pressure to dryness. Each dried fraction was weighted and subjected for identification by LC/MS [12, 13].

Phytochemical analysis by liquid chromatography mass spectrometry (LC/MS)

LC condition

The LCMS-8040 series system (Shimadzu, Japan) was used equipped with a degasser, binary gradient pump, column thermostat, autosampler, diode array detector (DAD) and controlled by chemstation software. This liquid chromatography system was coupled with Shimadzu, Japan mass spectrometer (LC/MS Ion trap VL.). For the separation, reverse phase analytical column was employed (C18, length 15 cm, pore size 3.5 µm, and inner diameter 4.6µm). The work temperature was 48 C^{o}. The mobile phase consisted of methanol (A) and water containing 0.1% acetic acid (B) at a flow rate 0.5 ml/min. The elution started with a liner gradient, beginning with 5% methanol and

ending at 42% methanol for 57 minutes, then 5% methanol for the next 10 minutes. The injection volume was 10µl. The mobile phase was prepared daily, filtered through 0.45µm membrane filter (Millipore), and sonicated before use.

MS condition

The mass analysis was performed on Shimadzu ion trap mass spectrometer with an electrospray interphase (ESI). The chromatographic data was processed using chemstation and data analysis software from Lab solution, japan. MS data were acquired in both positive and negative ionization modes and operated according to defined condition: nitrogen gas flow rate 3l/min, Desolvation line (DL) temperature 250 C°, heat block temperature 400 C°, drying gas flow rate 15 l/min and the mass spectra were recorded using the fill scan mode in the range of 100-800 Daltons [14-17].

Preparation of extract fractions

10mg from F2, F3 and F4 for each plant parts (fruits, leaves, stems and roots) that obtained by Soxhlet extraction methods were diluted separately with 10ml of ethanol and then centrifuged for 2 minutes at 14000 rpm and 25C° using Eppendorf centrifuge 5417R, then supernatant were taken to be filtrated by 0.45 µm pore size disposable filters before utilization for analysis by LC/MS.

Preparation of standards

Nine bioactive reference samples distributed in five major categories, hydroxyl cinnamic acid, hydroxyl coumarins, flavan-3-ol, flavonols and their glycosides and tropane alkaloids had been identified and qualified by LC/MS using 10 mg from each standard; illustrated in figure 2; dissolved in 20 ml of ethanol to obtain a solution at a final concentration of (0.5mg/ml), then filtrated by 0.45 µm pore size disposable filters before utilization for analysis by LC/MS.

RESULTS AND DISCUSSION

Sixteen extraction portions were obtained from extraction experimental work of whole Awsaj plant (fruits, leaves, stems and roots) in which the first fraction for each part was petroleum ether (F1), which was discarded, second fraction was chloroform (F2), the third fraction was ethyl acetate (F3) and finally the fourth fraction was n-butanol (F4). The percentage yields of each fraction (F1-F4) for each part of Awsaj are shown in Table 1.

LC-MS and MS detection in both positive and negative ionization modes were used to obtain more information on the structural features of analyzed compounds in addition to avoid or limit the interference from background. The multiple reaction monitoring (MRM) analysis mode was used instead of product ion monitoring. Nine bioactive compounds distributed in five major categories, hydroxyl cinnamic acid, hydroxyl coumarins, flavan-3-ols, flavonols and tropane alkaloids had been analyzed by LC-MS for the first time in Iraq and these compounds were designed as L1, L2, L3, L4, L5, L6, L7, L8 and L10. The full identification of compounds was performed by comparing the retention times and mass spectra with those of standards in the same chromatographic conditions (figure 3). The identification of the compounds within each class based on chromatographic behavior, fragmentation pattern and occurrence in each part of the Awsaj plant discussed below and summarized in table 2. Whereas the total ion chromatogram (TIC) and extracted ion chromatogram for each fraction (F2-F4) for each part (figure 4-15).

Identification of hydroxycinnamic acid

Two different polyphenols in the category of hydroxycinnamic acid were found to occur in all parts of the Awsaj plant in the ethyl acetate fraction (F3) (figures 5, 8, 11and 14). L1 and L2 were identified by comparing their retention times and characterized MS spectral data with those of authentic standards (tables 2-6). Generally, deprotonated phenolic acid $[M-H]^-$ produce a typical fragmentation pattern after collision induced dissociation (CID) characterized by the loss of COO (44 unit) from the carboxylic acid group, providing an anion of $[M-H-COO]^-$ [18]. Therefore, qualitative identification was performed using multiple reaction monitoring (MRM) mode in which L1 identified as Caffeic acid with target ions at m/z 179→m/z 135 for and L2 had$[M-H]^-$ at m/z 163.04 and M/S fragmentation ion in negative mode at 119 typical fragment of p-coumaric acid (figures 3A and 3B).

Identification of hydroxycoumarin

Both negative and positive ion modes were tried. Positive ion mode was found to be more sensitive for identification and characterization of L3 that was found for all parts of Awsaj plant in the chloroform and ethyl acetate fractions (F2 and F3) (tables 2-6) and (figures 4,5,7,8,10,11,13,14). Under collision induced dissociation, L3 undergo neutral loss from pseudo-molecular ion $[M+H]+$ at m/z 193, producing the corresponding fragment ions $[M+H-CH3]+$ at m/z 178, $[M+H-CH3-CO]+$ at m/z 149 and $[M+H-CH3-CO-OH]+$ at m/z 132.9, these fragmentation ions are in agreement with scopoletin fragmentation ions (figure 3C) [19].

Identification of flavan-3-ols

The positive ion mass spectra of ethyl acetate and n-butanol fractions (F3 and F4) in all parts of the Awsaj plant revealed two ions at m/z 291 with different retention times, due to the presence of diastereoisomers (L4 and L5) (tables 2-6) and (figures 5,6,8,9,11,12,14 and 15). The protonated precursor ion at m/z 291 showed the major product ions at m/z 139, 165 and 123, at which after CID, the ion at m/z 139 results from a Retro-Diels-Alder (RDA) cleavage of ring C, while the ion at m/z 165 was attributed to the elimination of ring A by heterocyclic ring fusion (HRF) in addition to the ion at m/z 123 that may be formed by benzo furan forming fusion (BFF) and loss of water molecule. L4 showed a pseudo-molecular ion at m/z 291, which produced daughter ions at m/z 165, which evidenced the elimination of ring A by heterocyclic ring fusion (HRF) and was identified as catechin. L5 showed a pseudo-molecular ion at m/z 291 producing a fragment at m/z 123 (formed by benzo furan forming fusion (BFF) and loss of water molecule) and was identified as epicatechin (figures 3D and 3E) [20].

Identification of flavonols and flavonols glycosides

According to the identical LC-MS characteristics as that of the authenticated standards, L6,L7 and L8 were identified in the ethyl acetate and n-butanol fractions (F3 and F4) in all parts of Awsaj plant (tables 2-6) and (figure 5,6,8,9,11,12,14 and 15). Upon fragmentation by CID, L6 was identified as quercetin as it produced ions at m/z 151 and 179 which resulted from a cleavage of heterocyclic C ring by RDA. On the other hand L7 and L8 were identified as flavonol-O glycosides (rutin and quercitrin respectively) since they showed the deprotonated $[M-H]^-$ molecule of the glycoside and the $[A-H]^-$ ion corresponding to the deprotonated aglycone. The latter ion was formed by losing the rutinose and rhamnose moiety from the corresponding glycosides, the identity being confirmed by comparison with an authenticated standard sample (figures 3F-3H) [21].

Identification of tropane alkaloid

The Awsaj plant belongs to Solanaceae family in which tropane alkaloids are the most well-known alkaloids to be found in this family. Therefore, screening for atropine as one type of this class of active metabolites in all parts of the Awsaj plant was done by using a highly sensitive and selective LCMS method. The LCMS chromatogram of L10 that

shown in figures 4,7,10 and 13 revealed a base peak at m/z 290, corresponding to the pseudo-molecular ion $[M+H]^+$ and the fragmentation at m/z 124 which is due to breakage of the ester bond between tropine and tropic acid. This result identical to the MS spectrum of atropine as shown in Figure 3I, the LCMS analysis proves the presence of atropine [the racemate of (+) and (-) hyoscyamine] in all examined parts of the Awsaj plant in the chloroform fraction (F2) (tables 2-6) [22].

This study represents the first chemical investigation of wild Iraqi *Lycium barbarum* (Awsaj) and demonstrates the presence of a variety of phytochemical compounds. For that purpose, LC-MS-MS fragmentation for each analyte was investigated by multiple reaction monitoring (MRM) analyzed method coupled with ion polarity switching mode (positive ion mode and negative ion mode) has been optimized, enabled identification of nine phytochemical compounds. This developed method was done by identifying target compounds with a high selectivity and sensitivity technique by comparison with authenticated standards. The negative ion mode was found to be more suitable for hydroxycinnamic acid and flavonols, while the positive ion mode was found to be more suitable for hydroxycoumarin, flanan-3-ols, and tropane alkaloids.

Nine compounds, including two hydroxycinnamic acids (Caffeic acid and p-coumaric acid), one hydroxycoumarin (scopoletin), two flavan-3-ols (catechin and epicatechin), three flavonols (quercetin, rutin and quercitrin) and one tropane alkaloid (atropine), were identified in all parts (fruits, leaves, stems and roots) of the Iraqi Awsaj plant using LCMS for the first time in Iraq.

Among the nine compounds identified in the whole Awsaj plant, two compounds (scopoletin and atropine) were detected in the chloroform fraction (F2). Eight compounds (Caffeic acid, p-coumaric acid, scopoletin, catechin, epicatechin, quercetin, rutin and quercitrin) were detected in the ethyl acetate fraction (F3). While the n-butanol fraction (F4) showed the presence of five compounds (catechin, epicatechin, quercetin, rutin and quercitrin). Further work is needed to quantify and isolate the above compounds identified in this study, in addition to using new bioactive standards to LCMS analysis method in order to confirm the presence of other compounds in the whole parts of the Iraqi Awsaj plant.

ACKNOWLEDGMENTS

This work was carried out with the support of Iraqi National Center for Drug Control and Research (INCDCR), with special thanks to pharmacist Dunya Fikrat Asker for her help.

REFERENCES

1. Benchennouf A., Grigorakis S., Loupassaki S. and Kokkalou E.: Phytochemical analysis and antioxidant activity of Lycium barbarum (Goji) cultivated in Greece. Pharmaceutical biology 2016; 55(1): 596-602.
2. Carnes J., Del Arramendi C.H., Ferrer A., Huertas A.J., Lopez-Matas M.A.et al: Recently introduced food as new allergenic sources sensitization to Goji berries (Lycium barbarum). Food chem 2013; 137: 130-135.
3. Potterat O.: Goji (Lycium barbarum and L. chinense): Phytochemistry, pharmacology and safety in the perspective of traditional uses and recent popularity. Planta Med 2010; 76: 7-19.
4. Endes Z., Uslu N., Ozcan M.M. and Er F.: Physicochemical properties, fatty acid composition and mineral content of Goji berry (Lycium barbarum) fruit. J Agroaliment Proc Technol. 2015; 21(1): 36-40.
5. Bucheli P., Vidal K., Sheu L., Gu Z., Zhang C.,Miller L.E. et al: Goji berry effects on macular characteristics and plasma antioxidant level. Optom Vis Sci 2011; 88(1): 257-262.
6. Magiera S. and Zareba M.L.: Chromatographic determination of phenolic acid and flavonoids in Lycium barbarum and evaluation of antioxidant activity. Food Anal. Methods 2015; 8:2665-2674.

7. Mohammed Z.Y., Al Samarrae K.W. and Hamza S.J.: Preliminary study for the anticancer activity of flavonoids extracted from wild Lycium barbarum. IJRSB 2015; 3(12): 88-94.
8. Amagase H. and Farnsworth N.R.: A review of botanical characteristics, phytochemistry, clinical relevance, inefficacy and safety of Lycium barbarum fruits (Goji). Food research international 2011; 44:1702-1717.
9. Zhang C.,Wang A., Sun X., Li X., Zhao X.Li Sh.,et al: Protective effects of Lycium barbarum polysaccharides on testis spermatogenic injury induced by Bisphenol A in mice. Evidence-based complementary and alternative medicine 2013: 1-10.
10. Pai P.G., Habeeba U.P., Ullaal Sh., Shoeb A.P and Ramya P.M.: Evaluation of Hypolipidimic effect of Lycium barbarum (Goji berry) in Murin model. Journal of natural remedies 2013; 13(1):4-8.
11. Kulczynski B. and Gramza- Michalowska A.: Goji berry (Lycium barbarum): Composition and health effects. A review .Pol.J.Food Nutr.Sci 2016;66(2):67-75.
12. Abdl-Jalil Th. Z.: Screening of rutin from seeds and leaves extract of dill, coriander and fennel cultivated in Iraq. Pharmacy global IJCP 2013; 3(2): 1-6.
13. Hamad M.N.: Detection and isolation of flavonoids from calendula officinalis (F.A steraceae) cultivated in Iraq. Iraqi J Pharm Sci 2016; 25(2): 1-6.
14. Moldovan R.I., Oprean R., Benedec D., Hanganu D., Duma M., Oniga I., et al: LCMS analysis, antioxidant and antimicrobial activities for five species of Mentha cultivated in Romania. Digest journal of nanomaterials and biostructures 2014; 9(2): 559-566.
15. Farias L.S. and Mendez A.S.: LC/ESI-MS method applied to characterization of flavonoids glycosides in B.forficata subsp. Pruinosa. Quim. Nova 2014; 37(3): 483-486.
16. Adams M., Wiedenmann M., Tittel G. and Bauer R.: HPLC-MS trace analysis of atropine in Lycium Barbarum berries. Phyto chem. Anal. 2006; 17: 279-283.
17. Kajdzonoska M., Gjamovski V. and Stefova M.: HPLC-DAD-ESI-MS identification of phenolic compounds in cultivated strawberries from Macedonia. MJCCAQ 2010; 29(2): 181-194.
18. Hossain M.B., Rai D.K., Brunton N.P., Martin-Diana A.B. and Barry-Ryan C.: Characterization of phenolic composition in Lamiaceae species by LC-ESI-MS/MS. J.Agric.Food Chem 2010; 58(19): 10576-10581.
19. Zeng Y., Li S. Wang X., Gong T., Sun X. and Xhang Z.: Validation LCMS/MS method for the determination of scopoletin in rat plasma and its application to pharmacokinetic studies. Molecules 2013; 20: 18988-19001.
20. Hamed A.L., Al-Ayed A.S., Moldoch J., Piacente S., Oleszek W. and Stochmal A. : profiles analysis of proanthocyanidins in the argun nut (Medemia argun- an ancient Egyptian palm) by LC-ESI-MS/MS. J.Mass spectrum 2014: 1-10.
21. Sun J., Liaug F., Bin Y., Li P. and Duan Ch.: Screening non-colored phenolics in red wines using liquid chromatography/ultraviolet and mass spectometry/ mass spectometry libraries. Molecules 2007; 12:679-693.
22. Boros B., Farkas A., Jakabova S., Bacskay L., Kilar F. and Felinger A.: LCMS quantitative determination of atropine and scopolamine in the floral nectar of Datura species. Chromatographia 2010: 1-7.

Table 1: Fractions yield percentage of whole Awsaj plant parts

Parts of Awsaj	Petroleum ether Fraction (F1)	Chloroform Fraction (F2)	Ethyl acetate Fraction (F3)	n-butanol Fraction (F4)
Fruits	8.5	1.9	3.9	7.8
Leaves	10.5	2.7	4.3	6.5
Stems	3.2	1.2	1.6	2.4
Roots	5.3	2.3	2.6	4.2

Table 2: characterization of identified compounds in whole parts of Awsaj by LC-MS analysis

Compounds	Retention time-min	Empirical formula	Ion polarity	Precursorion m/z	Production m/z	Detection in Whole parts
Atropine	2.313	C17H23NO3	Positive	290	124	F2
(-)-epicatechin	2,84	C15H14O6	Positive	291	123,165	F3, F4
Quercitrin	3.079	C21H20O11	Negative	447	301.1	F3, F4
Scopoletin	3.269	C10H8O4	Positive	193	132.9	F2, F3
Caffeic acid	24.704	C9H8O4	Negative	179	135	F3
P-Coumaric acid	33.38	C9H8O3	Negative	163	119	F3
Rutin	34.01	C27H30O16	Negative	609.1	301.1	F3, F4
Quercetin	60.357	C15H10O7	Negative	301	179	F3, F4
(+)-Catechin	61.6	C15H14O6	Positive	291	165,123	F3, F4

F2=chloroform fraction, F3=ethyl acetate fraction, F4= n-butanol fraction

Table 3: Retention time of identified compounds in fruits part of Awsaj plant and their authenticated standards.

Compounds	R_t St	R_t FF2	R_tFF3	R_tFF4
L1	24.704		25.112	
L2	33.38		33.588	
L3	3.269	3.285	3.318	
L4	61.6		60.852	61.579
L5	2.84		2.545	2.83
L6	60.337		59.48	59.35
L7	34.01		34.85	34.75
L8	3.079		3.019	3.149
L10	2.313	2.685		

R_t= retention time, St= standard, FF2=fruits chloroform fraction, FF3= fruits ethyl acetate fraction, FF4= fruits n-butanol fraction

Table 4: Retention time of identified compounds in leaves part of Awsaj plant and their authenticated standards.

Compounds	R_t St	R_t LF2	R_tLF3	R_tLF4
L1	24.704		24.855	
L2	33.38		33.189	
L3	3.269	3.15	3.83	
L4	61.6		61.01	61.655
L5	2.84		2.65	2.624
L6	60.337		59.75	59.655
L7	34.01		34.65	34.39
L8	3.079		3.156	3.229
L10	2.313	2.49		

R_t= retention time, St= standard, LF2= leaves chloroform fraction, LF3= leaves ethyl acetate fraction, LF4= leaves n-butanol fraction

Table 5: Retention time of identified compounds in stems part of Awsaj plant and their authenticated standards.

Compounds	R_t St	R_t SF2	R_tSF3	R_tSF4
L1	24.704		24.588	
L2	33.38		33.663	
L3	3.269	3.315	3.51	
L4	61.6		61.596	61.559
L5	2.84		2.885	2.39
L6	60.337		60.03	
L7	34.01		34.31	34.06
L8	3.079		3.28	3.079
L10	2.313	2.216		

R_t= retention time, St= standard, SF2= stems chloroform fraction, SF3= stems ethyl acetate fraction, SF4= stems n-butanol fraction

Table 6: Retention time of identified compounds in roots part of Awsaj plant and their authenticated standards.

Compounds	R_t St	R_t RF2	R_tRF3	R_tRF4
L1	24.704		24.922	
L2	33.38		33.35	
L3	3.269	3.346	3.314	
L4	61.6		61.569	61.85
L5	2.84		2.633	2.624
L6	60.337		59.56	59.748
L7	34.01		34.45	34.488
L8	3.079		3.058	3.423
L10	2.313	2.588		

R_t= retention time, St= standard, RF2= roots chloroform fraction, RF3= roots ethyl acetate fraction, RF4= roots n-butanol fraction.

Figure 1: *Lycium barbarum* plant

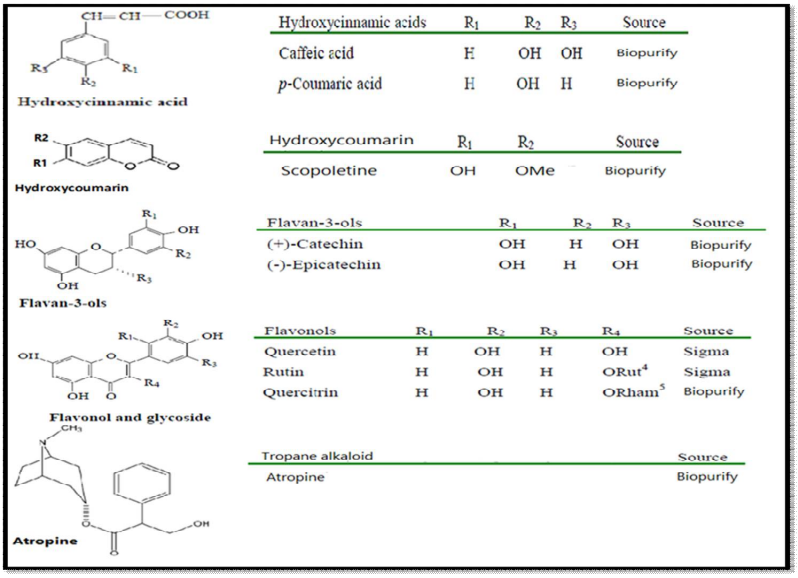


Figure 2: show the structure and the sources of purchased bioactive standards

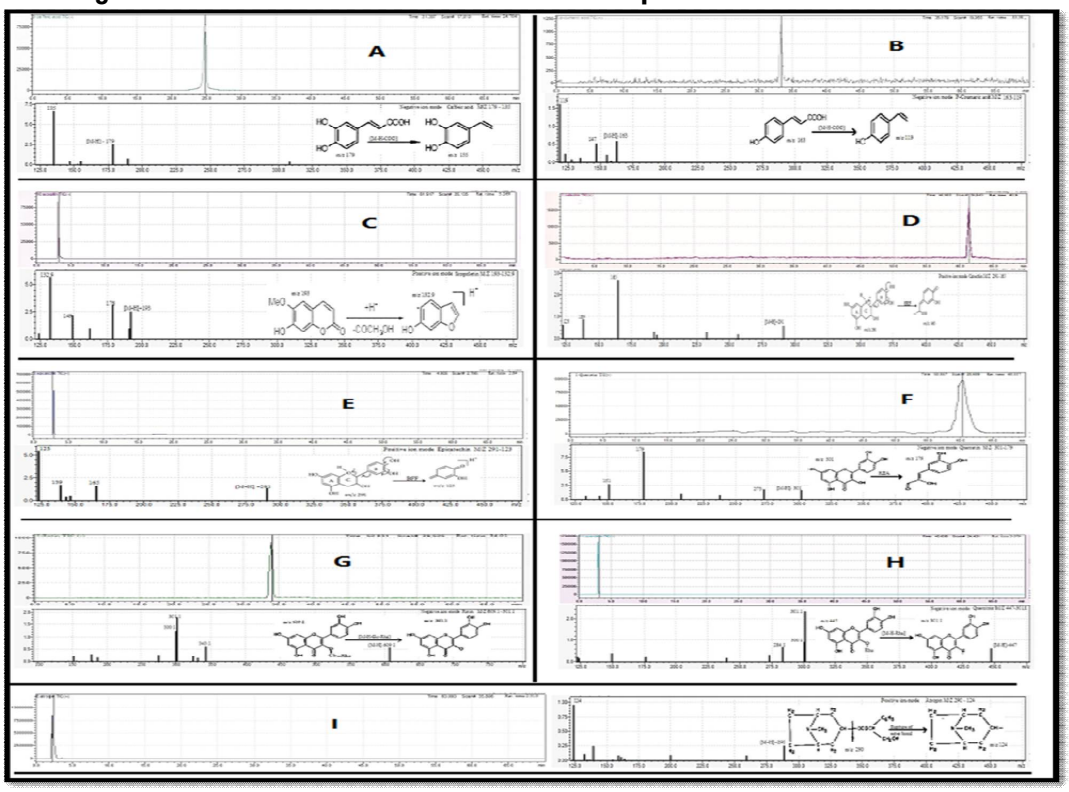


Figure 3: Representative multiple reaction monitoring (MRM) chromatograms obtained for reference standards using LC-MS method. A= Caffeic acid, B= P-coumaric acid, C= scopoletin, D= catechin, E= epicatechin, F= quercetin, G= rutin, H= quercitrin and I=atropine.

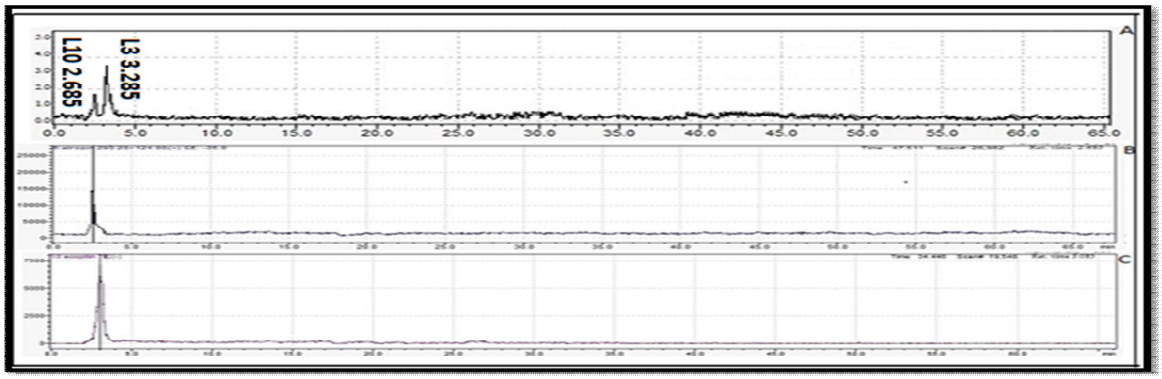


Figure 4: LC-MS chromatogram of chloroform fraction of Awsaj fruits (F2), A= TIC =Total ion chromatogram, B= EIC= Extracted ion chromatogram of L 10, C= Extracted ion chromatogram of L3.

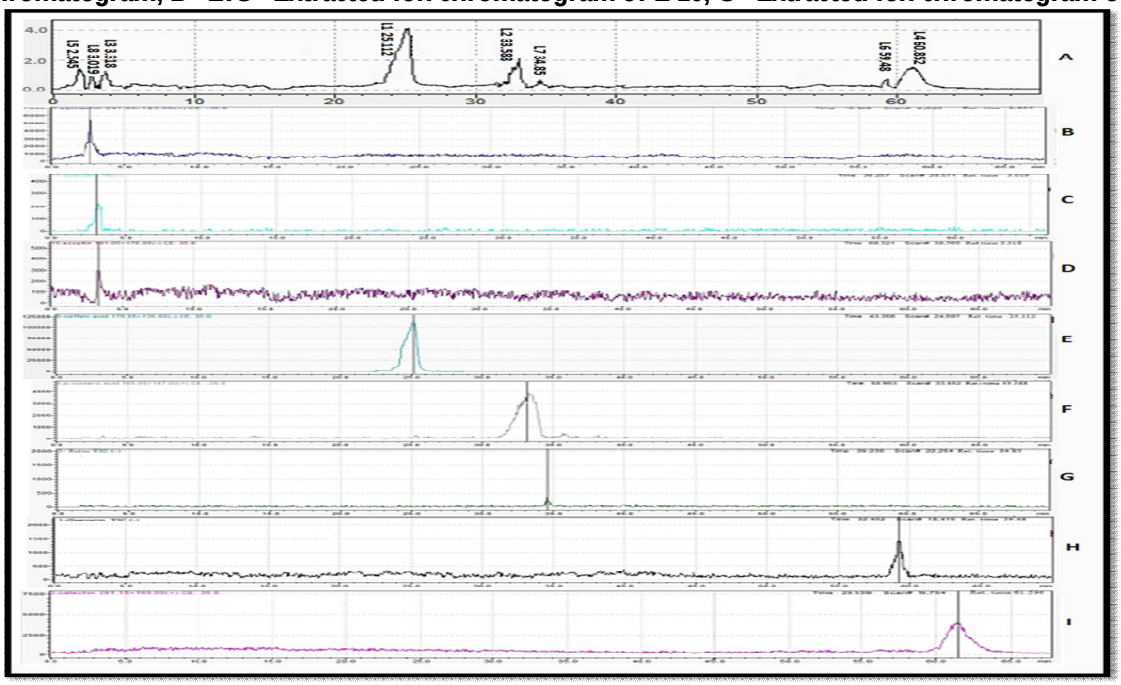


Figure 5: LC-MS chromatogram of ethyl acetate fraction of Awsaj fruits (F3), A= TIC, B to I= EIC of L 5,L 8, L 3, L 1, L 2, L 7, L 6 and L 4 respectively.

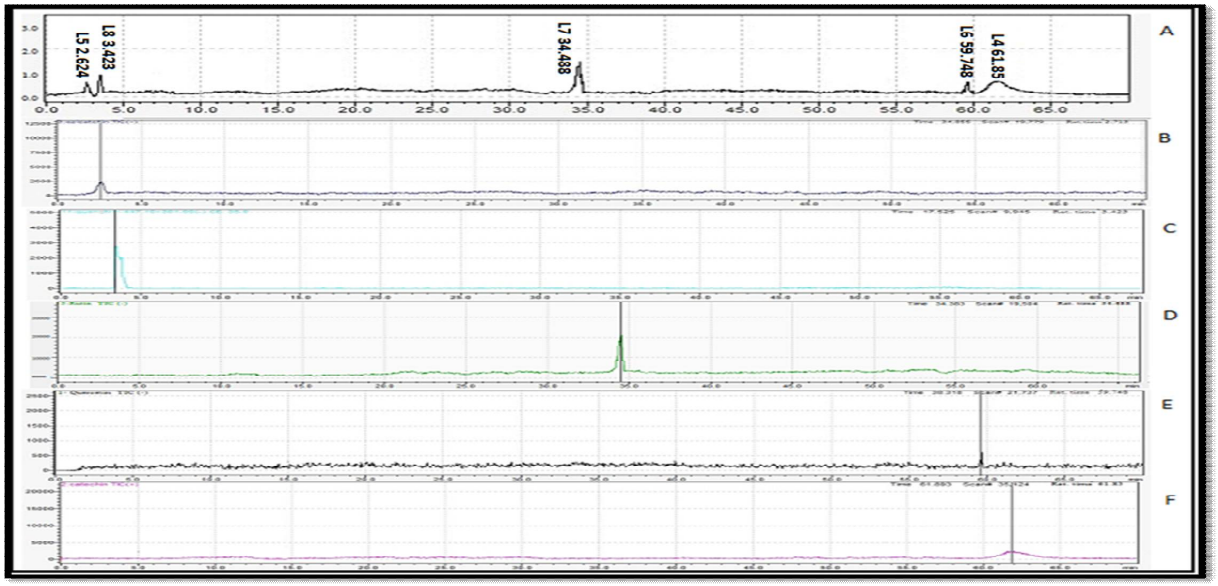


Figure 6: LC-MS chromatogram of n-butanol fraction of Awsaj fruits(F4), A= TIC, B to F= EIC of L 5, L 8, L 7, L 6 and L 4 respectively.

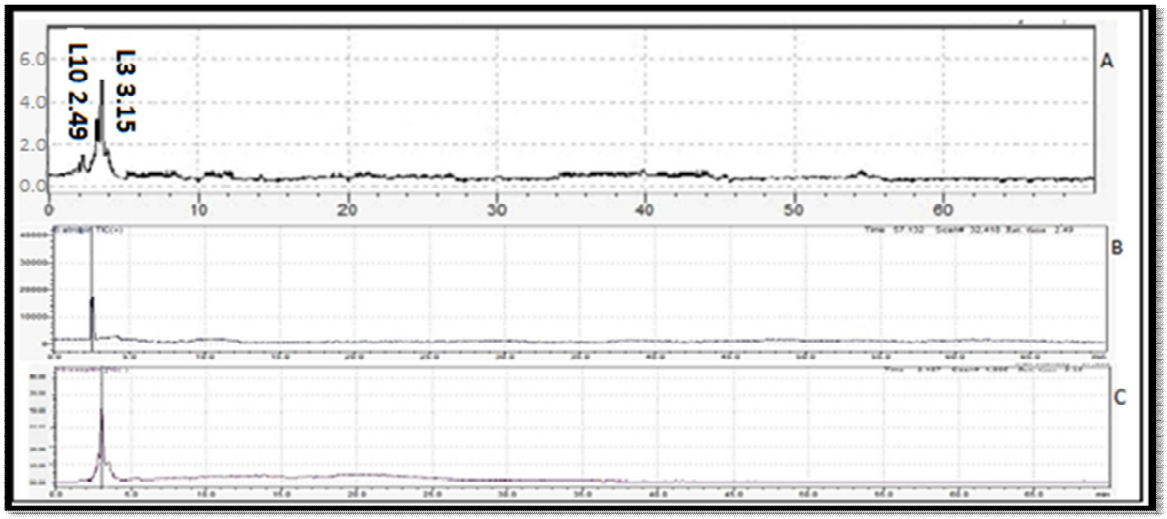


Figure 7: LC-MS chromatogram of chloroform fraction of Awsaj leaves (F2), A= TIC =Total ion chromatogram, B= EIC= Extracted ion chromatogram of L 10, C= Extracted ion chromatogram of L 3.

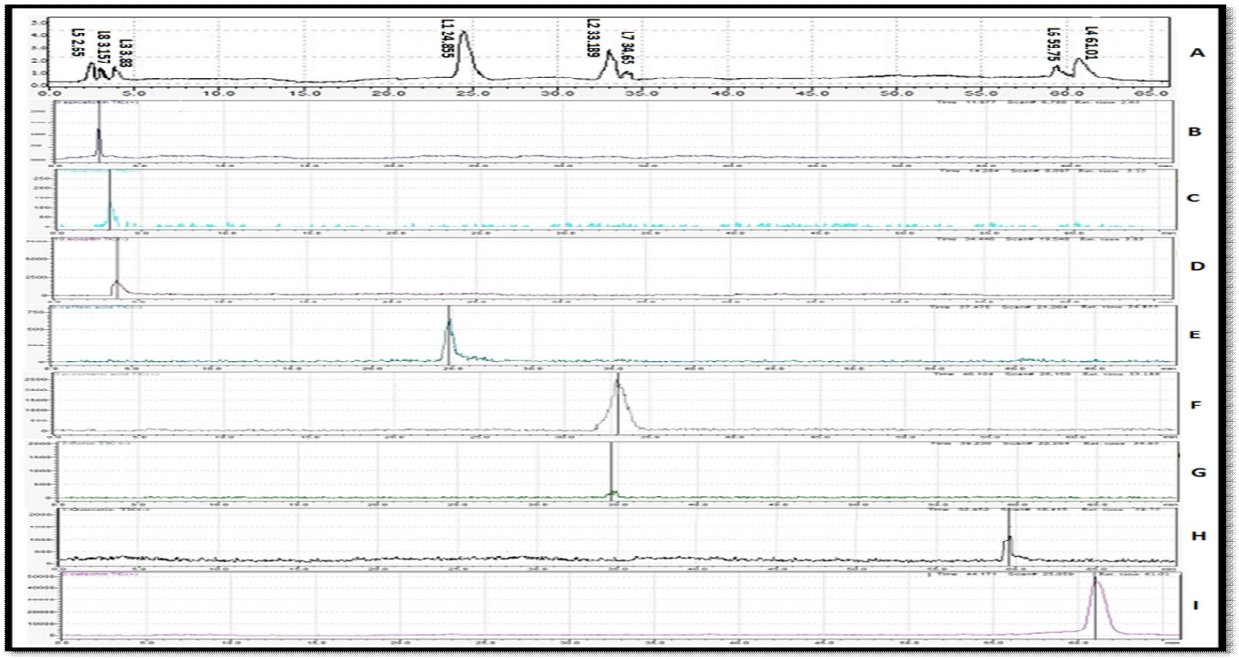


Figure 8: LC-MS chromatogram of ethyl acetate fraction of Awsaj leaves (F3), A= TIC, B to I= EIC of L 5,L 8, L 3, L 1, L 2, L 7, L 6 and L 4 respectively

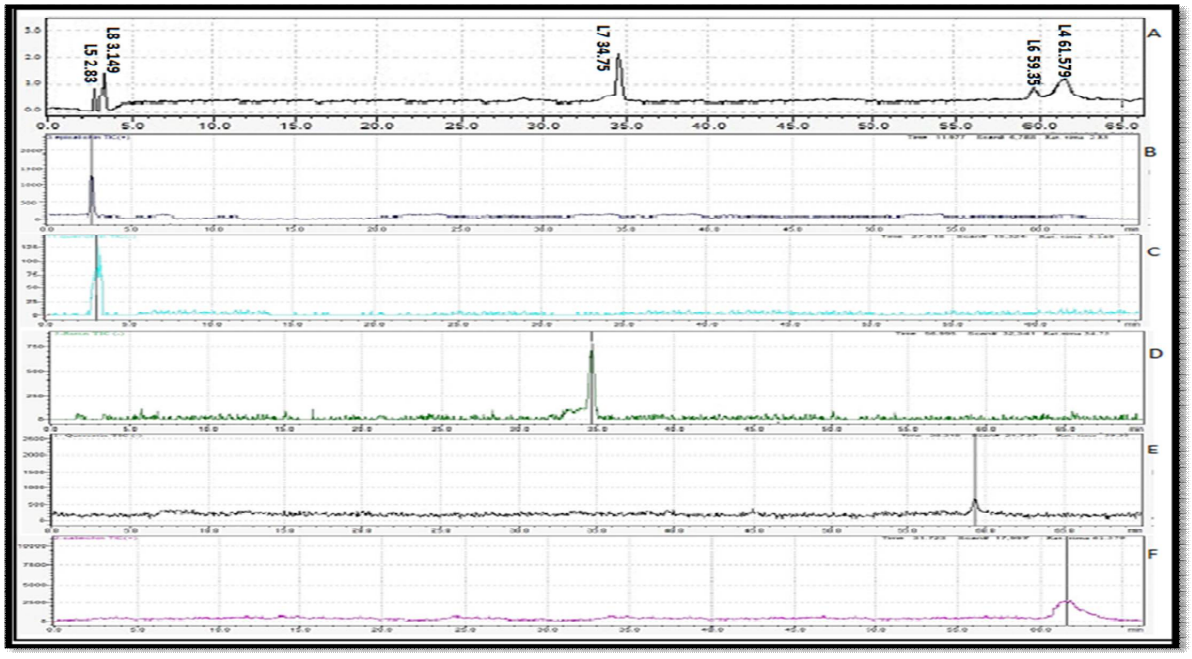


Figure 9: LC-MS chromatogram of n-butanol fraction of Awsaj leaves (F4), A= TIC, B to F= EIC of L 5, L 8, L 7, L 6 and L 4 respectively.

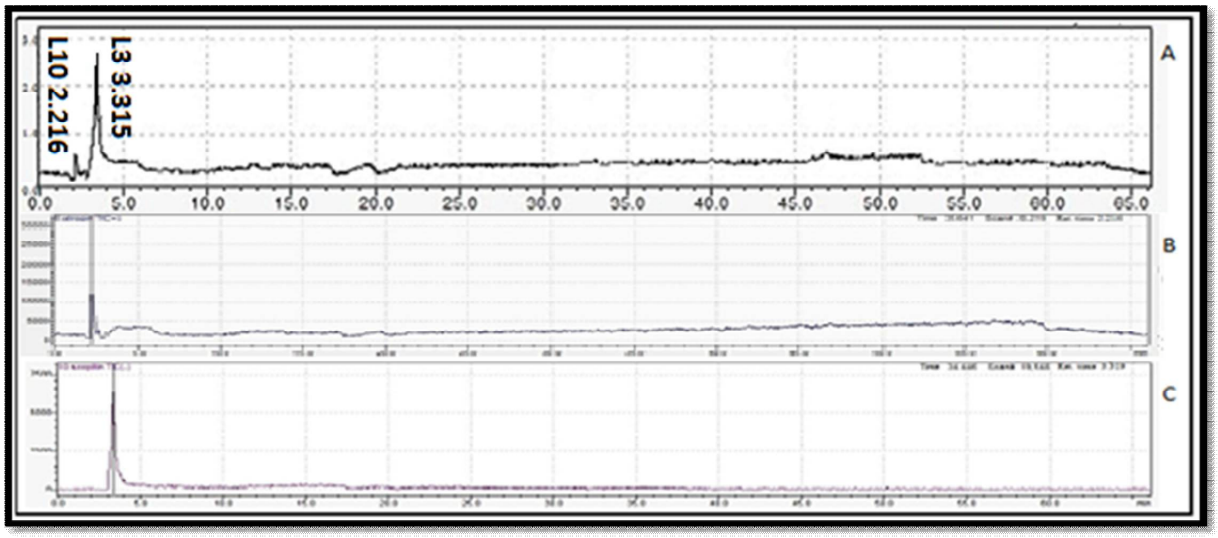


Figure 10: LC-MS chromatogram of chloroform fraction of Awsaj stems (F2), A= TIC =Total ion chromatogram, B= EIC= Extracted ion chromatogram of L 10, C= Extracted ion chromatogram of L 3.

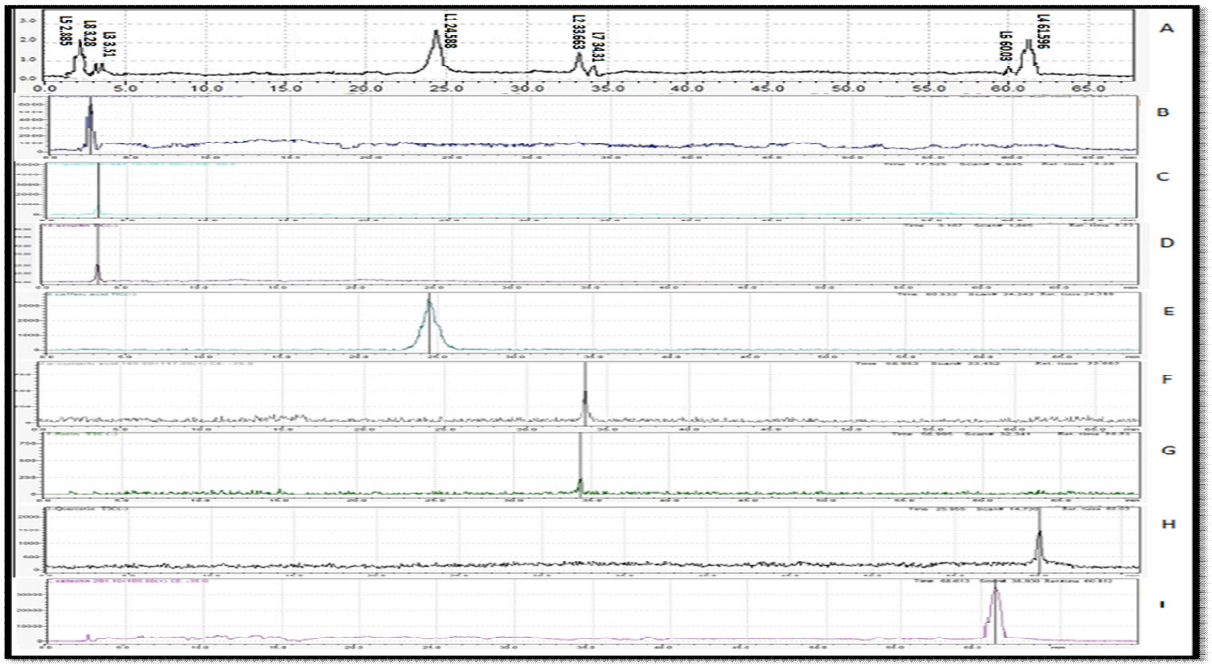


Figure 11: LC-MS chromatogram of ethyl acetate fraction of Awsaj stems (F3), A= TIC, B to I= EIC L 5,L 8, L 3, L 1, L 2, L 7, L 6 and L 4 respectively

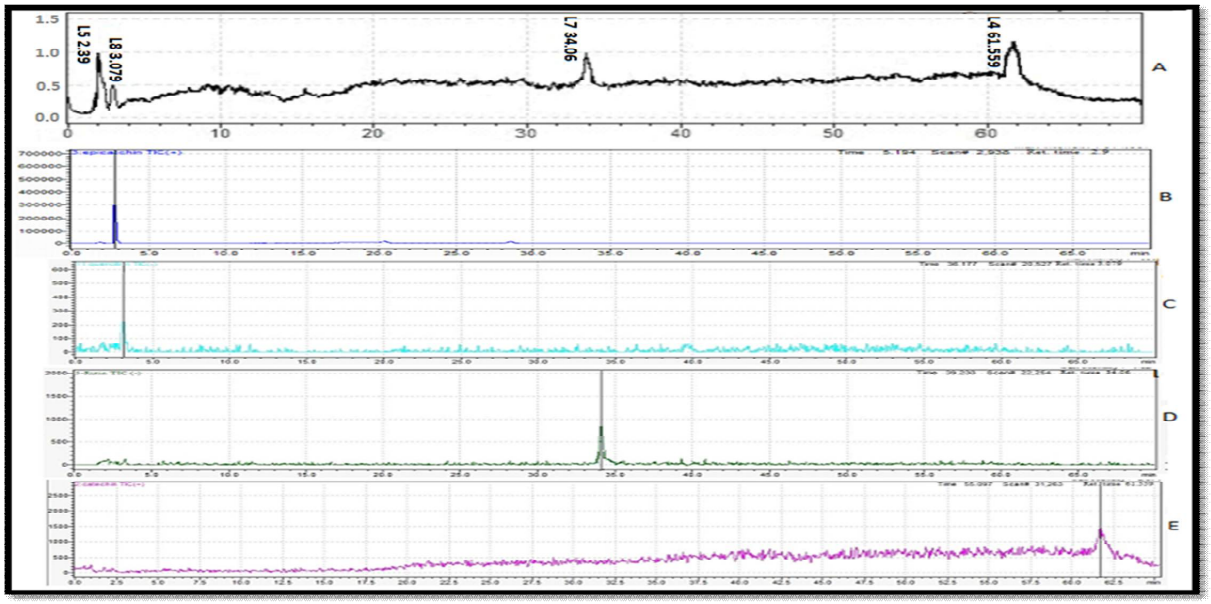


Figure 12: LC-MS chromatogram of n-butanol fraction of Awsaj stems (F4), A= TIC, B to F= EIC of L 5, L 8, L 7 and L 4 respectively.

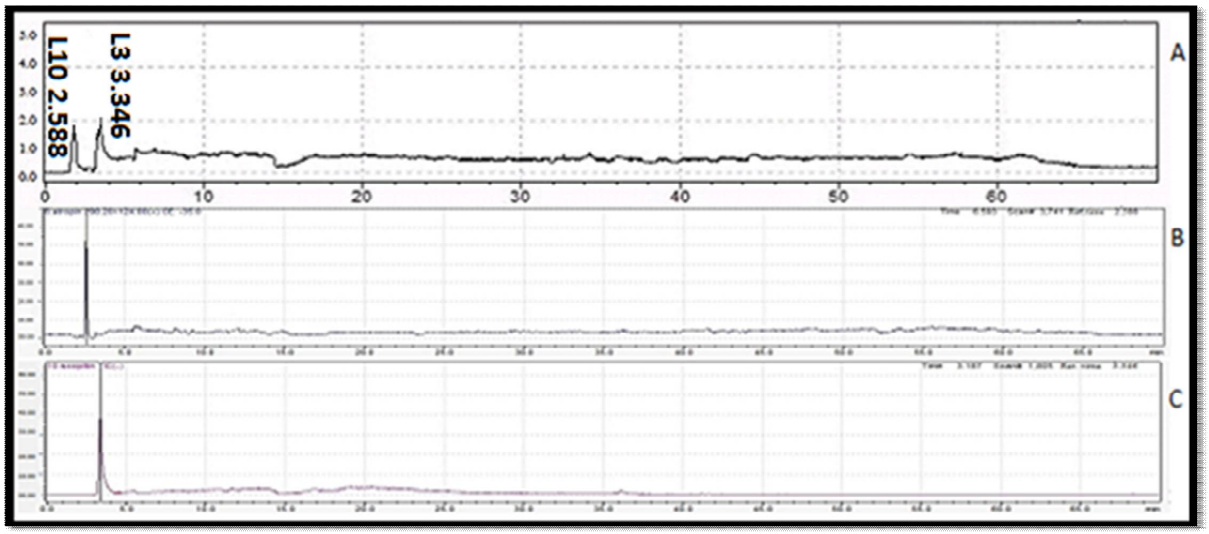


Figure 13: LC-MS chromatogram of chloroform fraction of Awsaj roots (F2), A= TIC =Total ion chromatogram, B= EIC= Extracted ion chromatogram of L 10, C= Extracted ion chromatogram of L 3.

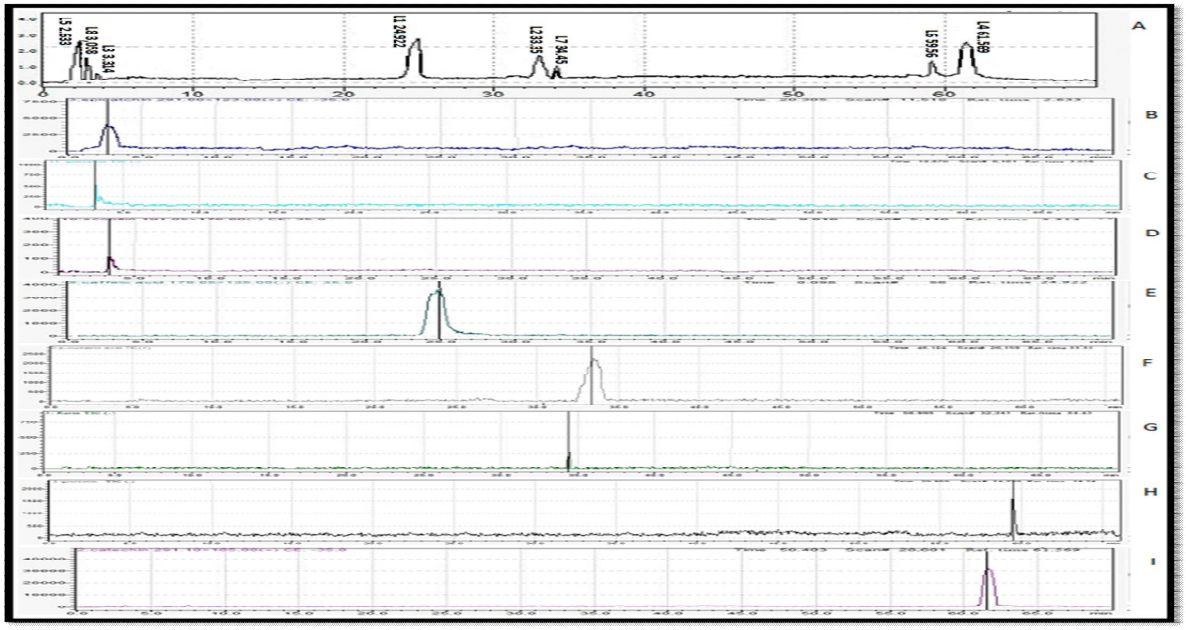


Figure 14: LC-MS chromatogram of ethyl acetate fraction of Awsaj roots (F3), A= TIC, B to I= EIC L 5,L 8, L 3, L 1, L 2, L 7, L 6 and L 4 respectively

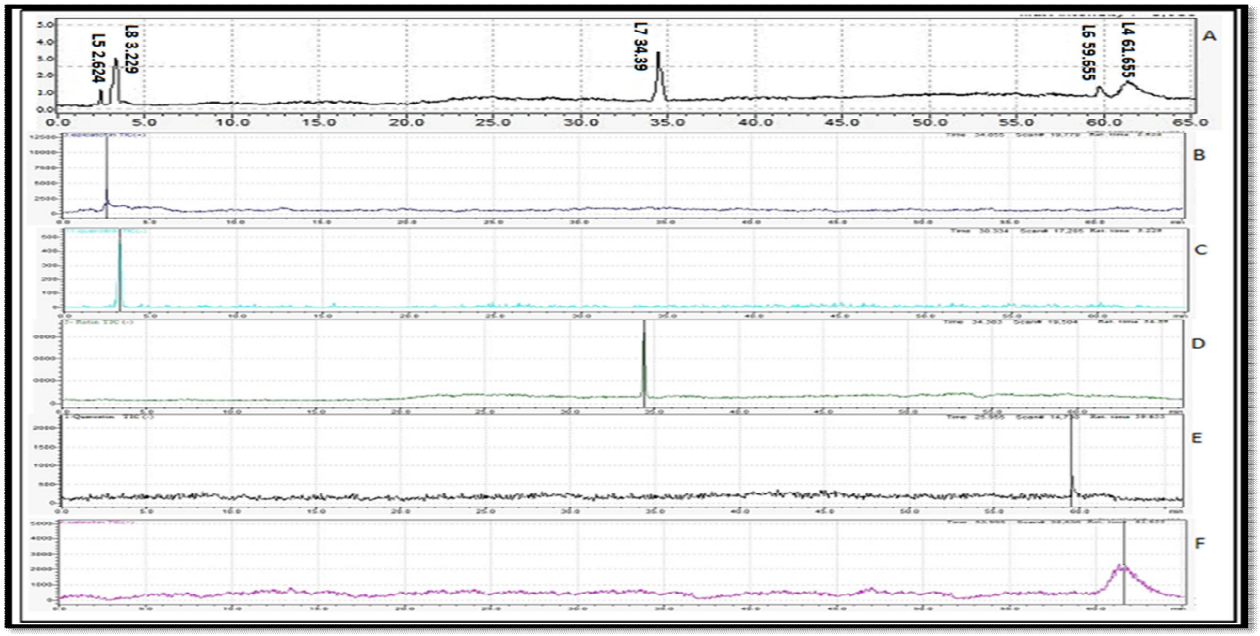


Figure 15: LC-MS chromatogram of n-butanol fraction of Awsaj roots (F4), A= TIC, B to F= EIC of L 5, L 8, L 7, L 6 and L 4 respectively.

Indian Journal of Natural Sciences www.tnsroindia.org.in ©IJONS

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

A Survey on Performance Metrics of Android Apps

G. Mohana Prabha

Assistant Professor, M.Kumarasamy College of Engineering,Karur,TamilNadu,India.

Received: 15 Feb 2018 Revised: 28 Feb 2018 Accepted: 23 Mar 2018

*Address for correspondence
G.Mohana Prabha
Assistant Professor
M.Kumarasamy College of Engineering,TamilNadu,India.
E mail: hodit@mkce.ac.in

ABSTRACT

Smart phones these days have turned out to be basic individual devices to bolster our exercises in practically every part of our lives. On account of the colossal headway of smart phone innovations, stages, and also the eagerness of individual designers, various portable (applications) have been made to serve an extensive variety of use purposes, making our everyday life more advantageous. While these applications are utilized, information logs are commonly created and mood setting is recorded framing a rich information wellspring of the smart phone clients' practices. In this paper, we study existing reviews on performance of smart phone and application designs, launch delay and various delays associated with it. Our exchanges of the reviews are sorted out as indicated by two fundamental research streams, to be specific application time based performance and space based close by a couple of other related reviews.

Keywords: Samrt phone; Launch delay; Performance

INTRODUCTION

With the developing fame of smart phones, individuals these days invest more time with their cell phones than any time in recent days. The use of the smart phones has reached out from fundamental correspondence needs, e.g. sending SMS and making calls, to many high-level needs that cover almost all the aspect of our daily life. Such needs are mostly bolstered by versatile applications (apps)which are particularly composed programming projects to keep running on portable devices such as smart phones and tablets. Major cell phone platforms, for example, Android and iOS give an extensive variety of applications for their clients. Major cell phone platforms, for example, Android and iOS give an extensive variety of applications for their clients. For example, Google Play, the authority application showcase for the Android framework, released more than 1 million applications in 2012. The number of downloads in the year was more than 45.6 billion. This number took off to more than 80 billion in 2013[1].

During the between a user and his smart phone, the information recorded from different sensors contains data on both smart phone use and user vibe setting, for example, area, association status and movement. Both sorts of data

together give an awesome chance to see how people collaborate with their smart phones, their mobility behavior, social connections and individual inclinations in different settings[1].

For smart phone users, existing reviews demonstrate that the number of utilized applications in every users smart phone ranges from 10 to more than 90, with a median of around 50[2]. While these applications are utilized, various metrics are associated with it. In this paper, we review some of the performance metrics associated with the android. Our examinations of the reviews are sorted out as per two fundamental research streams; specifically time based and space based performance of the android app.

Metrics of Smart Phone Apps

Metrics are the standard measurement of anything to be measured and analysed for their quality. It can be categorized as following,[2]

1. Performance metrics: This metrics measures focus on how the user is experiencing the app while they are using.
2. User and usage metrics: This metrics provide visibility into the user and their demographics
3. Engagement metrics: This metrics highlight how users are engaging with the app
4. Security Metrics: This metrics provides how securely app keeps data confidentially.

There few more metrics also associated with Apps but we concentrate only on performance metrics.

Performance Metrics

App Latency: Latency refers to the round trip time from app launch to app response. The thumb rule is to optimize the RTT to maximum of 1 second.

ETE Application Latency: It's not quite recently enough to track API latencies; you likewise require end-to-end reaction time to applications that are controlling the API's. Again the thumb rule is to optimize the RTT to maximum of 1 second.

App Load per Period: This is the time related to the number of calls or transaction over certain measurable period of time.

App Crash: App Crash is nothing but overloading of data that has injected to the application. the typical crash rate should be 1-2%. Different levels of cache becomes filled when multitasking takes place.

Application Interruption and Measuring its latency:

a. When the interfacing with an application Smartphone users may be hindered by expected or unexpected conditions. In such kind of situation , our work is to reveal how to reduce the cost when interruption occurs. Let T_{wo} be the time required by an app to work under normal condition. But due to unintended interruption, currently app affected by the overhead time, T_o and re launch time, T_r. where, $\{T_{wo}, T_o, T_r\} \in R^+$. Now the time has been increase from T_{wo} to T_o+T_r. .

Detecting Interruption

When the changes happening from application A to B, interruption is considered. $x \neq y$. There can be two restriction imposed 1) interruption will not consider for Launcher/Dialer application (L). 2) An interruption will nor affect an application. formally, $L \neq x \neq C$ and $y \neq C$. T_b refers to time taken by app before interruption, T_i refers to time taken by app during interruption T_a refers to time taken by app after interruption

Measuring Interruption

The user is truly exorbitant when the application interruption once in a while occurs on smart phones, yet when they do. It represents an abundance of such new difficulties for mobile designer Smartphone sellers,etc. Recent versions of android such as 6.0 and greater produces the negligible delay of Application Interruption even they are uninterrupted.

Reducing tail time in Android while connected to Data Service

Smart phones with 3G/4G networking frequently squander vitality in the purported "tail time" amid which the radio is continued despite the fact that no communication is happening. A framework DelayDroid, which permits a developer to add the required strategy to existing, unmodified Android (applications) with no human exertion and no SDK/OS changes.

At the point when data is to be exchanged to a base station, a phone's mobile radio (e.g., 3G or 4G radio) works in a powerful mode. At the point when no data. Compile time process initially locate network calls then refactors short-lived network calls followed by long-lived network callsThe run-time process consists of ContextService, DelayController and BatchPolices. ContextServiceis in charge of sending and receiving the Wakeup messages.

ContextService can give the Context Object to an arrangement to permit it to get to data in different settings. At the point when the application begins, setting pertinent classes are instantiated. At that point the instrumented code is executed to instate ContextService. Amid its instatement, Context Service enrolls a Wakeup message beneficiary for the application. At last ContextService introduces Delay Controller.

Delay Controller is accountable for clumping the system solicitations of all DelayDroid-empowered applications on the framework. Amid the introduction of DelayController, it picks a clustering approach as per a design record, and instantiates a case of the strategy.

At the point when data is to be exchanged to a base station, a phone's mobile radio (e.g., 3G or 4G radio) works in a powerful mode. At the point when no data is to be exchanged, the radio will first change to a purported tail mode that expends less vitality, and later will change to a sit still mode that devours even less vitality. The transitional tail mode is a bargain between vitality utilization and quick reaction to new information exchange demands.

We endeavor to naturally apply an arrangement inside the applications through static investigation and bytecode changes. Utilizing DelayDroid, designers can utilize the default bunching strategy or give their own groupingarrangement to spare the vitality squandered in tail time. In any case, regardless it meets the accompanying confinements. It can naturally examine and refactor unpredictable, existing Android applications, and uncover the open door for clumping system asks for crosswise over applications to designers.

Since most of the adequate functionalities are not available, it is difficult to do improvement in the application. It requires the android app with source code, which is very difficult to unpack the proprietary application. To extend, we need to deploy the techniques to individual developers.

The following network classes that need to refactored to reduce the tail-time energy.

Concurrent JavaScript Parsing for Faster Loading of Web Apps

JavaScript is one of the powerful scripting language for dynamic client-side applications. Before the java script is being executed, it must be parsed. So the parsing takes huge amount of time for client side application .The proposed research work deals with multithreading with parallel technique which reduce loading time of the app. When main thread is executed, All other remaining threads are getting parsed. It also ensures the synchronization among threads.

Working of JS Engine

To diminish the overhead, modern JavaScript engine compilation runs in parallel such that to enhance simultaneously with the main thread to shroud its overhead, JITC is performed on separate threads (compilation threads),.

Concurrent parsing approach

NOT-EXISTS: All the parsing request of a function has not been enqueued yet. The main thread will parse the function.
UN-PARSED : All the parsing demands are enqueued, but none of the parsing thread has dequeued and parsing begin for the function yet. When this problem arise ,main thread of the function will instantly cross out it and the string fundamental itself will parse the capacity because speedier than holding up until a parsing string parses the capacity.

IN-PARSING: All the parsing thread was dequeued and parsing started for the function, but the parsing is not yet completed. In this situation, the main thread will wait for some time until the parsing completes.
PARSED – All the parsing threads are completed the function of the parsing. The main thread read all the byte code and it will executed directly. Speculation-based Concurrent Parsing (SCP), Profile-based Concurrent Parsing (PCP), Original sequential parsing (OCP) algorithms are used to compare and evaluate the timings.

For original parsing only main thread gone under execution. For PCP and SCP maximum of four threads will undergo the execution. the reason threads are restricted to four is, we assume that CPU will be having maximum of four cores.

CONCLUSION

In Smart phone era, very low number of research took place regarding Application interruption and reducing the power and performance improvement. In 2012, initial research took place pertaining to application interruption in offline mode, followed by further research took online mode energy saving of the battery consumption and parallelizing the multithread in JavaScript core. Since Android 6.0, application interruption has reduced gradually and supports well optimized power saving of smart phones. By improving JavaScript core by parallelized multithreading, further more energy can be saved as well as performance of the android.

REFERENCES

1. Alexa. 2016. The Top 500 Sites on the Web. Retrieved July 10, 2016, from https://www.alexa.com/topsites.
2. Apple,Inc.2011. WebKit2 -High Level Document.Retrieved July 10, 2016, from https://trac.webkit.org/wiki/WebKit2.
3. Apple, Inc. 2016c. Safari browser. Retrieved July 10, 2016, from http://www.apple.com/safari.
4. Athivarapu P K, Bhagwan R, Guha S, et al. Radiojockey: mining program execution to optimize cellular radio usage. In: Proceedings of the 18th Annual International Conference on Mobile Computing and Networking, Istanbul, 2012. 101–112
5. Back MD, Stopfer JM, Vazire S, Gaddis S, Schmukle SC, Egloff B, Gosling SD (2010) Facebook profiles reflect actual personality, not self-idealization. Psychol Sci 21:372–374
6. Bianchi A, Phillips J (2005) Psychological predictors of problem mobile phone use. CyberPsychol and Behav 8:39–51
7. Biel J, Aran O, Gatica-Perez D (2011) You are known by how you vlog: Personality impressions and nonverbal behavior in youtube. In: Proceedings of international AAAI conference on weblogs and social media (ICWSM)
8. JinSeok Oh and Soo-Mook Moon. 2015. Snapshot-based loading-time acceleration for web applications. In Proceedings of the 13th Annual IEEE/ACM International Symposium on Code Generation and Optimization (CGO'15). IEEE, 179–189. DOI:http://dx.doi.org/10.1109/CGO.2015.7054198
9. Nirmalrani V., Saravanan P., Kalpana S Shesecure safety app–the hexa umbilical cord Pak. J.Biotechnol. Vol. 14 (Special Issue II) Pp. 128-133 (2017)
10. Nikzad N, Chipara O, Griswold W G. APE: an annotation language and middleware for energy-efficient mobile application development. In: Proceedings of the 36th International Conference on Software Engineering, Hyderabad, 2014. 515–526
11. Puustinen I, Nurminen J. The effect of unwanted internet traffic on cellular phone energy consumption. In: Proceedings of International Conference on New Technologies, Mobility and Security (NTMS), Paris, 2011. 1–5
12. Qian F, Wang Z, Gerber A, et al. Profiling resource usage for mobile applications: a cross-layer approach. In: Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services, Bethesda, 2011. 321–334
13. Sowmiya S, Sushma K and Kalpana S, Hospital Information Management using Android International Journal of Pharmacy & Technology 8(1): 11124 –11133(2016).
14. Vergara E J, Sanjuan J, Nadjm-Tehrani S. Kernel level energy-efficient 3G background traffic shaper for Android smartphones. In: Prcoeedings of the 9th International Wireless Communications and Mobile Computing Conference, Sardinia, 2013. 443–449
15. Vergara E J, Nadjm-Tehrani S. Energy-aware cross-layer burst buffering for wireless communication. In: Proceedings of the 3rd International Conference on Future Energy Systems: Where Energy, Computing and Communication Meet.New York: ACM, 2012. 24
16. Xu F, Liu Y, Moscibroda T, et al. Optimizing background email sync on smartphones. In: Proceedings of the 11th Annual International Conference on Mobile Systems, Applications, and Services, Taipei, 2013. 55–68
17. Yeh J-H, Chen J-C, Lee C-C. Comparative analysis of energy-saving techniques in 3GPP and 3GPP2 systems. IEEE Trans Veh Tech, 2009, 58: 432–448

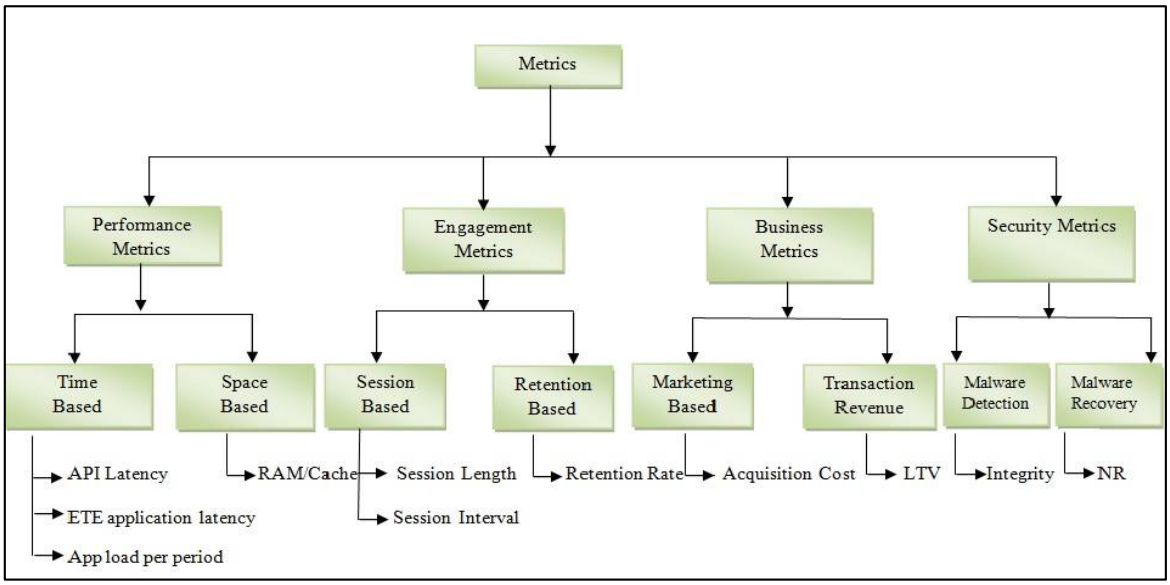


Figure 1.Metrics of Android App

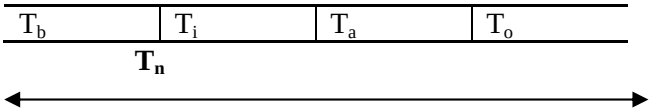


Figure 2. Interruption Measurement

Table.1 Classes to b Refatored

Category	Description
HTTPWebkit	This method is used to send Http request in package Android/webkit
HTTPApache	This method is used to send Http request in package org/apache
HTTPJavaNet	This method is used send Http request in package java/net
TCPSocket	This method is used to send TCP packets, for example, "send" method in class socket
UDPSocket	This method is used to that send UDP packets, for example "send" method in class DatagramSocket

Indian Journal of Natural Sciences www.tnsroindia.org.in ©IJONS

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

An Intelligent Irrigation System Using WSN Adorned with Zigbee Module

S.Kanimozhi[1]* and S.Saravanan[2]

[1]Assistant Professor, Department of Information Technology, M.Kumarasamy College of Engineering, Karur, Tamilnadu India.
[2]Assistant Professor, Department of Computer Science and Engineering, M.Kumarasamy College of Engineering, Karur,Tamilnadu India.

Received: 13 Feb 2018 Revised: 18 Feb 2018 Accepted: 23 Mar 2018

***Address for correspondence**
S.Kanimozhi
Assistant Professor, Department of Information Technology,
M.Kumarasamy College of Engineering,
Karur Tamilnadu India.
E mail: kanimozhis.it@mkce.ac.in

ABSTRACT

In the field of agribusiness, water framework accepts a basic part for the economy and progression of a country. The most ideal strategy for water framework is fundamental for agriculture. This wander is expected to develop a customized water framework structure for trading the draw motor ON/OFF on distinguishing the sogginess substance of the earth. The advantage of using this procedure is to reduce human intervention and moreover used to diminish the wastage of water. In this endeavor by executing GSM development, at whatever point the pump motor switches ON/OFF, a SMS will be passed on to the concerned individual with respect to the status of the pump. We can similarly control the pump through SMS. With the use of negligible exertion sensors and the direct equipment this work focuses ease thing, which can be obtained even by a poor rancher.

Keywords: Temperature sensor; Soil dampness sensor; Water level sensor; PIC microcontroller; ZigBee; Wireless sensor arrange; GSM.

INTRODUCTION

Prior agribusiness relies upon the storm season for water system which isn't the adequate water for the yields so programmed water system techniques were presented. It is troublesome for the ranchers to pick up data in regards to the rainstorm conditions which is useful for the specific product or not. In display age, water deficiency is getting to be noticeably one of the most serious issues on the planet. A wide range of techniques were presented for preservation of water. We require water in every last field. Farming is where water is required in colossal amount. At the present period, In India agriculturists utilize the water system procedure through the manual control. This

procedure here and there expends more water and the water achieves the field late because of which the products get dried. Wastage of water is a noteworthy issue in farming. These days diverse systems are accessible for watering horticultural harvests which is utilized to lessen the reliance of rain. As innovation is enhanced, remote sensor systems can be utilized as a part of agribusiness to enhance trim creation. Remote innovation utilizing distinctive sensors for exactness horticulture has turned into a well known research. The composed paper for programmed water system framework utilizing ZigBee innovation has sensors for estimating the dampness substance of the dirt. This sensor is utilized to begin water system naturally when the limit esteem is low. Checking parameters, for example, dampness substance of soil, temperature and mugginess is essential for acquiring astounding farming condition.

The Related Work

PDAs have almost transformed into a vital bit of individuals which expect a basic part in everybody's life. Veena Divya K et.al(2013) has discussed about the application makes use of the ZIGBEE feature of phone as a response for water framework control system. Mansour, H. A., et.al(2013) worked with a customized control of close circuit spill water framework structure as a balanced water framework system on yellow corn alter vegetative and yield parameters under conditions at Al-Hasa region. The proposed plan incorporates the change of an organized structure to modernize the spill treating water framework in green house. The course of action grasped incorporates a data obtainment card PCL-812PG controlled by PC. Guerbaoui.M (2013) have suggested about the water framework for the harvests is given by a weight driven circuit in perspective of an electric pump. Water needs are picked by evaluating soil water status by soil moisture sensor. Usually hand pumps trench water and precipitation were an important wellspring of water supply for water framework. This strategy has provoked genuine drawbacks like submerged framework, over-water framework and moreover it causes depleting and loss of supplement substance of soil. Robotized water framework structure is a machine based structure, which controls the water arrangement of land by joining diverse programming and gear approaches together for field water framework. This paper deals with a succinct outline of various GSM based robotized develop water framework systems. Purnima et.al(2012) has given about the GSM techniques which serves an indispensable limit since it is responsible for controlling the water framework office and sends them to beneficiary.

DESCRIPTION OF COMPONENTS

Temperature Sensor

LM35 is an IC temperature sensor which is used to check the temperature with its yield relating to the temperature in degree Celsius. The sensor equipment is settled and thusly it isn't subjected to oxidation and distinctive methodology. By using, LM35, temperature can be assessed more absolutely than with a thermistor. It moreover has low self warming. The working temperature extent of LM35 sensor is from - 55°C to 150°C. The yield voltage moves by 10mV on account of each oC rise/fall in encompassing temperature.

Soil Sensor

Soil clamminess sensors is used to measure the volumetric water content in soil by suggestion by using some other property of the earth, for instance electrical insurance, dielectric reliable, or coordinated effort with neutrons as a middle person for the moistness content. Assessing soil sogginess is basic for green applications to empower agriculturists to manage their water framework structures more beneficially.

Humidity Sensor

Mugginess estimation instruments normally in view of estimations of some other amount, for example, temperature, weight, mass or a mechanical change in a substance as dampness is ingested. By adjustment and figuring, these deliberate amounts can prompt a Measurement of moistness.

Water Level Indicator Sensor

Water sensor square is created for water distinguishing proof, which can be comprehensively used as a piece of identifying the precipitation, water level, even the trade spillage. The piece is generally involves three areas an Electronic square connector, a 1 Mω resistor, and a few lines of revealed driving wires. This sensor works by having a movement of gave takes after related ground and between the grounded takes after are the sense takes after. This thing has low power use, and high affectability, which is the key typical for this module.

PIC 16F877 Microcontroller

PIC microcontroller is a champion among the most celebrated microcontrollers used for the advanced reason. PIC stays for Peripheral Interface Controller. PIC is fiscally sagacious and is field programmable. The 16F877A is a capable microcontroller that can do different errands since it has an adequately broad programming memory (gigantic to the extent sensor and control wanders) 8k words and 368 Bytes of RAM. PIC16f877a finds its applications in a huge number of devices. It is used as a piece of various fields like remote sensors, security and prosperity devices, home robotization and in various current instruments. An EEPROM is featured in it which makes it possible to store a bit of the information always like transmitter codes and authority frequencies. The cost of this PIC controller is low, its managing is also basic and it is versatile.

ZIGBEE

ZigBee offers judgments to the devices which have low data rates, eat up low power and has long battery life. The IEEE 802.15.4 standard portrays two layers to be particular, the physical and MAC (Medium Access Control) layers. The physical layer supports three repeat bunches with net data rates: 2.450MHz, 915 MHz and 868 MHz .The target of ZigBee is to give versatility, compactness, security and gigantic framework restrain. ZigBee development is used as a piece of a broad assortment of things and applications across finished business, mechanical, government markets and around the globe.

Proposed System

System that comprises of a sensor module alongside the WSN, used to transmit the information bundles in remote. The dirt dampness sensor and temperature sensor are checking the farming field. At that point this information sent to the microcontroller unit. It shows through the LCD and transmitted to ZIGBEE. The estimation of the sensor range can be seen by the client in two way. One is by means of ZIGBEE and another path is through GSM. The client can control the engine when it is have to turn ON. The insight plot which is screens the dirt dampness and temperature of the horticulture field is a viable approach for the modernized agribusiness likewise it manages to diminish the human power wellspring of the agriculturist and the decrease the water wastage. Due to its vitality self-governance and minimal effort, the framework has potential use for natural harvests, which is essentially situated in geologically disconnected regions where the vitality lattice is far away. There are two primary units. The transmitter (Agriculture Field) and recipient (The User Part). The transmitter segment comprises of the temperature sensor, moistness sensor and water level sensor which is utilized to send their separate esteems to the microcontroller.

CONCLUSION

As indicated by our last outcome we have gotten, an ordinary rancher can interface this setup to PC and furthermore to both advanced cell and cell phone to consequently switch the draw engine ON/OFF for the water system of the harvests. The point of this proposed paper is to get data with respect to the mugginess, dampness substance and temperature in field. From this it is reasoned that the water system framework will controlled through WSN Communication between the sensor hubs and the information collector is by means of the ZigBee.

REFERENCES

1. Alaa Salih Ati, Abdualkareem Hassan, Sana Abd-Aljabar, Ammar Salah Role Of Bio Fertilization On Wheat And Water Productivity Under Water Scarcity, Pak. J. Biotechnol. Vol. 14 No-4 (2017).
2. Anwaar Arafa1 , Soha Farag2 , Manal A. Mahdaly1 , Mai M. Badr1, Biodegradation Of Petroleum Industry Oily Sludge By Bacterial Consortium And Its Application In Land Farming, VOL-13 NO-1 (2016).
3. Erdem.Y, L. Arin, T. Erdem, S. Polat, M. Deveci, H. Okursoy, and H. T. Gültas, Crop water stress index for assessing irrigation scheduling of drip irrigated broccoli (Brassica oleracea L. var. italica), Agriculture Water Management, Dec. (2010).
4. Guerbaoui.M, GSM based automated drip irrigation system, IOSR Journal of Mechanical and Civil Engineering - IOSR-JMCE (2013).
5. Jury W. A. and Vaux H. J., The emerging global water crisis: Managing scarcity and conflict between water users, Sep. (2007).
6. Mansour, H. A., and Yousif El-Melhem Impact The Automatic Control Of Closed Circuits Raingun Irrigation System On Yellow Corn Growth And Yield International Journal of Advanced Research (2013).
7. Muthukumar.P, Sakthivel E, Remote Sensing And Gis-Multicriteria Decision Analysis For Ground Water Resources Mapping Of Thalavadi Block - Erode District In Tamil Nadu, Pak. J. Biotechnol, Vol-14 No-3 (2017).
8. Nemali K. S. and Van Iersel M. W., An automated system for controlling drought stress and irrigation in potted plants, Nov. (2006).
9. Pravina B. Chikankar,Deepak Mehetre, Soumitra Das, An Automatic Irrigation System using ZigBee in Wireless Sensor Network, International Conference on Pervasive Computing - ICPC (2015).
10. Purnima, S.R.N Reddy, "Design of Remote Monitoring and Control System with Automatic Irrigation System are using GSM-Bluetooth" International Journal of Computer Applications - IJCA (2012).
11. Veena Divya K , Ayush Akhouri , Chandan Kumar,Raunak Rishabh , Rochak Bagla "A Real time implementation of a GSM based Automated Irrigation Control System using Drip Irrigation Methology" International Journal of Scientific & Engineering Research,May (2013) .
12. Wang X., Yang W., A. Wheaton, Cooley N., and Moran B., "Efficient registration of optical and IR images for automatic plant water stress assessment," Nov. (2010).
13. Yuan G., Luo Y., Sun X., and Tang D., "Evaluation of a crop water stress index for detecting water stress in winter wheat in the North China Plain", Jan. (2004)

Fig 1. LM35 temperature sensor

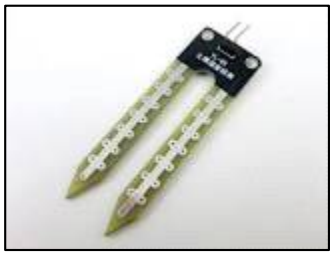


Fig 2. Soil moisture Sensor

Fig 3. Humidity sensor

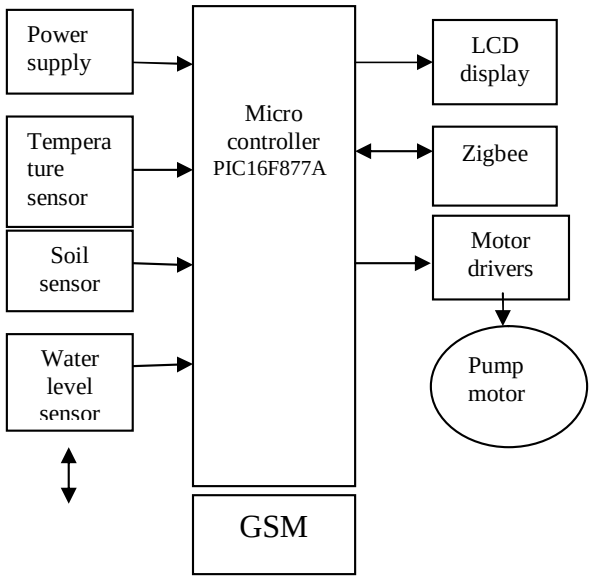


Fig 4. Transmitter section

Fig 5. Receiver section

Indian Journal of Natural Sciences www.tnsroindia.org.in ©IJONS

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Student Dropout Prediction Using Improved Decision Tree in EDM

M.Gunasekar[1*] and Shamugavelyutham.A[2]

[1,2]Assistant Professor, Department of Information Technology, M.Kumarasamy College of Engineering (Autonomous) Karur, Tamil Nadu,India.

Received: 10 Feb 2018 Revised: 19 Feb 2018 Accepted: 23 Mar 2018

[*]**Address for correspondence**
M.Gunasekar
Assistant Professor,
Department of Information Technology
M.Kumarasamy College of Engineering (Autonomous)
Karur, Tamil Nadu,India.
E mail: guna18it@gmail.com

ABSTRACT

Educational Data mining is an emerging research area where new methods are developed to predict the performance of the students. This work is about to predict the dropouts from school. Data has been collected from the school, the attributes of the dataset contains the demographic and academic data of students. Improved decision tree classifier algorithm is used to construct the model. This uses different algorithms for splitting a node into parts. Splitting is done based on all possible attributes and selection of attribute depends on the homogeneity of sub-nodes. The attributes are selected using the Information Gain and association function. Over-fitting is a concern in decision tree; this can be evaded using entropy and information gain. The proposed model has been applied on the student data and the results shows that the accuracy of the improved Decision tree algorithm outperforms other traditional models.

Keywords: Educational Data Mining, classifier, Entropy, prediction, Information Gain,

INTRODUCTION

Educational Data Mining is an emerging field of research, which does data mining on the largely evolving educational data's. Machine Learning algorithms are applied on these large datasets to study and improve the techniques practiced for betterment of the student. The data's for EDM comes from various sources such as from student database collected in the institution, interactive learning environment, online games, survey etc., the objective of EDM is to analyze the data and dig out the significant knowledge. There are number of factor which decides the future of youngsters. Since the current world is technology driven, we strongly believe that education is mandate for all to understand the living environment. The aim of this work is to predict the dropout students at the early stage.

The Machine Learning algorithm are good in making sense with the data we have, we apply one such algorithm for our work. The Decision Tree algorithm is deployed in our analytics. Though different classifiers are available, Decision Tree outperforms all of them. Logistic Regression, Naïve bayes also offers better result for a classification problem but Decision Tree performs well when dependent variable is categorical.

Related Work

A valuable amount of research works based on classifier and predictive model are developed to predict the performance of students. Pandey et al., (2016) (developed an integrated classifier model which combines the Decision Tree, K-NN and AODE. This model uses heterogeneous multiple classifier which combine three classifiers. Saurabh pal et al., (2017) defined BF Tree algorithm which is comprised of four decision Tree Algorithms. This data mining technique classifies the performance of the students. Paul cortez et al., (2008) used business intelligence and data mining techniques to predict the performance of higher secondary school students. They used four classifier models in their work namely Decision Tree, Random Forest, Neural Networks and SVM. Sivakumar et al., (2016) in their work they improved ID3 Algorithm using Renyi entropy, Information gain and association function for predicting the student dropout in Higher Education.

Proposed Work

The model starts with the data collection, pre-processing, transformation and splitting the data into train and test data. R Environment has been used for this work. In R the data's are divided into essential parts using seed and sample methods. The model is constructed based on Improved Decision Tree classifier. The data set was taken from a school in India. This contains demographic information of school students and it includes 25 attributes and 395 records. In the available attributes 6 attributes are considered for the analysis.The traditional Decision Tree algorithm is improved using Association factor and normalized information gain. At the start entropy for the class label is being calculated and based on this entropy value the Information gain for every attribute is derived. The attribute which has high information gain is selected as a root and this process is repeated until all the nodes are plotted in the Decision Tree. Figure 2. Shows the decision tree constructed from our model.

Features selected by Correlation Feature Selection

Medu - mother's education
Fedu - father's education
Mjob - mother's job
Fjob - father's job
schoolsup - extra educational support from school
Failures - number of past class failures

CONCLUSION AND FUTURE WORK

This paper proposed a classifier model to predict the student drop out in schools. The improved Decision tree algorithm illustrates its ability to enhance the decision trees and shows its classifier accuracy. The root of the tree is selected based on the feature selection done with the help of entropy and Information Gain. The entropy is calculated at every levels of the tree. The results has proven that the improved Decision tree algorithm out performs the traditional classifier models. Though the model has been constructed for predicting the drop out in schools, this can also be enhanced to apply for the higher education data sets. It can be further enhanced to decision support systems.

REFERENCES

1. Anitha and Nivetha Review Extractor and Performance Analyzer by Opinion Mining Techniques Pak. J. Biotechnol. Vol. 13 Pp. 428- 430 (2016).
2. Cortez, Paulo, and Alice Maria Gonçalves Silva. Using data mining to predict secondary school student performance. (2008).
3. Du W, Zhan Z. Building decision tree classifier on private data. InProceedings of the IEEE international conference on Privacy, security and data mining-Volume 14 (pp. 1-8). Australian Computer Society, Inc. 1 Dec (2002).
4. Heiner C, Heffernan N, Barnes T. Educational data mining. In Supplementary Proceedings of the 12th International Conference of Artificial Intelligence in Education Jul (2007).
5. Pal, Saurabh, and Vikas Chaurasia. Is Alcohol Affect Higher Education Students Performance: Searching and Predicting pattern using Data Mining Algorithms. (2017).
6. Pandey M, Taruna S. Towards the integration of multiple classifier pertaining to the Student's performance prediction. Perspectives in Science. 8:364-6. Sep (2016).
7. Ramaswami M, Bhaskaran R. A study on feature selection techniques in educational data mining. Journal of Computing.1(1):7-11 (2009).
8. Ranka S, Singh V. CLOUDS: A decision tree classifier for large datasets. InProceedings of the 4th Knowledge Discovery and Data Mining Conference (Vol. 2, No. 8) Aug (1998).
9. Rastogi R, Shim K. PUBLIC: A decision tree classifier that integrates building and pruning. InVLDB (Vol. 98, pp. 24-27) 24 Aug (1998).
10. Romero C, Ventura S. Educational data mining: A survey from 1995 to 2005. Expert systems with applications. 33(1):135-46. Jul (2007).
11. Safavian SR, Landgrebe D. A survey of decision tree classifier methodology. IEEE transactions on systems, man, and cybernetics. 21(3):660-74 May (1991).
12. Sivakumar S, Venkataraman S, Selvaraj R. Predictive modeling of student dropout indicators in educational data mining using improved decision tree. Indian Journal of Science and Technology. 9;9(4). Jan (2016).
13. Vinothan Damodaran and Saravanan Institution System Analysis By Using Similarity Based Clustering on Social Network Access Pak. J. Biotechnol. Vol. 13 Pp. 1-4 (2016).

Table 1: Cause of student dropout

Reasons	Percentage
mothers education	4.18
mother's job	4.16
father's education	2.43
father's job	1.54
Failures	0.65

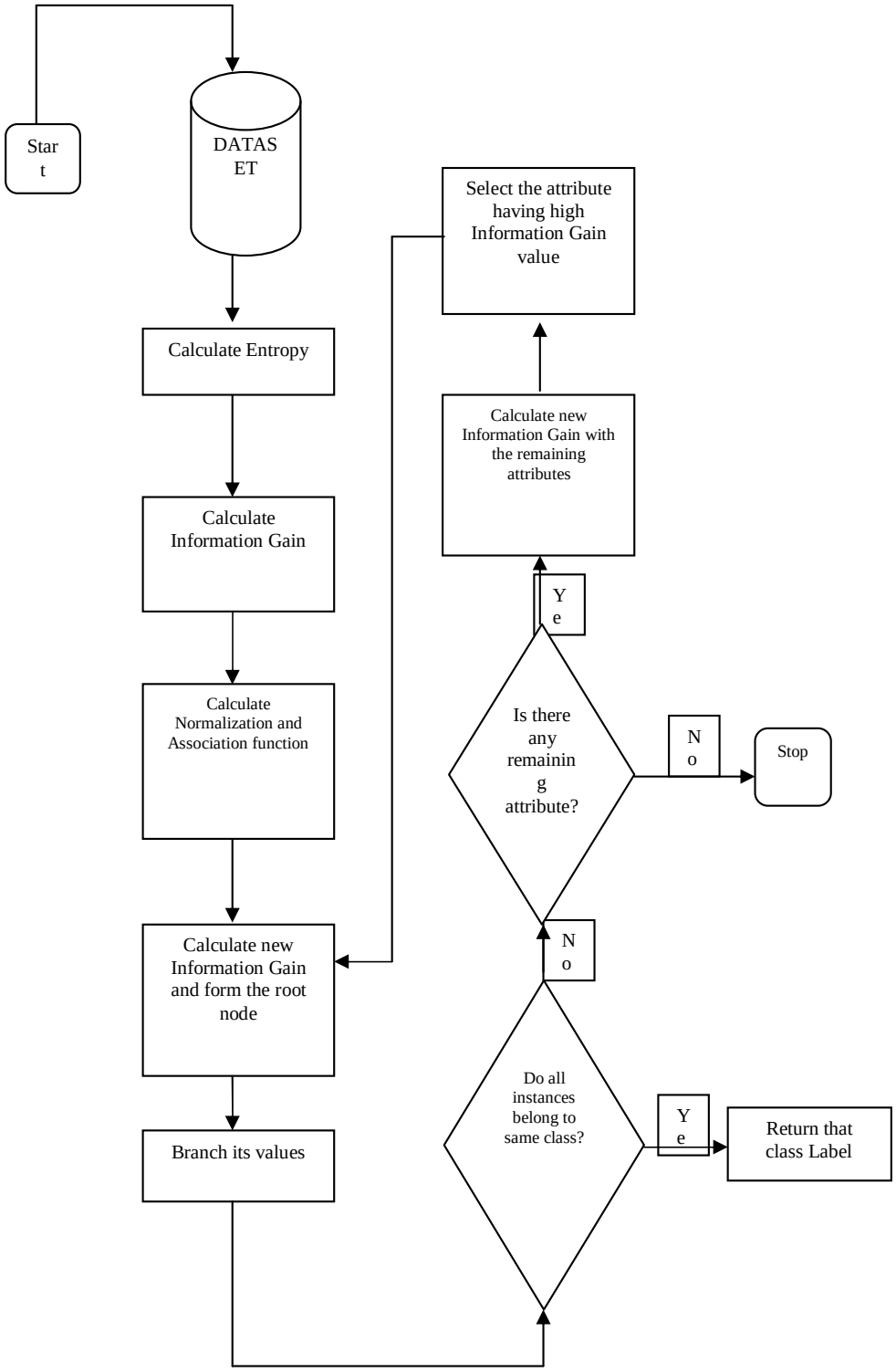


Figure 1: Decision Tree Algorithm

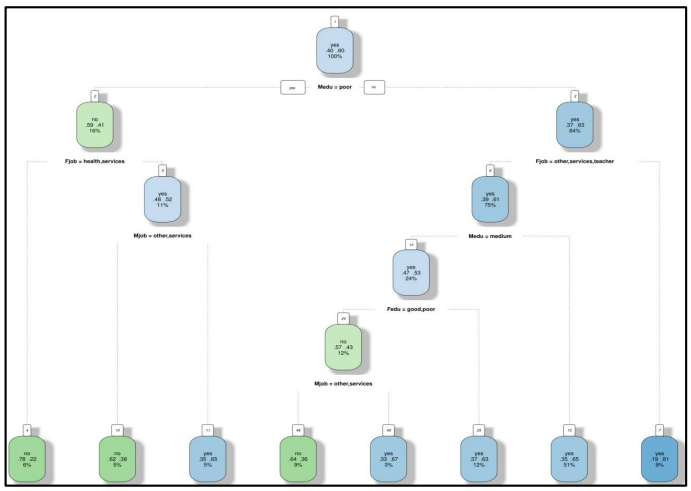


Figure 2: Decision Tree Model

Indian Journal of Natural Sciences *www.tnsroindia.org.in ©IJONS*

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Efficient Query Processing for Recommender System Using Skylines

M.Sindhuja*

Assistant Professor, Department Computer Science and Engineering, M.Kumarasamy College of Engineering, Karur, Tamilnadu,India.

Received: 08 Feb 2018 Revised:18 Feb 2018 Accepted: 23 Mar 2018

***Address for correspondence**
M.Sindhuja
Assistant Professor, Department Computer Science and Engineering,
M.Kumarasamy College of Engineering,Karur,
TamilNadu,India.
E mail: Sindhuvinith12@gmail.com

ABSTRACT

Data mining is the process of analyzing raw data from different ways and processing it in to useful information it can be used to boostup the turnover and also decrease in expenditure. Recommender system is an information filtering system which is used to keep track of the rating or preference of a user item. Recommender systems are used in applications that help the people in making decisions while choosing a product.A recommender system evaluate the user daily preferences by seeing to the user profile to some reference and suggest a personalized suggestion. These recommendations may be from the user referenced item or the user's societal environment or a combination of both. But in the recommender system there occurs a problem of long tail problem, in which the unpopular data that have low number of rating will be moved to the tail of the product distribution, these types of product should not be avoided, it is also utilized in an efficient manner.Hence to overcome the limitation of the existing system, a skyline query processing approach which is known to take into account the multi-dimensional data that has very little popularity will be proposed. Using skylines the users can be recommended with new or unpopular recommendations which solve the long tail problem.

Keywords: Data mining, Recommender systems, societal environment.

INTRODUCTION

Recommendation systems leads to a split up of two datasets,one is MovieLens and another is BookCrossing**.** MovieLens datasets has the customer attributes like age ,gender and also the product based attributes like movie year and general attributes. BookCrossing datasets consists the customer attributes like location, age and the product related variables such as author information, publisher, publication year.. Atleast 10 ratings for the product is

selected in order to carry out cross justification with 10 cross fold value and also the blank values of rating data will be eliminated.

Some of the recommender systems take no notice of unpopular products or newly introduced stuff having low ratings and focus only on the product that reach the rating of minimum set limit items of recommendation algorithms. Instead, such detested or newly introduced objects that remain in the system but it might require special conduct using various cold start methods. Detested or newly invented products leads to the long tail during product supply. In account of widespread of research in the Long Tail phenomenon, the product with low rating should nout be left unutilized it should also reach to the hands of the people in an effective manner.In such cases,likely to split the product into a head and as well as tail parts group items in the tail parts using certain clustering methods.

The cold start problem of the product exists in long tail with low ratings. A accepted solution to the cold start problem utilize the methods based on with the tail item inferred to be similar on the head items on their preferences,where grouping is made to avoid the sparsity problem for collaborative filtering.and the products are clustered in collaborative process to improve the prediction accuracy in reteriving the products.

LITERATURE SURVEY

Recommender systems goal is to gain the fame and consequence. The rise of various recommender systems lead to the complexity of originating the high-quality recommender system. Algorithm of recommender system is to meet needs of the users. So far the investigation carried out has been focused on civilizing the correctness of recommender systems.

The recommender system is supposed to shift away the predictable exactness measure and acquire some other criterion such as version, such as reporting, assortment, destiny, scalability, adaptability, risk, novelty and so on.The data results for the experimental process of velo leads to the people intersest on various aspects tend to prefer various algorithms. The various use of evaluating performance measure of the algorithm performance to improve searching efficiency and make it to be a meaningfull thing. E-Commerce has proliferate in terms of diversity and measure, the user at the terminal end expend the duration to review their products and services.Recommender system can be avail in various application that render the user to a enormous compilation of substance those systems will make the user with the preferred item with the prediction of an item according to the needs.The recommender systems use the fine founded and with effective algorithms. These algorithms also diverge with respect to their strength and flaw. Thus the users come across with choice for the choice of the most successful fact.

The majority of the algorithm at present focus on civilizing the exactness of the recommender system, nevertheless provide accuracy alone is insufficient.some of the recommendation algorithm based on separate item token recommendation. Some particular and narrative appraisal systems are selected and redefined to estimate our recommender system.

A recommender system keeps track of the user profile to some extent to know about the user preferences inorder to suggest the products according to it. Thus the uniqueness may be from the information item or a combination of both .Latest investigation shows that collaborative recommender systems are highly at risk to profile inoculation attack. Security mechanism are needed for shielding the recommender systems beside these attack.The multi agent system that uses the concept of Aspect Oriented Programming for structure protection part ensuring the protection in recommender system using a conservative manager based approach consequences not alone with the problem of system dispersion and policy tangle, but also leads to a bad security status. The security used in the recommender system is used to remove diffusion and tangle issues.

Recommendation algorithm is used to prevent the virus affected user profile even if predisposed profiles will be injected into the system, these cant be fraction of recommendation production. In our proposed approach, it explores the various securities relevant to the crosscutting concern, their surface property on modularity and build a safety aspect. This defense aspect is the incorporation of type-based and policy-based precautions strategy.

Evaluation of Recommender System to Detect the Long Tail Problem

The recommendation method builds the various data mining models for individual product to evaluate the undertermined rating. The each item method will not cluster the product with the other similar ems at all and it builds a various projecting model with the help of individual data item.

The each method method build a prognostic replica for every movies using the ratings of that particular movie and if the movie is about the toy story and then it can built a linear weakening model to forecast the unidentified ratings for that movie, here one can make use of the RMSE to measure the performance of the model and can able to apply 10 foldcross justification to compute RMSE for that particular movie the process should be repeated to N number of times for movie accord to the Movie by the usage of the movie lens. A long tail product with the low ratings is the drawback of the each method model. Hence the prognostic model for the tail items can be learned from a low training example through the each item method.

The LRTP problem occurs due to the lack of data to build good predictive models in the tail, and so that the clustering items can be the reasonable solution. The total clustering recommendation method clusters the whole item set i into different groups by applying conventional clustering methods which can be for example the k-means clustering and building rating-predictive models for each resulting group.

The TC method clusters 3260 movies into 100 groups by means of k-means clustering method and propose a predictive models. The unknown rating of the movie can be predicted by the other Boleyn girl for customer C, then the TC method first determines from where of the available 100 groups of that movie cause from. If the movie belongs to group G5, which includes 30 other moving having 10,000 transactions among them, then this method applies the SVM method to group G5 and computes RMSE Error rates by the fold cross-validation on these 10,000 ratings. This process will be repeated for 10 times on the Movie Lens data for all the available clusters.The clustering method is used to clusters the item along with the similar items when it has only a little amount of data, when it has a considerable amount of data then it will groups to a lesser extent or does not group at all. In the case of the Movie Lens dataset all the available movies are ordered based on the popularity for them. Then, the popularity of the each movie is compared with the criterion number of ratings consider that to be a. If it is larger than a, then the AC method does not apply any clustering method; instead, it keeps only the basic EI approach. However, if it is smaller than a, then the AC method clusters the movie with other similar movies one by one till the resulting group size reaches a. After that, the AC method builds rating predictive models using the resulting group for each item.

The movie secret sunshine which has only 50 ratings, then the AC method groups it with the most similar movie Beijing Bicycle, which has 35 ratings. Even though, the group size is still less than the criterion 100. Thus, the AC method with automatically finds the next similar movie poetry, which has 40 ratings, and then the group size becomes 125(50+35+40), which is larger than the criterion number 100. Next, the AC method builds the predictive model for the movie secret sunshine using the 125 ratings in cluster group. In this way, it can be able to build predictive models for 3260 movies in the Movie Lens dataset. If it is needed to predict the unknown rating of the movie secret sunshine for customer C, then the AC method finds the predictive model build only for the movie and estimates the rating using it.

Skyline Processing

Skyline Query Processing solves the limitation of the long tail problem for the effective use of the resources. Thus the skyline processing are needed to handle the long tail issues and make the users to handle the large amount of data by identifying the products without any negative oppositions or feedback. Using skylines the user can now be recommended with new or unpopular recommendations, which solves the long tail problem and hence detect the scarcity problem by providing sufficient amount data to build the good predictive model and it improves scalability and accuracy of the information reterival by the user.

According to results of various analysis, companies like Amazon that apply long tail process successfully make most of their profit not only from bestselling products, but from the long tail. Hence the skyline process solves the LTRP and makes the recommender system to be more robustness and diversity.

RESULTS

The skyline query processing is used to cluster item according to the popularities which makes the products with low rating to be available for the choice of customer in head items, thus solves the LTRP problem.

CONCLUSION

The recommender systems used to disregard the disliked or new products that bought to the market with low customer ratings and turn up its attention to focus on the products that satisfies the maximum customer ratings in high ,thus the unfavoured products remains in the system needs a special attention to hit the ratings or it should be referenced along with the head items based on the similarties of customer searching database To solve this issue, the skyline query processing is been identify to evaluate every dimensions of the database and to handle d-dimension and select points which are not conquered by any other point in the database. Skyline can support quries that have specific interest in different subsets of dimensions.

REFERENCES

1. Yoon-JooPark, “The Adaptive clustering method for the long tail problem of recommender systems” IEEE Transcation on knowledge and Data Engineering, 2013.
2. Adomavicius.G and Tuzhilin.A, “Expert-driven Validation of Rule-based User Models In Personalisation Application “IEEE Trans Data Mining and Knowledge Discovery,2005.
3. Adomavicius.G and Tuzhilin.A, “Toward the Next Generation of Recommender Systems: A Survey of the State of the Art and possible Extensions” IEEE Transaction on Knowledge and Data Engineering,2006.
4. Anderson.C, “The Long Tail.New York: Hyperion Press”,2007.
5. Bell.R.M and Yehuda.k, “Improved Neighbood – based collaborative Filtering” KDD Cup07,Calofornia,2007.
6. Fleder.D.M and Hosanagar.k, “Blockbuster Cultures Next Rise or Fall; The Impact of Recommender Systems on Sales Diversity”, NET Institute Working Paper No.07-10,2008.
7. Heli,JaesooYoo, “An Efficient Scheme for Continous Skyline Query Processing OverDynamic Data Set” , IEEE Transaction, 2014.
8. Hervas-Drane.A, “Word of Mouth and Recommender Systems: A Theory of the Long Tail”NET Institute Working Paper No.07-41,2007.
9. Linden.G, Smith.B, and York.J, “Amazon.com Recommendations: Item-to-Item collaborative filtering” , IEEE Internet Computing, 2003.
10. Park,Y.J,Tuzhilin.A, “ The Long Tail of Recommender systems and how to leverage It” , Proceeding of the 2008 ACM conference on Recommender Systems, Lausanne Switzerland, PP.no 11-18, 2008.

11. PunamBedi, Sumit Kr Agarwal, " Managing Security in Aspect-oriented Recommender System", IEEE Transaction, 2011.
12. Smola.A.J, Schoelkopf.B, "A tutorial on support vector regression, Statistics and computing", 2006.
13. Truong.K.Q.Ishikawa.F, and Honiden.S, Improving Accuracy of Recommender Systems by Item Clustering", IEICE Transaction on Information and Systems, 2007.
14. wittenL.H and Frank.E, "Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations" . Burlington, MA: Morgan Kaufmann, 2005.
15. Xiao Yong Li, Yijie Wang,Xiaoling Li, Guangdong Wang "Skyline Query Processing on Interval Uncertain Data" , IEEE Transaction on Knowledge and Data Engineering, 2012.

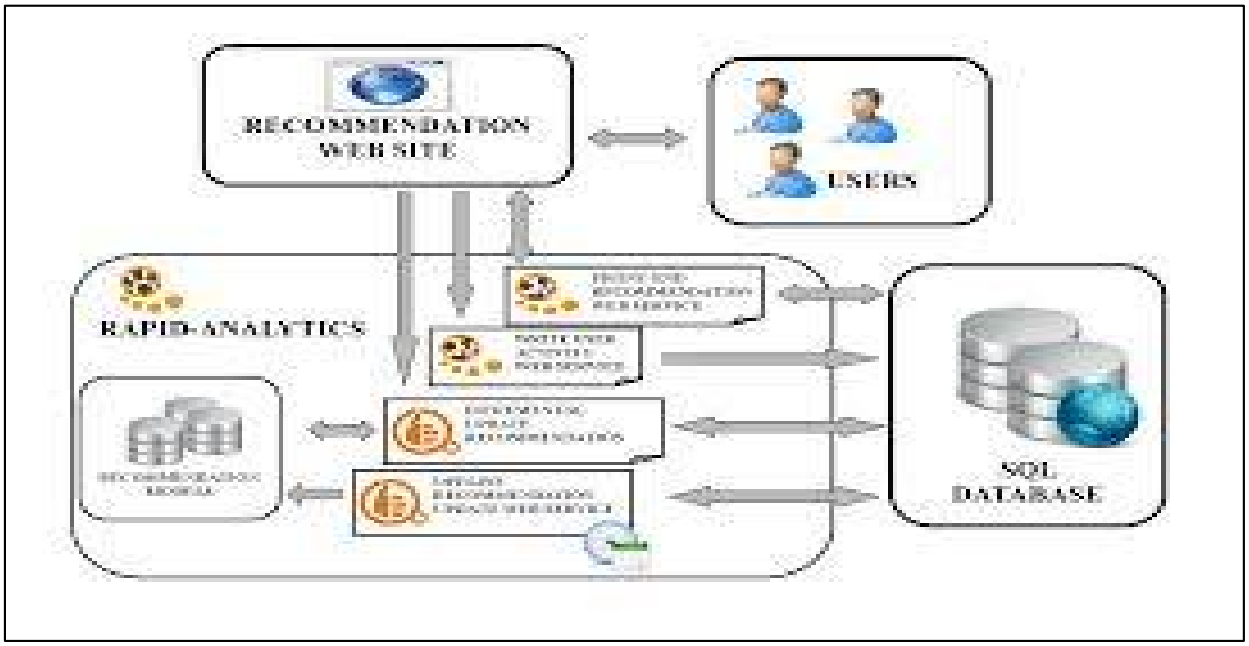


Figure 1: Recommender System

Indian Journal of Natural Sciences www.tnsroindia.org.in ©IJONS

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Stellar Reaction Rate Calculation for CNO-cycle Using non-ESM State Density for the Rigel and Procyon

Abdullah Ali Abd Al Kareem[1*] and Ahmed A. Selman[2]

Department of Astronomy and Space, College of Science, University of Baghdad , Iraq

Received: 12 Feb 2018 Revised: 21 Feb 2018 Accepted: 14 Mar 2018

[*]**Address for correspondence**
Abdullah Ali Abd Al Kareem,
Department of Astronomy and Space, College of Science,
University of Baghdad,Iraq
E mail: alrawi_90@yahoo.com

ABSTRACT

In this research, was calculated the state density of nuclear for nuclear reaction inside a two star nucleus the main sequence is different. This star is: the Rigel and Procyon, This star was selected as examples for the CNO-cycle. Therefore the elements selected for the current account oxygen - carbon - and nitrogen. Calculated the state density based on theory of exciton model of the nuclear reaction. It was also calculated cross-section and emission rate and extracts values from figures for cross-section and then calculates the reaction rate for those nuclei. The results were confirmed any change in configuration will result to different values from state density, the results showed that energy dependence is not linear, and found higher density in high energies but the change was less.

Keywords: Rigel and Procyon, CNO-cycle, nuclear reaction.

INTRODUCTION

In the exciton model, the nuclear state is characterized at any moment during the reaction by the total energy E_{tot} and the total number of particles above, and holes below the Fermi surface. Particles (p) and holes (h) are indiscriminately referred to as exciton. Furthermore, it is assumed that all possible ways of sharing the excitation energy between different particle-hole configurations with the same exciton number n = p + h have equal a-priori probability. To keep track of the evolution of the scattering process, one merely tracks the temporal development of the exciton number, which changes in time as a result of intranuclear two-body collisions. The basic starting point of the exciton model is a time-dependent master equation, which describes the probability of transitions to more and less complex particle-hole states as well as transitions to the continuum (emission). Upon integration over time, the energy-averaged emission spectrum is obtained. These assumptions makes the exciton model amenable for practical calculations. The price to be paid, however, is the introduction of a free parameter, namely the average matrix element of the residual two-body interaction, occurring in the transition rates between two exciton states. When this matrix element is properly parameterized, a very powerful model is obtained.[1]

Theory

The single-particle level density, *g*, is the key at which the expression of the state density depends. Earlier attempts were to consider variable Fermi level *F* [2] or more free energy dependence on *u* below and above Fermi surface [3,4]. Kalbach [5] discussed this dependence in some details and the conclusion was made is that, regardless the specified type of the potential well, *g* is expected to vary between that of the simple square well potential to the simple harmonic oscillator. A proper definition of *g* determines the state density type. Consider one-component system then in the FGM, *g* is given as,[1]

$$g(\varepsilon) = g_o \sqrt{\frac{\varepsilon}{F}} \qquad (1),$$

For particles and holes, this equation is,

$$\left.\begin{aligned} \varepsilon &= F + u_p \\ \varepsilon &= F - u_h \end{aligned}\right\} \qquad (2),$$

Eq.(2) simply means that particle (or holes) energies are above (or below) Fermi surface by the energy u_p (or u_h). Thus,[1]

$$\left.\begin{aligned} g_p(u_p) &= g_o \sqrt{1+\frac{u}{F}} \\ g_h(u_h) &= g_o \sqrt{1-\frac{u}{F}} \end{aligned}\right\} \qquad (3).$$

The state density then can be found from the relation [1],

$$\omega_1(p,h,E) = \frac{1}{p!h!} \int_0^\infty du_1^{(p)}\, g_p(u_1^{(p)}) \int_0^\infty du_2^{(p)}\, g_p(u_2^{(p)}) \ldots \int_0^\infty du_p^{(p)}\, g_p(u_p^{(p)}) \times$$

$$\int_0^\infty du_1^{(h)} g_h(u_1^{(h)}) \int_0^\infty du_2^{(h)} g_h(u_2^{(h)}) \ldots \int_0^\infty du_h^{(h)} g_h(u_h^{(h)}) \delta\left(E - \sum_{\lambda=1}^{p} u_\lambda^{(p)} - \sum_{j=1}^{h} u_j^{(h)}\right) \quad (4),$$

Then, the final equation of non-ESM state density is given by,[1]

$$\omega_{non-ESM} = \delta(E - \sum_{\lambda=1}^{p} u_\lambda - \sum_{j=1}^{h} u_j) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk\, \exp\left[ik\left(E - \sum_{\lambda=1}^{p} u_\lambda - \sum_{j=1}^{h} u_j\right)\right] \qquad (5).$$

in this work, equation(5) will be applied for stellar nuclear reactions based on CNO-cycle process inside the core of the Rigel and Procyon.

The Emission Spectrum

A very important quantity required from PE calculations is the energy spectrum of the emitted particle or type ν, that is: $I_\nu(\varepsilon) = (d\nu/d\varepsilon)_\varepsilon$. This is the goal of the preequilibrium statistical model. It was first Griffin's aim to calculate

this quantity at intermediate energies where other models failed. The significance of PE models all lay in the calculation of energy spectrum of emitted particles at these energies[1].

The task of calculating the energy spectrum at intermediate energies was the motivation for the present development of the preequilibrium statistical models. Important advances have been made and in all cases the need was to better fulfill and reproduce the experimentally measured spectra. Thus, various approaches have been proposed that describe the PE spectra[1].

In the exciton model, the energy spectrum is obtained by summing over all available emission rates, $W_\nu(n,\varepsilon)$, and weighting each term of the summation by the time-integrated probability $P(n,t)$ at each state. The energy spectrum is one of the quantities that are measured experimentally during nuclear reactions; therefore, comparing the experimental spectra with the calculated ones will indicate the validity of the present model[1].
Thus, in the exciton model, PE spectrum $I_\nu(\varepsilon,t)$ is given as [6],

$$I_\beta(\varepsilon,t)\,d\varepsilon = \frac{(2s_\beta+1)}{\pi^2\hbar^3}\mu_\beta\,\varepsilon\,\sigma_\beta(\varepsilon)\,d\varepsilon\sum_{\substack{n\\ \Delta n=2}}T(n,t)\,\frac{\omega(p_\pi-Z_a,h_\pi,p_\nu-N_a,h_\nu,U)}{\omega(p_\pi,h_\pi,p_\nu,h_\nu,E)} \quad (6)$$

This formula describes the exciton model entirely. Physically it means that the measured PE spectrum composes from individual emission rates of different exciton states, weighted by the equilibration time of each state[1].

The CNO cycle

The CNO cycle is the combination of two independent cycles: the CN cycle and the NO cycle. The presence of some isotopes of C, N or O are necessary for either cycle to begin. Being both produced and destroyed during these cycles, these elements act as catalysts. The reaction networks involved are the following:[7]

CN cycle	**NO cycle**
$^{12}C + {}^{1}H \rightarrow {}^{13}N + \gamma$	$^{15}N + {}^{1}H \rightarrow {}^{16}O + \gamma$
$^{13}N \rightarrow {}^{13}C + e^{+} + \nu_e$	$^{16}O + {}^{1}H \rightarrow {}^{17}F + \gamma$
$^{13}C + {}^{1}H \rightarrow {}^{14}N + \gamma$	$^{17}F \rightarrow {}^{17}O + e^{+} + \nu_e$
$^{14}N + {}^{1}H \rightarrow {}^{15}O + \gamma$	$^{17}O + {}^{1}H \rightarrow {}^{14}N + {}^{4}He$
$^{15}O \rightarrow {}^{15}N + e^{+} + \nu_e$	
$^{15}N + {}^{1}H \rightarrow {}^{12}C + {}^{4}He$	

In the CN cycle the isotopes of C and N act as catalysts, and so behave as ' secondary elements '. As a consequence the cycle can start almost from any reaction if the involved isotope is present, and during a complete loop around the cycle the isotope is consumed and then produced again. However, this does not mean that the concentrations of the different isotopes will be unchanged as the final abundances depend strongly on the relative rates of the nuclear reaction in the cycle. Only at a high enough temperature ($T \gtrsim 15 \times 10^6$ K) will the isotopes achieve their equilibrium abundance, i.e. the rate of production is exactly equal to the rate of destruction. At this point, the abundance of each isotope is inversely proportional to the nuclear cross section of the reaction by which it is destroyed. Since the slowest reaction of the CNO cycle is ^{14}N (p,γ) ^{15}O, the most abundant element in the CNO cycle processed material is ^{14}N[7].

The temperature of the complete CNO cycle is much larger than that of the p-p chain, where $\varepsilon_{CNO} \alpha T^{18}$ at T ≈ 10 × 10^6 K. This means that the p-p chain dominates at low temperature – T ≤ 15 × 10^6 K – i.e. in stars with lower than ≈ 1.3 $M_\odot$, while the CNO cycle dominates at higher temperatures, thus for larger stellar masses[7].

The quite different temperature sensitivities of the p-p chain and CNO cycle have an important consequence: if the H-burning process is dominated by the CNO cycle, it is confined to the central regions of the core. This results in a large central energy flux, an occurrence which favours the presence of a central convective region[7].

RESULTS AND DISCUSSION

Through the type of interaction of the Rigel and Procyon and it CNO cycle we use the following equations for drawing in all cases.

configurations n1 (p=2,h=1) A=12 Z= 6 E=0:100 MeV and
configurations n2 (p=2,h=1) A= 14 Z= 7 E=0:100 MeV and
configurations n3 (p=2,h=1) A=16 Z= 8 E=0:100 MeV

The figure (1,3and4) explains the relationship between the state density and energy, results for non-ESM and Ericson and Williams formula for Rigel and Procyon for ^{12}C,^{14}N and ^{16}O, and the figure (2) show the non-ESM relative error with the Ericson's and Williams formula for the number of program iterations.The figure (1) highest value for state density when the atomic mass (A) great this is noted that in case (n_3), and noted the configurations (p,h) same values for all cases.

From figure (2) noted that relative error with the Ericson's formula the increase starts at approximately 1MeV and continue to energy approximately 80MeV here begin to decrease. While the relative error with Williams's formula when all cases are approximately equal, where at low energies Rapid decline to energy 20MeV then it starts to increase to the value of energy 80MeV then it starts to decrease again. The error ratio is Ericson's formula better than Williams's formula

The figure (3,4) comparison between non-ESM and Ericson's and Williams's formula for Rigel and Procyon in case (p=2,h=1) and A=12,14 and 16 noted that from figure (3) Ericson's be higher state density from non-ESM in all case with increased energy, while the figure (4) the non-ESM be higher state density from Williams's formula. And noted that the stat density from non-ESM with Ericson the values are between10^4 and 10^6 while the non-ESM with Williams's formula the values are between10^2 and 10^4.

Cross-Section for Different Channels

The figure show the relation between cross section and energy, we note that the total curve have a maximum value at low energy and decrease smoothly. The following figures have been calculated cross-section of elements (C-12,C-13 and N-14). The figures represents the relationship between the differential cross-section (emission spectrum) and the energy (MeV). The form consists of three important parts where it includes the pre-equilibrium which has the Significant effect in the calculation of cross-section, the second part is called evaporating component, the third part is called the nuclear transfer component. The second and third part have less effect than the effect of the pre-equilibrium. The total or sum of these parts gives us a total calculation for cross-section, Through these forms have been taken values of cross-section at different energies, the reaction rate for each figure was calculated using the following law.

Reaction rate = $\sigma * \nu$ ----------(1)

Where : σ= cross-section and ν= speed of interaction

In general there is an increase and decrease in emission spectrum When increasing energy.

Through the figure(5) we extracted the value of cross-section The value was obtained equal ($3*10^{2}$ mb/Mev) when the energy values equal ($1.6*10^{-12}$J) of the equation (1) was calculated the reaction rate was equal ($2.4*10^{-61}$ m.sec/Kg).

In the same way was calculated reaction rate for all other figures the results are in a table (1).
Figure (5) pre-equilibrium increase at an energy equal to ∼2MeV and be the increase gradually until the reactionreaches an energy value of ∼11MeV starts to decrease and be a little. Evaporation Its value is less than PE and increases with increased energy. The cross-section $(d\sigma/d\varepsilon)_{total}$ it is approximately equal with values of pre-equilibrium.

From figure (6) the pre-equilibrium increase at energy equal to ∼20MeV there is a decrease in reaction pre-equilibrium, this decreases to less than 100MeV. The evaporation for C-12 the reaction time is very low, it starts at a little more than zero and ends at energy equal 10MeV. the nuclear transfer for C-12 its value is higher than the value of H-3 and increases with increased energy when the energy value becomes at 50MeV where it decreases then gradually increases the difference between C-12 and C-13 in the nuclear transfer where they are less valuable at C-13. And the difference in the total cross-section $(d\sigma/d\varepsilon)_{total}$ at C-12 there is a sudden rise while C-13 is in a sudden decline.Note that other figures(7,8,9 and 10) have the same behavior but the values are different for each reaction when the energy increases with different elements involved in the interaction.

From figure (11) the values for pre-equilibrium in all cases. Noted that the elements of H-2 and H-3 the approximation same behavior where there is an increase and decrease of energy 5MeV to 30MeV Then there is a gradual decrease with increasing energy. While the elements of He-3 it begins with energy 5MeV then increases and decreases and it ends at energy 20MeV approximation. While the elements of C-12,C-13 and N-14 is the same behavior and starting from the energy 5MeV and obtain an increase and then a decrease to energy 50 MeV then a gradual decrease and ends at energy 90MeV. And noted that the **Reaction rate** $\left(\frac{m.sec}{Kg}\right)$.

CONCLUSIONS

- Through the results obtained the change atomic mass (A) effects on the state density. in Ericson ,Williams and non-ESM methods there is on direct relationship between atomic mass and state density apparent in the figure(1,3 and4).
- The formulas of methods after adding corrections gives a better result and is more accurate.
- The error ratio is variable with all formulas. The best line ratio when the Ericson's formula.
- The values of the state density from non-ESM with Ericson higher than the non-ESM with Williams.
- The pre-equilibrium it has a greater effect in calculating the cross section.
- Best case for calculating the cross-section for the elements H-2 and H-3.
- Element mass's affect the increase and decrease in the reaction rate.

REFERENCES

1. A. Selman, " Neutron Induced Preequilibrium Nuclear Reactions Using the Exciton Model", Ph.D. Thesis, Uni. Of Baghdad, (2009).
2. E. Gadioli, E. Gadioli-Erba, and J.J. Hogan, Phys. Rev. C16(1977)1404.
3. G. M. Braga-Marcazzan, E. Gadioli, E. Gadioli-Erba, L. Milazzo-Coli and P. G. Sona, Phys. Rev. C6(1972)1398.
4. Kalbach, Phys. Rev. C 24(1981)819.
5. Kalbach, Phys. Rev. C32(1985)1157.
6. K. Cline and M. Blann, Nucl. Phys. A172(1971)225.
7. Evolution of Stars and Stellar Populations – copyright 2005 – John Wiley and Sons Ltd, The Atrium, Southern Gate, Chichester,West Sussex PO19 8SQ, England -pag 105-106 pag 118

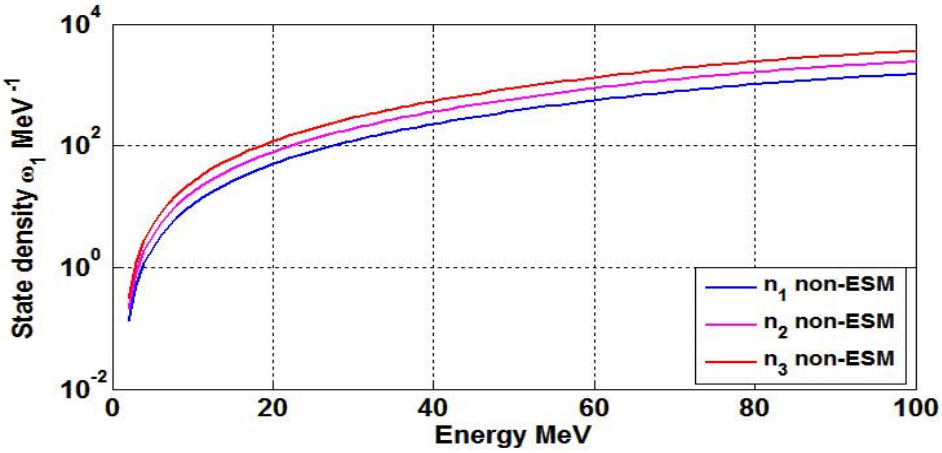


Figure.1: The relation between state density and the excitation energy, E. with non-ESM Ericson and Williams formula where mass number A=(1,2,4) for Rigel and. Procyon.

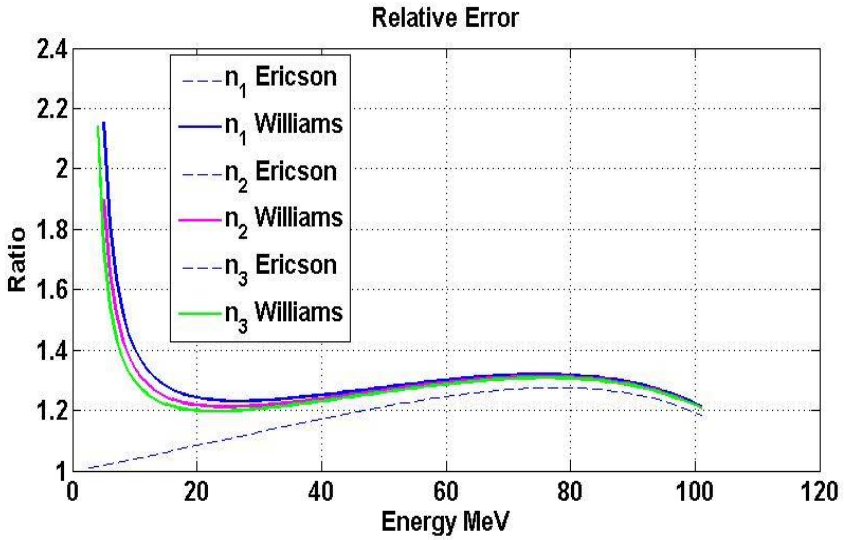


Figure.2: The non-ESM relative error with the Ericson's Williams's formula the number of program iterations are shown for each case, the mass number A=(1,2,4)for Rigel and. Procyon.

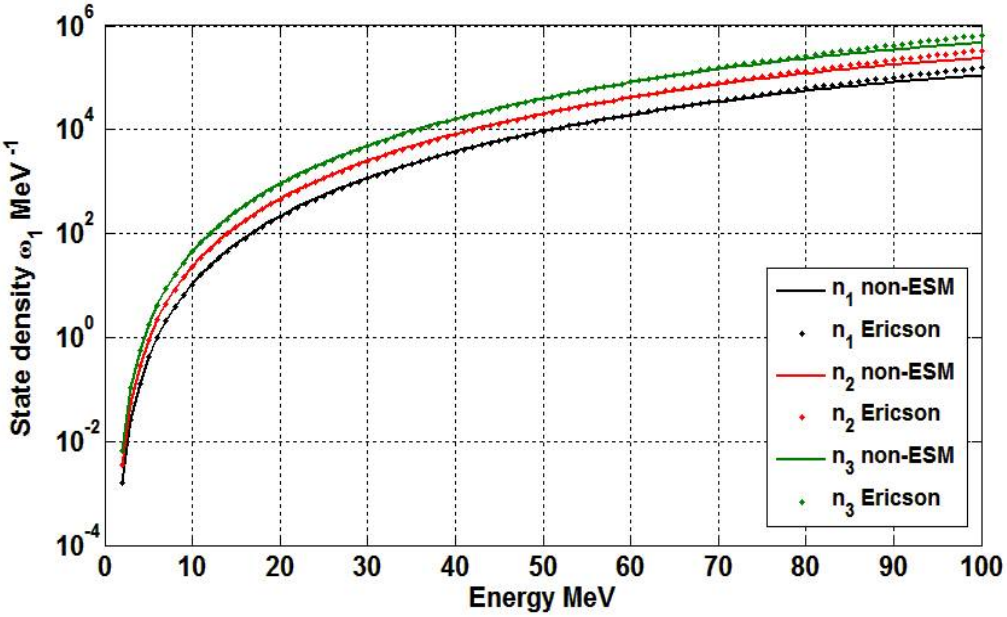


Figure.3: The results of the state density, as a function of excitation energy *E*, for one-component with Rigel and. Procyon for Ericson formula and non-ESM.

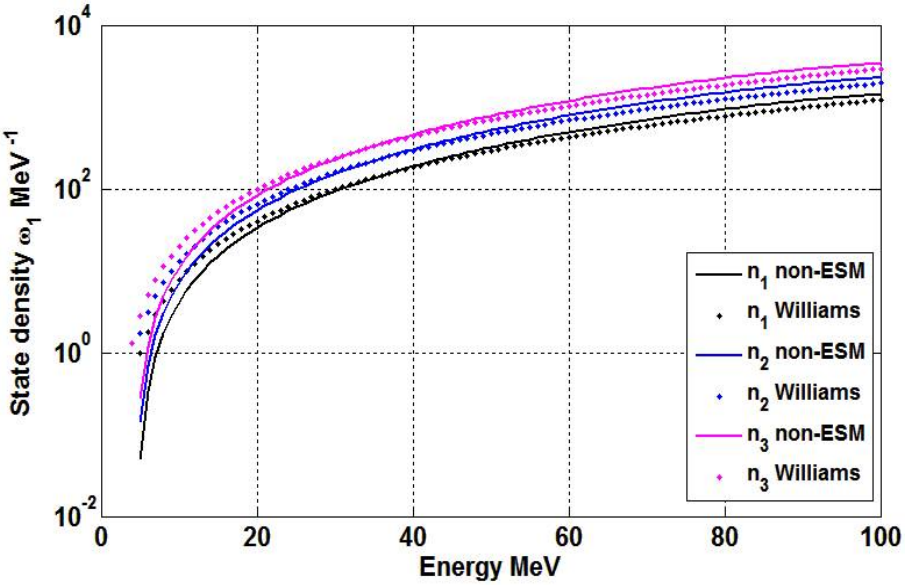


Figure.4: The results of the state density, as a function of excitation energy *E*, for one-component with Rigel and Procyon for Williams formula and non-ESM.

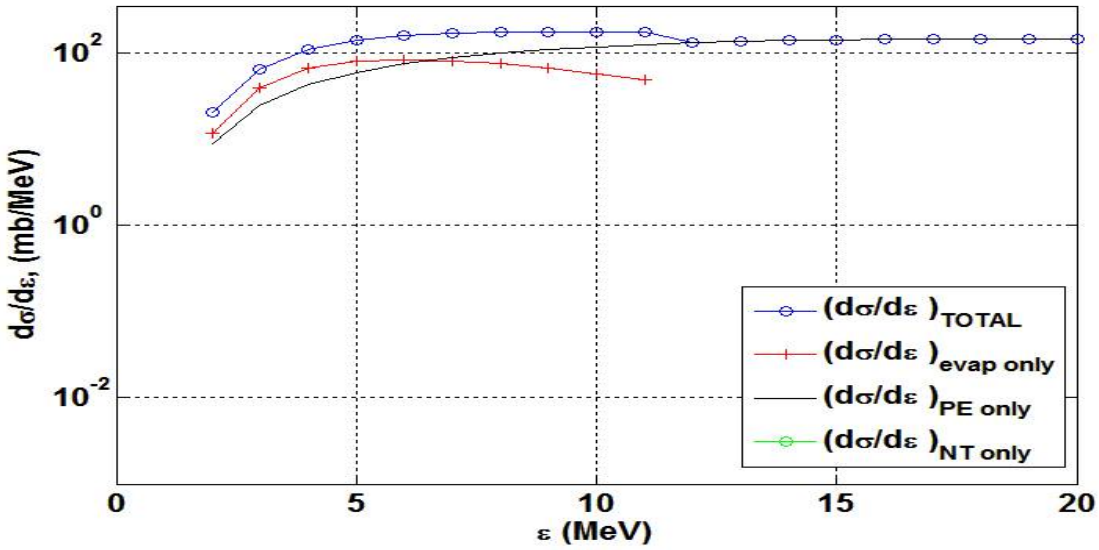


Figure.5: The results of calculated cross-section for C-12 (P,P) Reaction at 100 MeV. and state density is: Pauli two-component state density, ESM

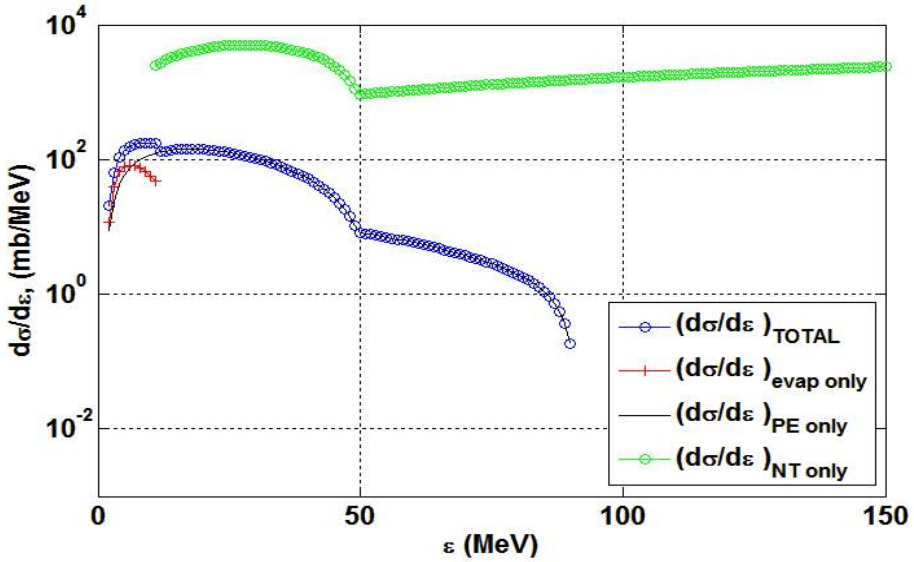


Figure.6: The results of calculated cross-section for C-12 (P,P) Reaction at 100 MeV. and state density is: Pauli two-component state density, ESM

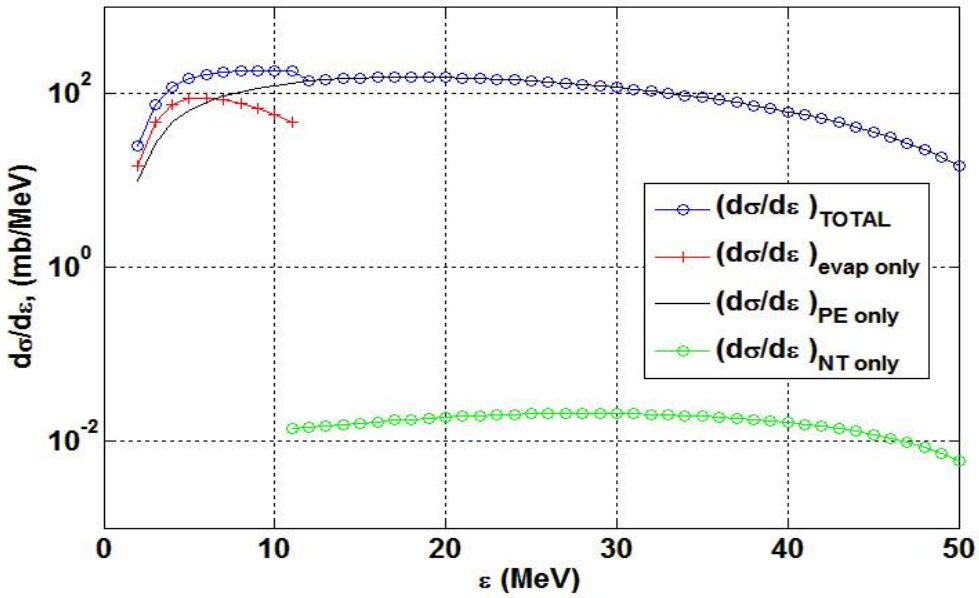


Figure.7:The results of calculated cross-section for C-13 (P,P) Reaction at 100 MeV. and state density is: Pauli two-component state density, ESM

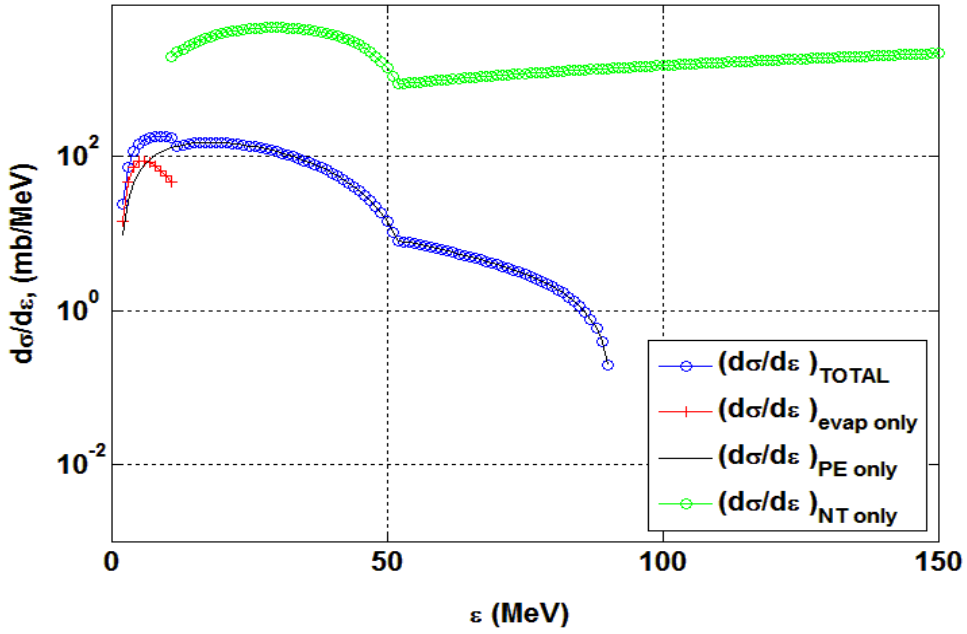


Figure.8: The results of calculated cross-section for , C-13 (P,P) Reaction at 100 MeV. and state density is: Pauli two-component state density, ESM

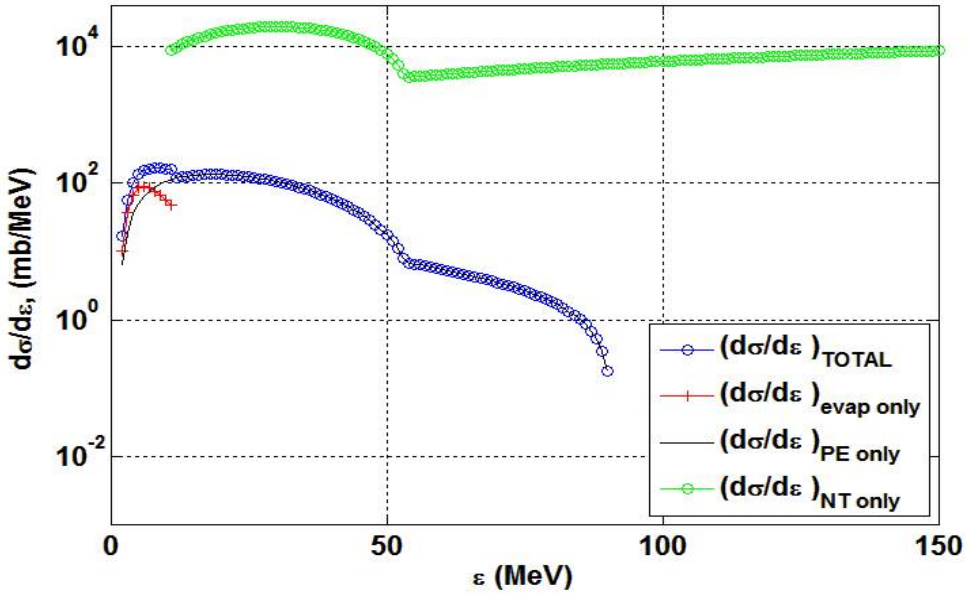


Figure.9: The results of calculated cross-section for , N-14 (P,P) Reaction at 100 MeV. and state density is: Pauli two-component state density, ESM

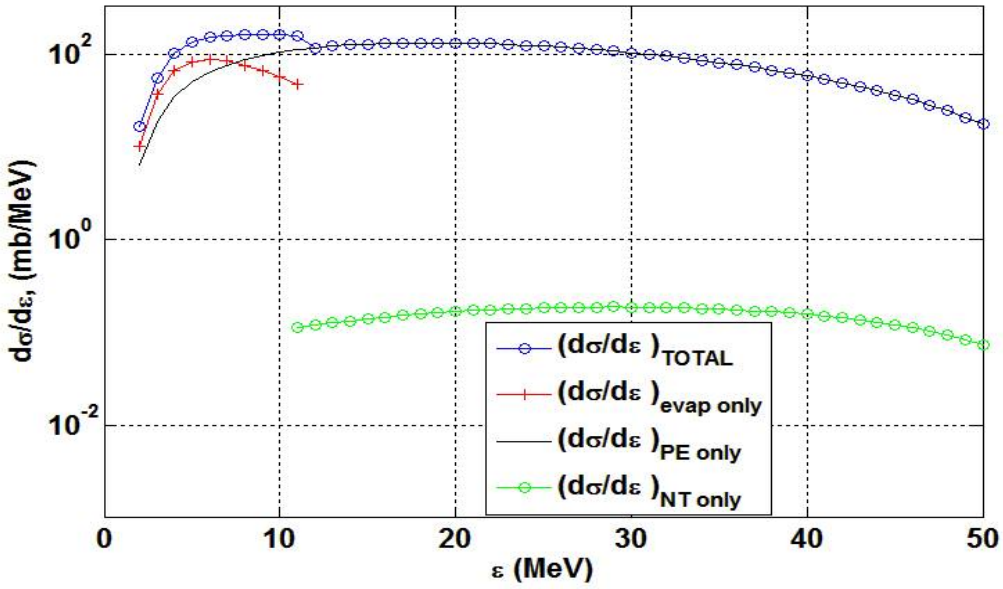


Figure.10: The results of calculated cross-section for , N-14 (P,P) Reaction at 100 MeV. and state density is: Pauli two-component state density, ESM

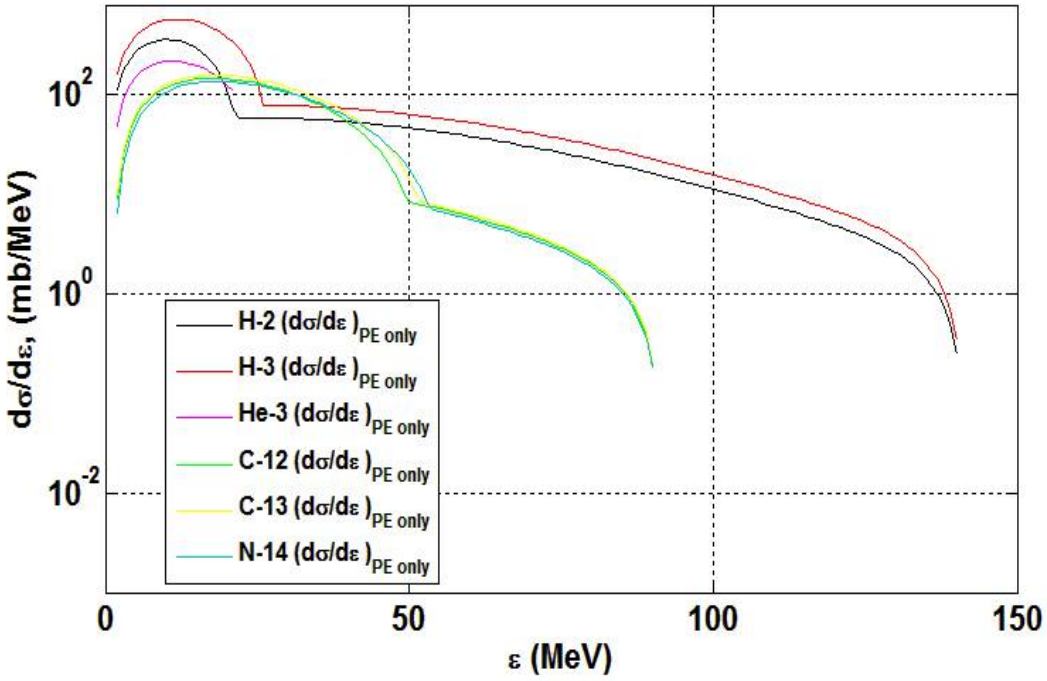


Figure.11: The results of calculated pre-equilibrium for all cases (P,P) Reaction at 100 MeV and state density is: Pauli two-component state density, ESM

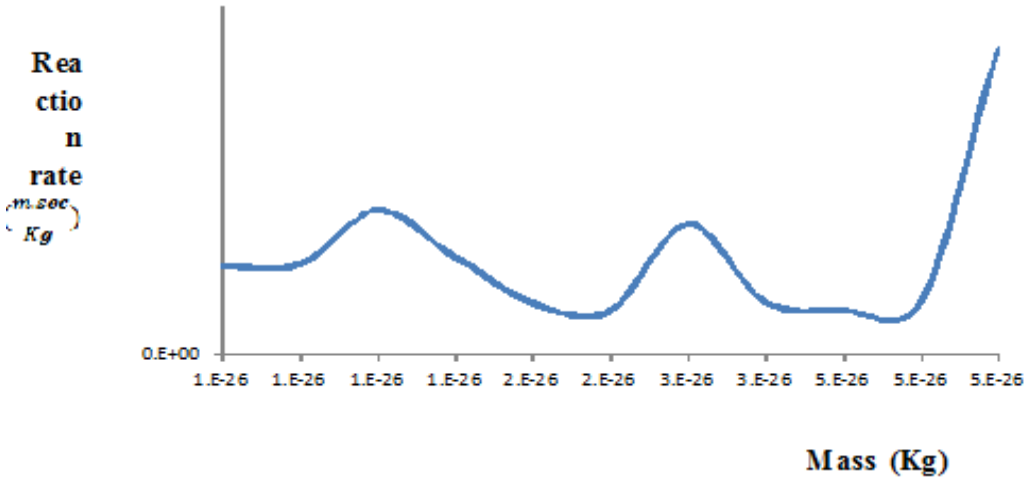


Figure 12. The relationship between reaction rate and mass.

Table 1: calculated reaction rate

Element for (p,p) reaction	**Mass M (kg)**	**Cross Section σ ($\frac{mb}{MeV}$)**	**Energy E(J)**	**Speed= $\sqrt{\frac{2E}{M}}$ V ($\frac{m}{sec}$)**	**Reaction rate = σ * V ($\frac{mb.sec}{Kg}$)**
C-12 Pauli	$1.92*10^{-26}$	$3*10^{2}$	$1.6*10^{-12}$	$1.28*10^{-19}$	$2.4*10^{-61}$
C-12 Pauli	$1.92*10^{-26}$	10^{3}	$1.2*10^{-11}$	$3.201*10^{-18}$	$2.0006*10^{-59}$
C-13 Pauli	$1.12*10^{-26}$	$2*10^{-2}$	$4*10^{-12}$	$2.672*10^{-19}$	$3.313*10^{-65}$
C-13 Pauli	$1.12*10^{-26}$	$9*10^{2}$	$1.2*10^{-11}$	$4.487*10^{-18}$	$2.523*10^{-59}$
N-14 Pauli	$2.338*10^{-26}$	$6*10^{3}$	$1.2*10^{-11}$	$3.16*10^{-18}$	$1.185*10^{-58}$
N-14 Pauli	$2.338*10^{-26}$	10^{-1}	$4*10^{-12}$	$1.849*10^{-19}$	$1.155*10^{-64}$

Indian Journal of Natural Sciences www.tnsroindia.org.in ©IJONS

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Survey on Recognize and Analyze Forgery Documents Based on Feature Detection Method

D.Maalini[1]* and E.Balraj[2]

[1,2]Assistant Professor, Department of Information Technology, M.Kumarasamy College of Engineering, Karur, TamilNadu, India

Received: 07 Feb 2018 Revised: 17 Feb 2018 Accepted: 23 Mar 2018

***Address for correspondence**
D.Maalini
AssistantProfessor,
Department of Information Technology,
M.Kumarasamy College of Engineering, Karur, TamilNadu, India

ABSTRACT

Counterfeited documents appraised to be an exigent work for forensic experts. Counterfeited documents are defined to be replicas of the originals. In general, when counterfeing documents were taken into account, it was observed that many fraudulent have occurred among pan cards, voter id, mark sheets, degree certificates, driving license, etc. The forgers were progressively adopting own masteries to produce fraudulent documents of an individual for criminal offense. Further the forgers were furnishing the fraudulent of an individual without one's own intelligence. Besides forger have also tangled in fabricating illegal replicas of security documents which affects business transactions and leads to smuggling and drug trafficking. Earlier recognition of counterfeited documents was done manually and an information system is developed to discover the fake documents. The pitfall of the prevailing system consumes manual work and it does not provide the better results. It is not feasible to generate authentic results only by adopting edge and corner feature detection. The proposed system incorporates edge and corner features of digital image processing to detect and match counterfeits.

Keywords: Counterfeited, forgers, criminal offense and illegal replicas.

INTRODUCTION

Digital image processing is the technology which could find its own application in various sectors popularly in the fields of medical imaging, satellite imaging and image mining etc. Though it finds applicable in many sectors the image processing technology were used by the individuals or group of criminals to produce counterfeited documents. The difficulties were culminated in various domain knowledgeable areas of social collision, terrorism, and security related procedures. The digital forgeries leads to cultivation of fraudulent in diverse fields such as in students mark sheet, passport, visa, voter id and driving license. Due to the advanced and proceeding digital imaging techniques and with the availability of digital software's, forgers were incorporating their knowledge to

generate digital forgeries. The fraudulent documents were generated through various digitalized techniques such as splicing, copy and paste methodology and resampling. Hence it is difficult for a human being to discriminate between native and the copies.

In the past few years the field of forensics has been involved to work and solve the problems in restoring the digital images. In order to detect the copies of the original they were adopted various algorithms to detect such counterfeited documents. These algorithms trace out the digital image tampering [23]. The trace which has been analyzed is the proofs were the document gets counterfeited. On the other hand if the traces were not found after analysis means it is the result that document remains same and it does not gets counterfeited.There are three ways to generate counterfeited documents which includes,

1. Producing entire copy to the original
2. Altering only the content of the image
3. Altering only the context of the image

Of these all were considered to be counterfeited. These counterfeited documents were identified through pattern recognition and edge and corner detection techniques. The objective of this project is to vanquish the fabrication of fake documents.

Organization

The rest of this paper is organized as follows: Section 2 deals with approaches for classifying the counterfeited documents, Section 3 deals with algorithms available for detecting counterfeited documents. In section 4 we reviewed the related works on counterfeited document detection. The proposed system and workflow of algorithm has been discussed in section 5.

Counterfeited documents

Image tampering is the technique which is mainly involved in generating the counterfeited documents. The image tampering is done by either changing the entire content of the image, or moving the part of the document or replacing or copying and moving the entire image [18]. These counterfeited documents have to be analyzed in such a way that to find discrimination between original and copies. The replicas were commonly found in driving licenses, passports, voter ID, student mark sheets etc. The counterfeited documents have to be analyzed through physical and chemical examinations. The chemical examination which includes some of them are analyzing printing ink, patterns, fluorescent radiations, image locations in pixels, embossing, signatures etc. There were two techniques adopted to identify the replicas.

Active approach

The active approach is mainly concerned with changing the structure of the image and also it invokes altering the security concerned with the image. There are two main active approaches were adopted by the forgers namely Digital watermarking and digital signature. In digital watermarking an element is added to the original image to change the structure of the originality. If the special information is cannot able to extract from the image means then it is true that the original gets tampered. The Security related elements in the image is altered by enabling the digital signature methodology [17]. The structure or result which is obtained has been used for analyzing the tampered of the image. For image tampering analysis whereas algorithms and methodologies were available to detect the tampered or unaltered area in the image. Some of the methodologies were adopted to analyze the tampered images include checksum value and hashing values and analyzing the authentication codes.

Passive approach

The passive approach of forgery detection involves in analyzing semantics and statistics of the image in order to localize the content of the image[15]. This methodology involves in analyzing the security related elements concerned with the images, hence this approach is also coined as the raw image analysis. The core theme of passive approach technique is to locate the tampered areas on the given image. In section 3 automatic detection of falsified document has been diagnosed using digital image processing algorithms were discussed.

Image processing algorithms

The increasing accessibility of elegant editing software's and the increased usage of internet, digital image forgery and document forgery becomes a universal problem. To deal with this problem the following edge, corner detection algorithms has been used.

- Harris corner
- Scale Invariant Feature Transform (SIFT)
- Speeded-Up Robust Feature (SURF)
- Oriented FAST and Roatated BRIEF (ORB)

Harris corner detection algorithm

The Harris corner algorithm was developed by Harris and Stephens [9]. It uses a mathematical representation to detect corners and edges. It uses a function to calculate correlation. The Correlation function considers a window in given image and calculated the change of intensity of image by shifting the window in various directions. The Harris corner window function is given in Eq [1]. $(,) = \sum , (,)[(+ , +) - (,)]$ [1]where, x and y are gray value of pixels and U, V are shifting intensity values

$\Sigma x.yW(x, y)$ is Window function

I[x+U, y+V] is shifted intensity I[x, y] is intensity.

SIFT

The concern related in Harris corner algorithm (zoom in & zoom out) has been resolved by adopting Scale Invariant Feature Transform (SIFT). SIFT is a key point detector and descriptor. This algorithm computes 128 dimensional feature descriptor for every key point. SIFT features are powerful aspirants for image forgery detection because it features are invariant to rotation, scaling and illumination variations. Huang et al. [10] used SIFT key point local features for finding matched regions within the same image.

SURF

Bay et al proposed [22] a new variation of SIFT called SURF, i.e. Speeded-Up Robust Feature. It is also a point detector and descriptor. The standard version of SURF is several times faster than SIFT because of integral image processing. To identify interest points SURF uses an integer approximation of the determinant hessian blob detector, which can be computed with pre computed integral images. Hessian matrix is used to compute local change around the points. SURF identifies the blob like structure at location where the determinant is maximum. SURF summarizes the pixel information within a local neighborhood by convolving pixels with horizontal and vertical wavelet filter.

ORB

ORB is a new version of feature descriptor. ORB means Oriented FAST and Rotated BRIEF. ORB is a corner detector and BRIEF descriptor. ORB comprises of FAST and BRIEF algorithms.

FAST

Features from Accelerated Segment Test was proposed by Rosten et al [20] for identifying interest points in the image. Interest point is a pixel which is well defined in its position and can be robustly detected. High local information content interest points should repeatable between images. FAST algorithm selects appropriate threshold value for selected pixel and process the pixel based on selected value.

BRIEF

Binary Robust Independent Elementary Feature is general feature point descriptor and it can be combined with FAST feature point detector. It loads detected feature point position in pixel via FAST. It computes feature descriptor by having a smoothed image path using Gaussian kernel.

Related Works

The first step in document image forgery detection is identifying corner features in the image.In region duplication detection based on Harris corner points Likai Chen et al [14] used Harris corner points to detect the duplicated regions. Firstly, Harris corner points are detected and image region descriptor method is involved to denote the image area about each Harris corner point with a feature vector. Then the small circle regions are compared using the best bin algorithm to find out replica regions in the image. In Passive copy move forgery detection using SURF and SIFT Features, Ramesh Chand et al [19] proposed a passive copy and move forgery detection using Scale Invariant Feature Transform (SIFT) and Speeded-Up Robust Feature (SURF) features. SIFT and SURF features are resistant totranslation, rotation,scaling transformations. Key points of SIFT and SURF are determined and Euclidean distance is calculated between all key points. Threshold is applied to the least Euclidean distance and the cluster is formed to find the best match. Copy-move forgery detection by robust clustering with J-Linkage byIrene Amerini et al [11] used robust clustering algorithm which focus only on copy move forgery detection. SIFT algorithm is used to group key points that are spatially close.SIFT matching follows a clustering process that groups key points which are spatially close. In some cases this would be a difficult process and does not yields effective results. So J-Linkage clustering is applied which performs effective clustering in the space of geometric transformations. Copy-move forgery detection using SURF feature point detection, Nishadha et al [16] brought up copy and forgery detection based on SURF feature points, in which the analyzed image contents were provided in terms of feature coordinates. These key points or coordinates are found to be invariant to translation and rotation. Here a matrix called Hessian matrix is used in which it stores the key points and these key points were clustered in order to obtain the counterfeited area or region in the input image or document.

Copy-move forgery detection using multi resolution local binary pattern method is robust to illumination variations and geometric distortions of duplicated regions. In this approach image is divided into blocks and each blocks are extracted using local binary patterns i.e. LBP. Duplicated image blocks are determined in the block matching phase using K-d trees for more time reduction. To detect false matching RANSAC algorithm is used. The prevailing methodologies focuses only on passive approaches and in the proposed work is concentrated mainly on automatic analysis and detection of counterfeited documents using digital image processing algorithms. The proposed system incorporates edge and corner features of digital image processing to detect and match counterfeits.

PROPOSED METHODS

Architectur

The proposed work uses the desktop application for uploading document images and verifying the document. Documents are scanned and uploaded to a directory using local server. The proposed workprocedure is to analyze the document images and finding the counterfeited document images automatically. The Figure 3.1 shows the proposed architecture of automated document forgery detection.The architecture is wed based and it is composed of web client and web server. User scans and uploads the document using HTTP request. Uploaded documents will be sent to the server and a new database is created for storing the document images. Then the web server queries the database to find the image associated with the category of the document and country. The web server process the script which establishes the image processing module of corner detection and key point matching. After finishing this web server returns the result to the client of the document being examined. The database is able to store the images and their information for providing better results with corner detection and matching algorithms.

CONCLUSION

In this paper, we have proposed an automatic web client method for analyzing and detecting counterfeited documents using digital image processing algorithms. Speeded-Up Robust Feature (SURF) has been employed for extracting corner and edge features of counterfeited documents. Despite of other digital image processing algorithm SURF algorithm will deals with large set of data set and hence it provides better results.

REFERENCES

1. V. Anand, M. F. Hashmi, and A. G. Keskar, “A Copy Move Forgery Detection to Overcome Sustained Attacks Using Dyadic Wavelet Transform and SIFT Methods”, Springer International Publishing Switzerland, (2014), pp.530–542.
2. Y. Cao, T. Gao, L. Fan and Q. Yang, “A robust detection algorithm for copy-move forgery in digital images”,Forensic Sci Int., January (2012).
3. V. Christlein, C. Riess and E. Angelopoulou, "On rotation invariance in copy-move forgery detection", IEEE International Workshop on Information Forensics and Security (WIFS), (2010).
4. P. Deshpande and P. Kanikar, “Pixel Based Digital Image Forgery Detection Techniques” International Journal of Engineering Research and Applications (IJERA), vol. 2, no. 3, (2012) May-June, pp. 539-543.
5. H. Farid and A. Popescu, “Exposing Digital Forgeries by Detecting Traces of Resampling, “Proceedings of the IEEE Transactions on Signal Processing. (In Press). (2004).
6. J. Fridrich, J. Lukas, and D. Soukal, “Detection of Copy-Move Forgery in DigitalImages.” Proceedings of DFRWS 2003. Cleveland, OH, (2003) August.
7. P. G. Gomase and N. R. Wankhade, “Advanced Digital Image Forgery Detection: A Review”, International Conference on Advances in Engineering&Technology-2014, e-ISSN: 2278-0661, p-ISSN: 2278-8727, pp. 80-83. Hanny Farid “Image forgery detection”, IEEE signal processing magazine, march 2009,pp 16-25.
8. C. Harris and M. Stephens, “A Combined Corner and Edge Detector”, in 4th Alvey Vision Conference, pp 147-151, 1998.
9. H. Huang, W. Guo, Y. Zhang, Detection of copy-move forgery in digital images using SIFT algorithm, in: Pacific-Asia Workshop on Computational Intelligence and Industrial Application (PACIIA), vol. 2, IEEE, Wuhan, China, 2008, pp. 272–276.
10. Irene Amerini , Lamberto Ballan , Roberto Caldelli , Alberto Del Bimbo ,Luca Del Tongo , “Copy-move forgery detection and localization by means ofrobust clustering with J-Linkage, I. Amerini et al. / Signal Processing: Image Communication 28 (2013) 659–669.

11. Y. Ke, Q. Zhang, W. Min and S. Zhang, " Detecting Image Forgery Based on NoiseEstimation" International Journal of Multimedia and Ubiquitous Engineering, vol. 9, no. 1, (2014), pp.325-336.
12. W.N.Lie, G.Shin and S.L.cheng," Dual protection of JPEG Images based on informed embedding and two stage watermark extraction techniques," IEEETrans Information forensics and security,vol.1,no.3,pp.330-341,sep 2006.
13. ikai Chen a, Wei Lu a,⇑, Jiangqun Ni a, Wei Sun b, Jiwu Huang,"Region duplication detection based on Harris corner points and step sector statistics", Vis. Commun. Image R. 24 (2013) 244–25
14. Muhammad, M. Hussain and G. Bebis, "Passive copy move image forgery detection using undecimated dyadic wavelet transform", Elsevier doi:10.1016/j.diin. 2012.04.004.
15. Nishadha S. G., Jaseela S,"Copy Move ImageForgeryDetection Using SURF Feature PointExtraction ",International Journal Of Scientific Engineering Research, Volume 7, Issue 7, July-2016 653 ISSN 2229-5518
16. A.c.popesu and H.Farid, "Statistical tools fordigital forensic", in proc, the 6th international workshop on information hiding, Toronto, Can13 V.P. Raval, "Analysis and Detection of Image Forgery Methodologies", International Journal of Scientific Research and Development 2013, vol. 1, no. 9, (2013).
17. S. Ryu, M. Lee and H. Lee, "Detection of copy-rotate-move forgery using zernike moments" in Proc.Int.Workshop Information Hiding, Springer , (2010), pp. 51–65.ada 2004.
18. Ramesh Chand Pandey, Rishabh Agrawal, Sanjay Kumar Singh, and K.K. Shukla."Passive Copy Move Forgery DetectionUsingSURF,HOG and SIFTFeatures".
19. E. Rosten and T. Drummond, "Machine learning for high speed corner detection", in 9th Euproean
20. Conference on Computer Vision, vol. 1, 2006, pp. 430–443.
21. R. Sekhar and A S. Chithra, "Recent Block-based Methods of Copy-Move Forgery Detection in DigitalImages",InternationalJournalofComputer Applications March 2014, vol. 89, no. 8, (2014) March.
22. Tinne Tuytelaars2, and Luc Van Gool12, "SURF: Speeded Up Robust Features Herbert Bay1".
23. N. Tiwari, N. Hemrajani and M. K. Ramaiya, "A Novel Technique For Tampering Detection" International Journal of Computer Science Engineering and Information Technology Research (IJCSEITR), ISSN 2249-6831, vol. 3, no. 3, (2013) August, pp. 341-346.
24. A.K. Yadav, D. Singha nd V. Kumar, "Forgery (Copy-Move) Detection In Digital Images using Block Method", International Journal of Collaborative Research in Engineering Sciences, (2014), April.
25. J. Wu, Z. Cui, V. S. Sheng, P. Zhao, D. Su and S. Gong, "A Comparative Study of SIFT and its Variants",Measurement Science Review, vol. 13, no. 3, (2013).

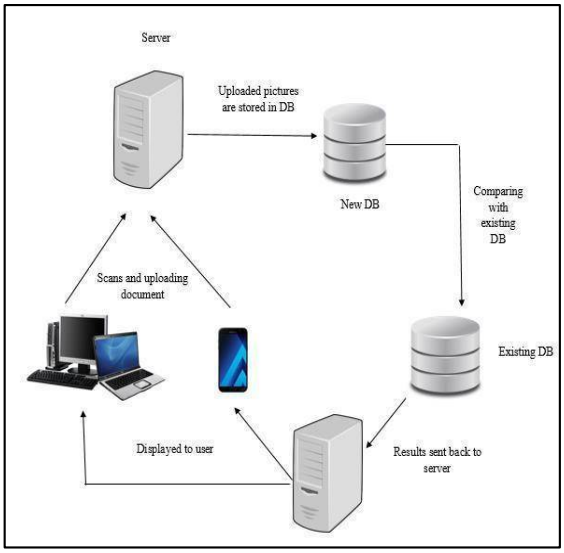


Figure:1.Web – client Architecture for automatic detection of counterfeited documents

Indian Journal of Natural Sciences *www.tnsroindia.org.in ©IJONS*

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Public Auditability and Data Dynamics for Security in Cloud Storage

N.Valarmathi[1]* and A.Shanmugavelaytham[2]

[1,2]Assistant Professor,M.Kumarasamy College of Engineering, Karur,TamilNadu,India

Received: 11 Feb 2018 Revised: 19 Feb 2018 Accepted: 23 Mar 2018

***Address for correspondence**
N.Valarmathi
Assistant Professor,
M.Kumarasamy College of Engineering,Karur,TamilNadu,India
E mail: mathi.nataraja@gmail.com/ shanmugavelaytham.it@mkce.ac.in

ABSTRACT

Cloud computing refers again to the availability of computational belongings on demand by way of a computer community and also it has been anticipated as the following new release structure of IT corporation. The principle in the back of the cloud is that any laptop linked to the web is connected to the equal pool of computing energy, features, and records. The appliance program and expertise's are resides within the centralized giant knowledge core's but the core or server will not be thoroughly dependable. To make certain the reliability of competencies cache in cloud gauge 1/3 party auditing (TPA) is used on behalf of the consumer to avert the usual checking by way of purchaser to verify the information's are saved in proper area, and it's surely intact. In addition, to reap economies of scale, which is one of the advantages of cloud computing. The aid for knowledge dynamics by way of almost probably the most general varieties of understanding operation corresponding to dam amendment, renew, remove and add , as well be a giant stage toward reasonableness, considering the fact that choices in cloud gauge will not be restricted to collection or support know-how. This paper objectives at modifying the current resistant of cache items near way of utilising mold the Logarithmic merle mess hierarchy progress for hunk label verification, to acquire effective in sequence and BLS algorithm for error correction, bilinear blend signature to help multiuser auditing. This paper gather enormously amazing and provably copy. Cloud atmosphere is created making use of eucalyptus open supply.

Keywords: information storage, public audit ability, knowledge dynamics, cloud computing.

INTRODUCTION

Cloud gauge is a original time period on behalf of whatever so as to entails offering hosted forces over the web. They offerings be commonly separated into three classes: Infrastructure-as-a-supplier (IaaS), Platform-as-a-supplier (PaaS) and utility-as-a-carrier (SaaS). Cloud packages have three specific features that distinguish it beginning typical internet site web hosting. It's provided on particular, most of the time via the minute or the hour; it's bendy - a man

or woman be able to have as so a lot or as tiny of a facility as they crave at some known point and the check is completely driven with the aid of the cloud supplier. The competency of blur is expense fiscal funds. The high challenge is safety.

It is utilized through a lot of Software Company at the present a days. From time to time the cloud service provider could bury the info corruption to maintain the name. To prevent this trouble, we introduce an strong 1/three celebration auditor to audit the man or woman's outsourced knowledge as soon as required. The defense is completed near using utilising hint the information section. Hint is carried out with BLS algorithm. We increase unrestricted key established homomorphic security by means of unintentional blind receive privations preserve guiding practice.TPA make the guiding enterprise for both and every person i.E. sole auditing.

Cloud storage is a model of community pc expertise cupboard space the place knowledge is saved on many digital servers, mostly hosted by way of zero.33 parties, rather than being hosted on committed servers. Website hosting company operate massive expertise facilities; and individuals who need their information to be hosted buy or hire storage space ability from them and use it for his or her storage wishes. The info center operators, within the historic past, virtualize the property in keeping with the standards of the patron and expose them as digital servers, which the consumers can themselves manipulate. Bodily, the resource may just span throughout more than one servers. Cloud storage offerings can be accessed by way of an internet provider application programming interface (API), or by the use of an internet-headquartered individual interface.

Related Works

C.Wang et al., [2009] This paper don't forget active records garage in a allotted state of affairs, and the future project-response procedure can each find out the facts accuracy and discover likely errors. They just remember unfair carry for energetic statistics operation. It use visibly supportable homomorphic authenticators built beginning BLS on which the clue can be accurate into small authenticator price, and open retrievabiltiy is finished. This scheme offer honesty certification for distinctive information catche systems, the problem of supporting each public guiding and information dynamics has no longer been absolutely addressed. To gain a secure and proficient plan to faultlessly integrate statistics garage services.

C.Erway et al., [2009] they prolonged the PDP form to aid verifiable renew to save records files using rank- primarily based authenticated pass lists. This scheme is basically for block adds, they remove the index statistics within the label calculation and make use of leave out listing records shape to authenticate the tag in sequence of challenged or efficient blocks first earlier than the verification process. The current schemes goal at imparting integrity verification for different records garage structures, the trouble of supporting each public audit ability and statistics dynamics has now not been completely addressed. How to reap a comfortable and green layout to effortlessly combine those crucial apparatus for information catche provider stays an open challenging challenge in blur.

System Design

Blur guage mechanism are classify as 1) Cloud User (CU), 2) Cloud Service provider (CSP) and 3) Cloud server,(CS) Client: an thing, which has big data files to be saved in the cloud and is based at the cloud for records protection and computation, can be either character clients or businesses.

Cloud Storage Server (CSS): an entity, that's controlled with the aid of Cloud Service Provider (CSP), has sizeable storage area and computation aid to preserve the customers' data.

Third Party Auditor (TPA): an entity, which has expertise and skills that clients do no longer have, is relied on to assess and divulge threat of cloud garage Services on behalf of the clients upon request.

In the blur pattern, through placing the big records documents on the far flung servers, the customers may be thankful of the trouble of garage and calculation. while customers now not own their records regionally, it's miles of important substance meant for the customers to make sure that their figures are human being efficaciously stored and followed. With the intention of is, customers ought to be organized through positive safety way so to facilitate it will at times prove the accuracy of the remote records even without the existence of neighborhood copies. In case those customers do no longer always have the time, feasibility or resources to display their facts, they are able to delegate the tracking mission to a trusted TPA.

The 1/3 birthday party auditor (TPA), who has records and skills that cloud clients do not have and is depended on to assess the cloud garage service safety on behalf of the patron upon request. Users rely upon the CS for cloud records garage and maintenance. They may additionally dynamically interact with the CS to get right of entry to and replace their saved facts for severa software capabilities. The clients may also inn to TPA for making sure the garage protection of their outsourced information, at the equal time as hoping to hold their records non-public from TPA. However, in a few unspecified time within the future of imparting the cloud information garage based totally services, for their personal advantages the CS may probable forget about approximately to preserve or deliberately delete now not regularly accessed information files which belong to normal cloud clients. additionally, the CS might also choose to cowl the records corruptions due to server hacks or Byzantine disasters to maintain popularity. We wait for the TPA, who's within the company of auditing, is reliable and independent, and for that reason have refusal enticement to plan with both the CS or the customers ultimately of the guiding way. TPA ought to be able to efficiently audit the cloud records storage without close by duplicate of facts and devoid of coming in extra on line up weight to blur customer.

Design Goals

TPA has to be designed for Integrity checking with out neighborhood reproduction preserving in purchaser. It need to paintings in allotted surroundings. It must achieve all types of safety requirements in cloud garage. Support information dynamic operation. Highly green and resilient to attackers.

To create an green BLS set of regulations for signing the information. To reap the statistics dynamics for block records logarithmic Merkle hash tree is used. To perform a couple of auditing responsibilities Bilinear combination signature is used.

Proposed System

To efficaciously help communal review skill while not have to recover the information section themselves, Homomorphism authenticators are unforgivable more than data appeared from man or woman statistics blocks, which can be securely aggregated in such a manner to assure a verifier that a linear mixture of facts blocks is correctly computed by verifying simplest the aggregated authenticator. In this scheme PKC based homomorphic authenticator to equip the verification protocol with public auditability. Suppose so as to folder F is separated into n hunk m1,m2, . . . ,mn

Here we are assigning

M=a;

K=b;

1, where mi ∈ Zp and p is a biggest prime no..

Let e: G×G → GT be a bilinear map, with a hash function H : {0, 1}∗ → G, viewed as a arbitrary vision .Let g be the creator of G. h is a cryptographic confusion task.

The BLS name format belongs to three stages,

1. In the key generation phase, a dispatcher select a chance figure x € Z and guage $y = g_1{}^x \in G_1$. The secret key is x and the open key is y.

2. set a communication $m \in \{0,1\}$in the clue stage the dispatcher initial gauge $h = h(m) \in G_1$, everywhere h () is a confusion function, and after that gauge $\sigma = h^x \in G1$. The signature of m is σ.

3. During the verification segment, the receiver first computes $h = h(m) \in G_1$after which test whether or not. If the proof succeed, next the communication m is true.

Logarithmic Merkle Hash Tree

This set of rules is implemented in TPA. Merkle timber have determined many makes use of in theoretical cryptographic structures, having been especially designed sincerely so a leaf fee may be hooked up with recognize to a freely recounted core price and the verification in sequence of the side. This authentication statistics carry one node fee at every top, in which those nodes are the relative of the bump on top of the course between the leaves to the premise. The Merkle tree traversal trouble is the assignment of locating an inexperienced algorithm to output this authentication information for successive leaves. The trivial answer of storing every node charge in memory requires an excessive amount of vicinity. On the other hand, the technique of computing the authentication nodes at the round they're required can be very luxurious for a few nodes. The mission is to keep each vicinity and computation thru the use of amortizing the charge of computing such luxurious nodes.

The upcoming wished authentication nodes are computed as inside the conventional traversal, but the numerous stacks do now not all acquire equal hobby. Each TREEHASH example may be characterized as being both not began out, in component finished, or finished. Our schedule to prefer to finish Stack for the lowest h values first, except a few distinct stacks has a decrease tail node. We unique this choice by using manner of defining l min be the minimum of the h values Stack.Low, then selecting to awareness our interest at the smallest degree h undertaking this minimum. (putting Stack.Low = 1 for finished stacks efficiently skips them over). In extraordinary phrases, all stacks ought to be finished to a point wherein there are not any tail nodes at top h or a lot much less in advance than we start a cutting-edge day Stack TREEHASH computation

Multiple Batch Auditing

TPA also can moreover concurrently deal with multiple auditing delegations upon one of a kind customer' needs. The person guiding of those obligations used for TPA may be boring along with simply ineffective. Agreed K guiding delegations on K notable statistics documents from K splendid clients, it's far extra exquisite for TPA to batch the ones more than one obligation collectively and guiding at one time. Maintenance that herbal call for in intelligence, the proposed method of blinder combination autograph, whichever wires the gathering of several autograph by manner of superb signers on superb messages right into a single signature and because of this offers green verification for the authenticity of all messages. Using this signature aggregation approach and bilinear assets, it aggregate K verification equations proper right right into a unmarried one, in order so as to the immediate guidiing of many duties be able to be done.

Diagram for spread records cache protection.

In the direction of in addition decorate the supply of the information garage safety, person character's information may be daily cached in multiple objective places. so as to is, as well creature broken at man or woman users, facts cached and it be hired diagonally a couple of users to accept blunder or user blast as buyer information gain in volume as well as significance so widely known that deleting-perfect rules be able to tolerate many loss in dispensed

catche structures. During cloud statistics catche, we are able to depend upon this method to disperse the records report F redundantly at some point of a hard and fast of n = a+okay distributed servers. A [a+b, b] method is to make adequate daily parity rule from m records rule in this type of technique so as to the actual m facts rules could live restore beginning several m away of the a+adequate information and equality angle. Next to insertion every of the + okay angle on top of a special attendant, the authentic information record preserve live on the letdown of a few okay of the a+ok servers exclusive of any information defeat. Such a allotted cryptographic machine to permit a tough and speedy of servers to show toward a patron so as to a filled record is unbroken and retrievable.

CONCLUSION

The scheme is the primary to maintain flexible and inventive unrestricted audit in the blur storeroom. The method of Bilinear combined autograph is second-hand to get group guiding and additionally multi patron guiding, and where TPA can perform more than one guiding obligations at once. The information within the shade does now not stay put still. In this system make use of Logarithmic merle hash tree algorithm for supporting data dynamic operations. Its performance is high when compared to other auditing services. Log. Merle hash tree algorithm t increase the speed of TPA to provide high performance. It wires secure and capable active process on data block stored in the blur, including: information renew, remove and add.

REFERENCES

1. G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable data possession at untrusted stores," in *Proc. of CCS'07*. New York, NY, USA: ACM, 2007, pp. 598–609.
2. A. Juels and B. S. Kaliski, Jr., "Pors: proofs of retrievability for large files," in *Proc. of CCS'07*. New York, NY, USA: ACM, 2007, pp. 584–597.
3. H. Shacham and B. Waters, "Compact proofs of retrievability," in *Proc. of ASIACRYPT'08*. Melbourne, Australia: Springer-Verlag, 2008, pp. 90–107.
4. K. D. Bowers, A. Juels, and A. Oprea, "Proofs of retrievability: Theory and implementation," Cryptology ePrint Archive, Report 2008/175, 2008.
5. M. Naor and G. N. Rothblum, "The complexity of online memory checking," in *Proc. of FOCS'05*, Pittsburgh, PA, USA, 2005, pp. 573–584.
6. E.-C. Chang and J. Xu, "Remote integrity check with dishonest storage server," in *Proc. of ESORICS'08*. Berlin, Heidelberg: Springer-Verlag, 2008, pp. 223–237.
7. M. A. Shah, R. Swaminathan, and M. Baker, "Privacy-preserving audit and extraction of digital contents," Cryptology ePrint Archive, Report 2008/186, 2008.
8. A. Oprea, M. K. Reiter, and K. Yang, "Space-efficient block storage integrity," in *Proc. of NDSS'05*, San Diego, CA, USA, 2005.
9. T. Schwarz and E. L. Miller, "Store, forget, and check: Using algebraic signatures to check remotely administered storage," in *Proc. of ICDCS'06*, Lisboa, Portugal, 2006, pp. 12–12.
10. Q. Wang, K. Ren, W. Lou, and Y. Zhang, "Dependable and secure sensor data storage with dynamic integrity assurance," in *Proc. Of IEEE INFOCOM'09*, Rio de Janeiro, Brazil, Appril 2009, pp. 954–962.

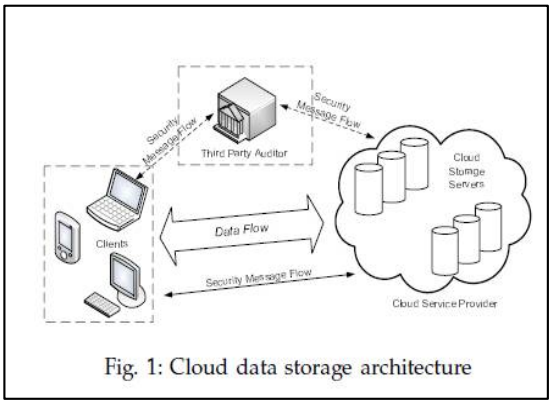
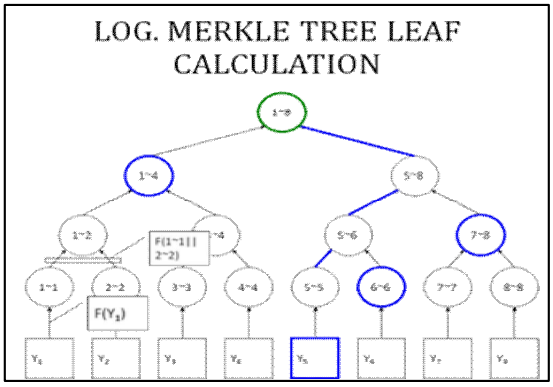


Fig. 1: Cloud data storage architecture



1. Set $leaf = 0$.
2. **Output:**
- Compute and output $\Phi(leaf)$ with $LEAFCALC(leaf)$.
- For each $h \in [0, H-1]$ output $\{Auth_h\}$.
3. **Refresh Auth Nodes:**
For all h such that 2^h divides $leaf + 1$:
- Set $Auth_h$ be the sole node value in $Stack_h$.
- Set startnode $= (leaf + 1 + 2^h) \oplus 2^h$.
- $Stack_k.initialize(startnode, h)$.
4. **Build Stacks:**
Repeat the following $2H - 1$ times:
- Let l_{min} be the minimum of $\{Stack_h.low\}$.
- Let $focus$ be the least h so $Stack_h.low = l_{min}$.
- $Stack_{focus}.update(1)$.
5. **Loop:**
- Set $leaf = leaf + 1$.
- If $leaf < 2^H$ go to Step 2

Fig.2.Algorithm for Log. Merkle hash tree

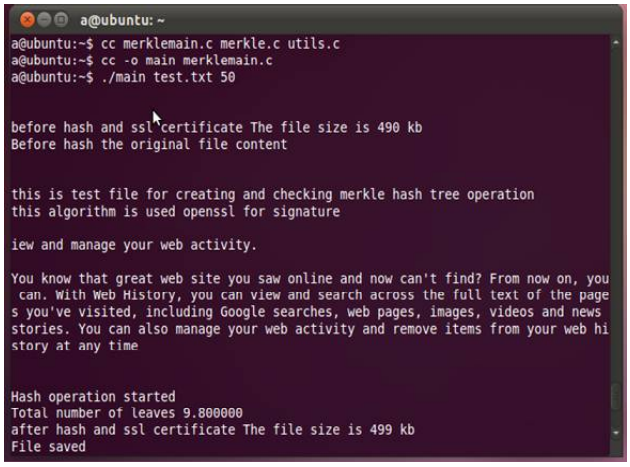


Fig.3.Integrity checking of log merkle algorithm at sending.

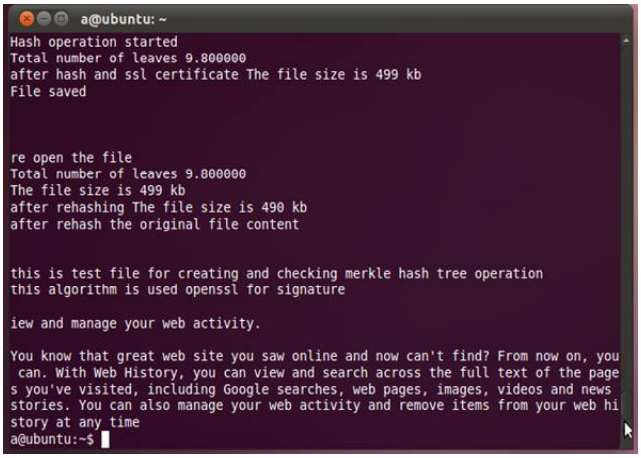


Fig.4. Integrity check at receiving

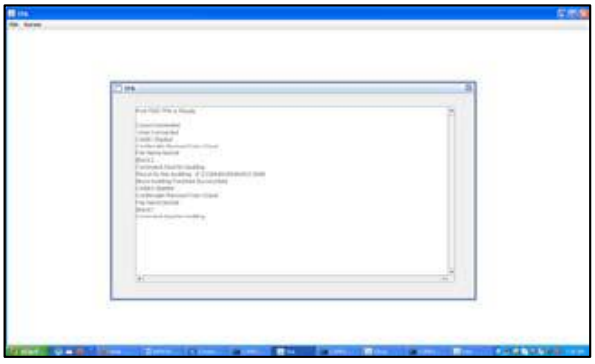


Fig.5. Multiple auditing

Indian Journal of Natural Sciences *www.tnsroindia.org.in ©IJONS*

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Improved Object Selection and Object Retrieval Using Prediction Errors

R.Sujatha[1]* and C.Selvarathi[2]

[1,2]Assistant Professor, M.Kumarasamy College of Engineering,Karur,TamilNadu,India.

Received: 18 Feb 2018 Revised: 21 Feb 2018 Accepted: 23 Mar 2018

***Address for correspondence**
R.Sujatha
Assistant Professor, M.Kumarasamy College of Engineering,Karur
TamilNadu,India.
E.Mail:sujathar.it@mkce.ac.in, selvarathic.it@mkce.ac.in

ABSTRACT

Interactive Object Segmentation plays a major role in Image segmentation. Many algorithms and approaches have been proposed earlier. An attempt as been made in this work to retrieve the desired object focusing on the storage space, data security and bandwidth requirement. The two important factors considered for object localization are object proposals and location prior.The proposed work used prediction errors as important features features for image retrieval. The images are grouped into a set of classes using a clustering system which is implemented in two steps. The first step uses wavelet-based contourlet and Fuzzy C means clustering algorithm. The prediction is done with the help of Feed forward artificial neural network. The pixel difference between the predicted and the original image is mainly used for performing prediction. The similarity matching between the prediction errors of the required image and the image stored in database is performed.

Keywords: Interactive Object Segmentation, algorithm, Image segmentation

INTRODUCTION

Object segmentation is important concept in image processing. Image retrieval plays a major role after the desired object is detected. Many methods such as Magic wand and Graph cut methods have been used in object selection and segmentation. Many factors such as complexity, background color distributions and tuft image boundaries play a vital role in accurately selecting the object with the help of user interaction.

The main purpose of this work is to reduce the user effort spent in selection of an object and to improve the object retrieval. Our approach focuses on selecting the small objects and potentially weak saliency objects which are associated with many other saliency objects. Our approach aims at performing image editing using Natural language input. Our approach uses a very simple object selection where the user is allowed to specify the class of objects and select from a image. The image search is done in two folds:

1.Classification of object proposals in the input to check the presence of selection class.

2.Presenting the Localization information and appearance models to select the object.

Positive and negative images are gathered using a text based image search. Object proposals are made to select the required object finally. The object proposals thus made are used for computing location prior and appearance models. Energy minimization framework is utilized to compute the final object selection.Object segmentation is very vital for object detection from an image. It focuses on converting the image into gray scale images and then subdivisions of the images are done. Pattern matching is done for every part of the image. The pattern matches found are then combined together form a segmented image. FCM method based on feature extraction using Wavelet based contourlet transformation is used for better image retrieval.

Related works

Many state of the art methods were used for object selection failed to select the named objects. The main objective of our approach is to perform improved interactive object selection. The main aim of our work is to select the required object using the salient feature of the image region. Optimization of the energy function determines the salient feature of the required image. Salient pixels are found in highly contrast region. Tag based selection process and Grabcut method are used for selecting an object to reduce the effort of the user in object selection. These works have helped us in better understanding of our approach.

Tag based selection approach was introduced by Ezaj Ahmed, Scott Cohen and Brain Price in their work named Semantic Object segmentation reduces the user effort which is required for object selection. Their work also focused on the aspects of segmentation and midlevel clustering, key factors for deciding the homogeneity of a sub region.The desired object is localized by using a Bounding box, and Grab cut method is initialized during the segmentation process. Co segmentation is performed on multiple images which is provided as input. A common object is selected in every image. These approaches are used for our work to query the internet and to perform the same operation on the results and the query image. Sparseness is computed for the function to perform the co segmentation.

Proposed work

In our approach we use the tag based search to formulate the object proposal. Object proposal is required to calculate the location prior which is indeed the key factor for identifying the location of the object. The object selection uses a voting scheme which is based on image search localizations .Our approach aims at reducing the storage space and to handle the object classes which have larger appearance variations and contain less rigid objects.

OVERVIEW

Our approach uses the name of the object as input as shown in the figure. Input tag is used to perform text query from the image database. Thus the exemplars are gathered related to the object . The positive and negative exemplars are collected and preserved in the image database. The validation of the object is done based on the object proposals and the queries pertaining to the object proposal.

Potentially contained object proposals are considered to estimate the location of the object for its respective exemplars. SIFT Flow based image warping is used to find the location prior. Our approach combines the image specific appearance model and location prior framework. The localization framework is obtained by using the Grab cut methods.

Localization

Location prior is found by using the tag of the object. We collect the exemplars by using the WBCT and FCM to create a image retrieval database. The target image is divided and the validation of the object proposals are

completed. SIFT Flow is used to validate the exemplar data for its corresponding the target images and to transfer the location information.

Exemplar Retrieval Database

Input tag is used to collect both the positive and negative exemplars and to create the positive and negative exemplars. Here we use WBCT and FCM clustering for more texture based image retrieval. Wavelet Based Contourlet Transform which consists of two filter banks stages. The first stage provides the sub band decomposition and the second provides the segmented contours for the images.

Contourlet is formed by the combination of the DFB basis along with the multiscale subspaces of the Laplacian pyramid as given by:

$$\{ g_k^l [n - S_k^l m] \}, 0 \le k \le 2^l \quad m, n \in Z^2$$

Texture is very important properties of the images. It provides a regional geometric descriptor which helps to retrieve the images. The WBCT block diagram is shown below:
Fuzzy C Means Clustering :

Algorithm I:

Step 1: Initialize membership $U = [u_{ij}]$, $U^{(0)}$.
Step 2: Evaluate the fuzzy centroid $C^{(k)} = [cj]$, where $j = 1, \ldots, c$, using Eq. 3
Step 3: Update the fuzzy membership $U^{(k)} = [u_{ij}]$, $U^{(k+1)}$ using Eq. (2).
Step 4: If $||U^{(k)} - U^{(k-1)}|| < \varepsilon$,, then STOP, else again compute centroid using step 2.

After the decomposition phase LH,HL<HH which are the three high pass bands are obtained.These coefficients of the wavelets are used for texture discrimination. Fuzzy C means algorithm is used to assign datapoints, each of which has its own degree of membership for its respective cluster.

This algorithm separates each input image Y = { y1,……yn}into a collection of cluster centers C ={c1,……cc}using the image texture features respective to the objective function Jm. The Degree of membership u_{ij} and cluster center c_j can be computed using the below:

$$u_{ij} = \frac{1}{\sum_{j=1}^{c} \left(\frac{\| y_i - c_j \|}{\| y_i - c_j \|} \right)^{\frac{2}{m-1}}}$$

$$c_j = \frac{\sum_{i=1}^{n} u_{ij}^{m} y_i}{\sum_{k=1}^{n} u_{jk}^{m}}$$

Artificial Neural network comprises of an input, hidden, output layer. It is used to predict the image predict the image pixel values using the neighboring pixels. An optimal ANN predictor can be designed by selecting the appropriate number of input and hidden layer neurons as well the best activation functions for the hidden and output layers. The ANN model with three layers and the activation function required for the ANN predictor is shown in the figure below:

$f(x) =$

x *Linear*

$\frac{1}{1+e^{-x}}$ *Sigmoid* (*Logistic*)

$\frac{1-e^{-2x}}{1+e^{-2x}}$ *Hyperbolic tangent*

Optimal ANN Configuration:

Prediction Error (PE) matrix of an image is determined as the difference between actual and the predicted pixel values.The optimal ANN is selected based on the input,hidden,output layer neurons, activation functions and the training algorithms which gives the minimum prediction error for that required class. The performance of PE matrix is given by

$$ARMSPE = \frac{1}{K}\sum_{k=1}^{K}\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}PE_{ij}^{k}}{N^2}$$

Where K represents the total number of images in the class.

ANN Prediction

For each cluster a separate ANN model which is optimal is obtained by using the procedure given above.The Prediction error matrix (PE) of an image is given by:

$$PE = A - A'$$

where A and A' represent the original and the predicted matrices of the image.

Object retrieval

Object Retrieval is done based on the query image and the Prediction done by using the ANN prediction based on the prediction errors . The bounding boxes are used based on the prediction result.

DETECTION via OBJECT PROPOSAL VALIDATION

Voting map is used for determining the object center locations and the scores. The votes are casted based on the corresponding voting maps. The similarity scores of the exemplars and the potential location of the objects based on the exemplars are ranked. The topped ranked exemplars t are taken for validation only if the exemplars are positive.

Location Prior

The best possible exemplars are considered for segmentation transfer. Saliency map is computed and the output of the computation is used to find the soft segmentation mask. SIFT flow warping is used to find the soft segmentation mask.

Segmentation

The binary segmentation for the desired object is computed based on the retrieve positive exemplars and the location prior for each positive object proposal. We update our model at every iteration in order to minimize the energy framework.

Algorithm II

1. Energy function is determined by using the input from the appearance model and the exemplar data.
2. Formulate the pair wise potential.
3. Evaluate the potential to find the color of the pixel.
4. Use Location prior to find the foreground and background for the desired selection.
5. Positive exemplars contribute to Appearance model.(APM)
6. Input of (APM) is used to find the location prior of an object.
7. Check for the probabilistic nature of the soft segmentation obtained.

RESULTS

We have tried various datasets and viewed the results based on qualitative and quantitative results. Our approach was found to give better result when compared to the other state-of-art approaches.

CONCLUSION

In this paper we have tried a new approach where we focus on reducing the storage space by using WBCT approach and FCM algorithm in maintaining the exemplar retrieval database. Object Localization and energy minimization framework are used to identify and classify and validate the objects. Our approach shows better results when compared to many other different datasets.

REFERENCES

[1]R.Sujatha,E.Balraj, Interactive Object Selection and Segmentation,Journal of Chemical and Pharmacetical Sciences,2017.
[2]A.Levin,D.Lischinski,andY.Weiss.A closed form solution to natural image matting. In *CVPR*,2006.
[3]A.Joulin,F.Bach,andJ.Ponce.Multi class co-segmentation.In *CVPR*,2012.
[4]J.Shotton,J.Winn,C.Rother,and A.Criminal.Text on boost: Joint appearance, shape and context modeling for multi-class object recognition and segmentation. In *ECCV*, 2006.
[5]K.E.A.vandeSande,J.R.R.Uijlings,T.Gevers,and A.W.M.Smeulders.Segmentation as selective search for object recognition. In *ICCV*,2011.
[6]M.Rubinstein, A.Joulin, J.Kopf,and C.Liu. Unsupervised joint object discovery and segmentation in internet images.*CVPR*,June2013.
[7]Q.Yan,L.Xu,J.Shi,andJ.Jia.Hierrchical saliency detection.In *CVPR*,2013.
[8]Rother,V.Kolmogorov,and A.Blake. Grabcut: inter-active foreground extraction using iterated graphcuts. In *SIGGRAPH*, volume23,pages309–314,2004.
[9]Semantic Object Selection,Ejaz Ahmed,Scott Cohen,Brain Price,IEEE Explore,2014.
[10]C. B. Akgül, D. L. Rubin, S. Napel, C. F. Beaulieu, H. Greenspan, and B. Acar, "Content-based image retrieval in radiology: current status and future directions," *Journal of Digital Imaging*, Vol. 24, 2011, pp. 208-222.
[11]H. L. Tang, R. Hanka, and H. H. Ip, "Histological image retrieval based on semantic content analysis," IEEE Transactions on Information Technology in Biomedicine, Vol. 7, 2003, pp. 26-36.

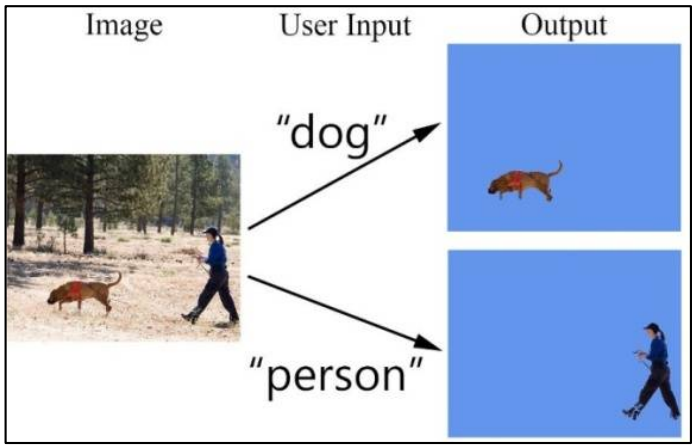


Figure 1. Input image and object segmentation using our approach.

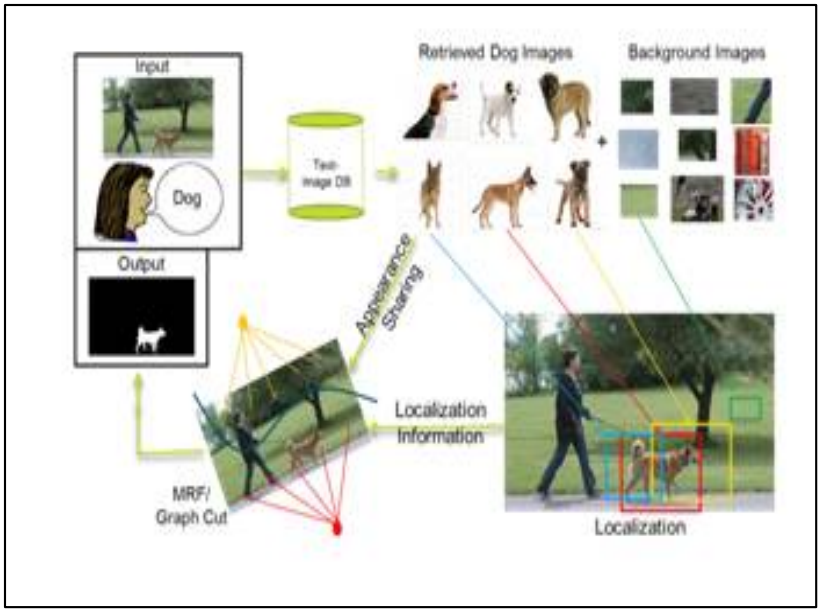


Figure 2. User input, Positive Exemplars and Localization process

Figure 3. Database with Positive Exemplars.

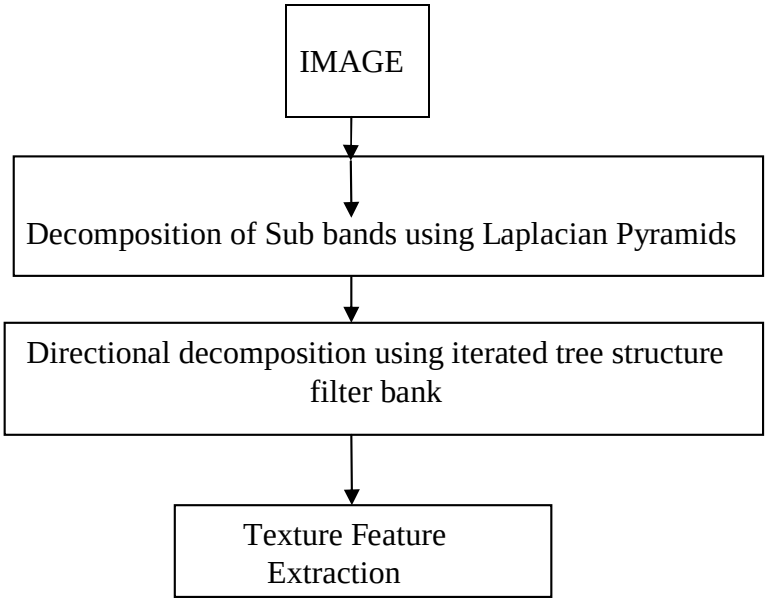


Figure 4.WBCT Block diagram

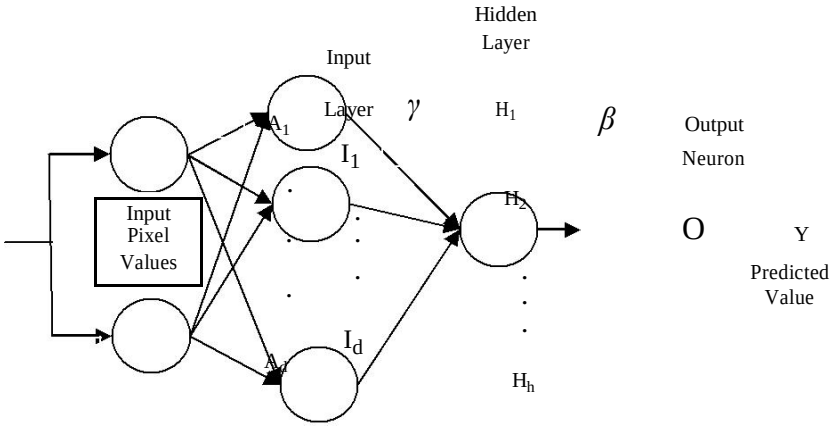


Fig. 5. Neural network structure.

A_{11}	A_{12}	A_{13}	. . .	A_{1N}
A_{21}	A_{22}	A_{23}	. . .	A_{2N}
A_{31}	. . .	. . .	. . .	. . .
. . .	. . .	. . .	. . .	. . .
A_{N1}	A_{N2}	A_{N3}	. . .	A_{NN}

6(a) Original image matrix (A)

0	0	0	0	0
0	A'_{22}	A'_{23}	. . .	A'_{2N}
0	A'_{32}	. . .	. . .	. . .
0	. . .	. . .	...	....
0	A'_{N2}	A'_{N3}	.....	A'_{NN}

6(b) Predicted matrix (A')

$PE_{11} = A_{11}$	$PE_{12} = A_{12}$	$PE_{13} = A_{13}$	. . .	$PE_{1N} = A_{1N}$
$PE_{21} = A_{21}$	PE_{22}	PE_{23}	. . .	PE_{2N}
$PE_{31} = A_{31}$	. . .	. . .	. . .	. . .
....	. . .	. . .	...	....
$PE_{N1} = A_{N1}$	PE_{N2}	PE_{N3}	. . .	PE_{NN}

Fig.7. PE matrix of an image.

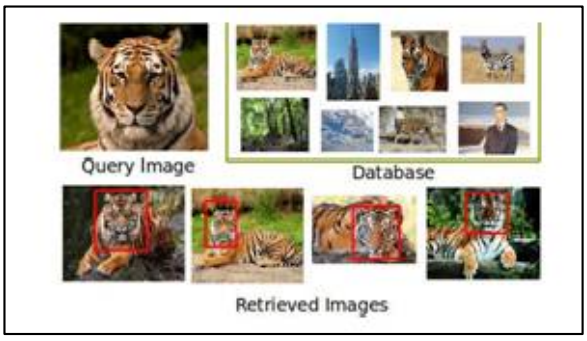


Figure 8. Object retrieval

Figure 9.Validation.Green box indicates positive results and red box indicates negative results.

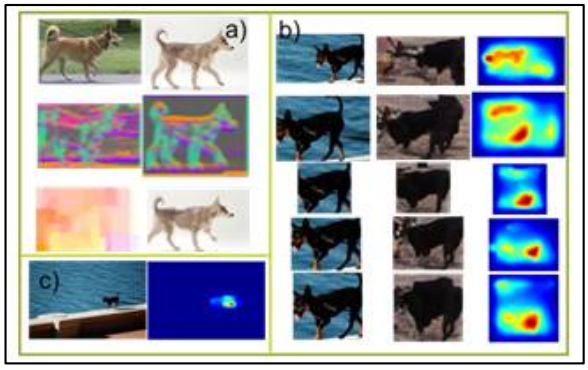


Figure 10. Mask Transfer

Table1: Results on MSRC data set. Comparison of our approach against other state-of-art methods.

class	our approach	[8]	[6]	[7]
bike	57.1	54.1	43.3	29.9
bird	61.1	67.3	47.7	29.9
car	67.2	66.7	59.7	37.1
cat	68.1	66.2	31.9	28.7
chair	59.8	62.2	39.6	28.7
cow	78.1	79.4	52.7	33.5
dog	71.3	67.5	41.8	33

Indian Journal of Natural Sciences *www.tnsroindia.org.in ©IJONS*

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Location Based Clustering of Proximity Nodes Using Prefix Hash Tree

S.Gowsika*

Assistant Professor, Department of Information Technology, M.Kumarasamy college of Engineering, Karur, TamilNadu,India.

Received: 05 Feb 2018 Revised: 15 Feb 2018 Accepted: 23 Mar 2018

***Address for correspondence**
S.Gowsika
Assistant Professor, Department of Information Technology,
M.Kumarasamy college of Engineering, Karur,TamilNadu,India.
E mail:gowsikas.it@mkce.ac.in

ABSTRACT

Peer to Peer Network is a distributed network where the insertion and removal of any node does not have any impact on network. P2P is mainly used for sharing of files such as audio, video, or data, File sharing application in P2P mainly relies on the distributed hash table which distributes the responsibility of storing data on very large set of node and access them very efficiently. DHT uses a structured overlay which maps key to the node. The main problem that relies in using this data structure is that it works well with the exact match of queries in a straightforward approach. To overcome this inefficiency and to improve the file storing, file querying and load balancing in peer to peer network a concept called Locality Aware interest clustering is introduced where the peers are clustered according to the node proximity based on the location. Second, it constructs an overlay of the network by using the cluster to reduce node overload. Third, it reduces the file sharing and file querying delay by using efficient data structure called Prefix Hash Tree a tire based data structure that enables more refined quires over DHT. PHT is resilient than DHT. To further increase the file query efficiency, and to improve load balancing a cost effective file replication algorithm is introduced to reduce the cost as well as to improve the query efficiency.

Keywords: Peer to Peer, Prefix Hash Tree, File replication, cluster

INTRODUCTION

Peer to Peer network is known for its scalability, reliability and ease of administration. In this kind of network architecture there is no need of any centralized server, the peers can act as both client and server, it means that P2P network is more scalable that is when the number of clients increases when the number of server's increases linearly. Peer to Peer network is generally referred as an Overlay network which is on the top of an ip infrastructure. There are two different kinds of P2P network they are Unstructured and Structured Peer to Peer network.

Unstructured Peer to Peer network: In this type the overlay associations are established randomly. The popular unstructured p2p network such as freenet and Gnutella When a peer wants to discover the file, the requests are flooded through network to discover as many peers as possible that share the data. This flooding process creates an enormous amount of signal traffic. In this type when a file become older there is no assurance that file will be found.

Structured Peer to Peer network: This type of network can overcome the drawbacks of the unstructured p2p with its scalability, reliability, efficiency and data location. The most common form of structured p2p is the Distributed Hash table which guarantees a structured pattern of overlay associations DHT is a lookup service, which allows any participating node to competently get back the value related with a given key whether the file is new or older. This will be efficient even when there is a continuous arrival and departure of nodes.

The main function of the Distributed hash table is to plot a given key to node in the network asset the object related with the key. The lookup (key) operation is used to carry the put(key, value) and get(key) hash table operations. The main problem that relies in using this data structure is that it works well with the exact match of queries in a straightforward approach.

Due to this inefficiency of DHT in this paper an efficient data structure called prefix hash indexing is used which is more resilient than DHT. Prefix hash tree uses a lookup interface of DHT to build a tire based data structure which is more efficient than DHT.

Related Works

Here we converse about the works that are related to the data structure used in structured peer to peer and about the proximity awareness, file replication and load balancing techniques. Robert Morrisz [3] provides a solution to the problem that confronts Peer to Peer application i.e. The problems like decentralization, availability, load balancing and supple naming. To overcome these kinds of problems he proposed a structured p2p network protocol called chord. It is one of the most popular, scalable, lookup protocols in structured Peer to Peer network. Chord has some library functions. It provides lookup (key) which yields IP address of a peer with respect to its key. Chord uses finger table as its routing table and skip list as its routing data structure. Query will be passed to the successor node when a client queries a key to find its successor (k). In case key is not found the query time will be increased to O(N). To avoid this situation by using a finger table a fastest searching method is implemented. The main disadvantage of this method is that it handles the queries which match exactly and maximum request is for replicated files.

Paul Francis [6] proposes one of the other DHT systems called Scalable Content Addressable Network. The main motive behind this paper is to provide a scalable indexing system for large scale decentralized storage application on the internet. CAN use a d dimensional Cartesian space coordination system, where the keys are hashed to map on to the values. The main operations performed in this DHT overlay protocol are insert (key,Value) and retrieve(key). In this paper francis deals with CAN construction, routing, and maintenance. The performance is simulated and it is compared.

Shen [2] comes up with an idea for locality aware services which deals with the clustering of resources and for the competent and healthy location in a grid system an algorithm called discovery algorithm is introduced. For load balancing concept liu[5] proposed a secure load balancing which balance the load among the nodes in a hierarchical manner.

Sylvia Ratnasamy[7] comes up with a scheme called distributed binning scheme. The main motive is to optimize routing path in network. This scheme requires a well known landmark machine stretch across the internet. For extracting the proximity information network latency is used. Here, the nodes are partitioned into bins. In this partition the peers that drop within a given bin will be relatively close. From a bin of nodes, it measures the round

trip time in the increasing order of RTT to each landmark. In this paper Ratnasamy deals with three applications of distributed binning scheme, they are on structured overlay, unstructured overlay and server selection and compared these three applications with their experimental result and concluded that the nature of underlying topology affects the behavior of the scheme.

Daniel Baue [1] proposes a replica placement algorithm that works on the Distributed hash table. The novel method used here is replica enumeration in a fully decentralized manner. This algorithm does not depend upon the location. The random probing is used and the simulation result shows the performance of the system in a logarithmic sequence.

Haiying Shen [2] comes up with an idea of clustering peers by using a technique called proximity aware interest clustering t in structured peer to peer network where the clustering is done by both the proximity of nodes as well as the peer interest. In this method he uses a Distributed data structure called DHT for indexing and mapping of (key, Value) pair.

PROPOSED METHODOLOGY

In proposed work to improve the file storage and file querying process a new concept called LAIC is introduced where the peers are clustered in the location based on proximity and the interest of the file. Here the files are distributed among the cluster. Since this work goes in peer to peer network there are no permanent server, the server and client's changes according the node resilient and load in a particular node. Among the peers in cluster, a node with a high storage capacity is selected as a server where the distributed file ids are stored. For organizing and storing files and to map node id a hashing method called prefix hash tree is used.

Prefix hash tree is a routine binary tree constructed on the top of the DHT using the data set, it consists of N numbers and D bit binary keys that are indexed by PHT. The originality of this hashing method relies on how the vertices are mapped to the nodes. This becomes possible when the prefix label of PHT nodes is hashed over the identifier space of DHT. One of the most advantages in using PHT is direct access. For example consider the node ticket m is assigned to the node to which m is mapped to peer whose identifier is nearer to HASH (m). This labeling and assignment shows that it feasible to situate its equivalent PHT peer via a Single lookup using DHT. The above process shows the property of direct access. Some other invariant propertied involved in PHT are key storage, universal prefix, split, merge, threaded leaves. The operations involved in PHT are lookup; lookup operation is performed in two ways lookup linear and lookup binary. By using Prefix hash tree it supports range queries, load balancing and node failure.

Range Queries

In PHT range queries can be implemented in many ways, one of the methods is parallelizing. In this method the node with label corresponds to the lowest prefix range which is within the given rage is located. Using DHT if the located node is an internal node then the quires are forwarded recursively to those children who matches with the given range.

Load Balancing

By considering uniform lookups, PHT allows load to spread over 2 D/2 nodes. Thus the bottle neck problems are eliminated here.

Fault Tolerance

In a traditional hash table and tree models the top down traversal is involved. So in this case if the internal node losses the whole sub trees will be rooted to failure node. This kind of problem is avoided by using PHT; it allows direct access instead of traversal. So the failure of any node in PHT does not affect the data stored in other nodes.

File Access

When a requester is requesting for a file it first checks with the cluster server, whether the requested file is available in the cluster or not by using the id that matches with the incoming query. If the incoming query matches with the id's of file stored in the server the file path is discovered easily and the file can be accessed with less delay. In-case the requested file is not found in the particular cluster it goes with the next cluster. On other hand the most popular and files with high interest are replicated at the server.

File Replication

For handling the file replication an algorithm called cost effective file replication is used. In this algorithm the query rate and the threshold value is considered. If the query rate of the file is high that file is taken as a popular file and for that file the replicas are created. If the query rate and the user interest of a file becomes less then it becomes unpopular file and the replicas will produce more load. To Balance this load and to avoid the crash of peer the cost effective replication algorithm helps in removing the underutilized replicas.

Performance Analysis

This methodology aims to show that Locality aware interest cluster uses Prefix hash tree as a distributed algorithm and works well with the load balancing, range queries by avoiding the issues generally occurred in structured peer to peer network. To test this method the PHT lookups on the scattered keys and lookup traffic is observed. To improve the efficiency in replicated and removing the underutilized replicas of a file, a simulation analysis is carried and the graph is given below.

CONCLUSION

In this paper the new concept called cost effective replication in locality aware interest cluster to improve the file querying, load balancing and node failure in peer to peer network is discussed. Several methods will organize the data stored in peer but Prefix hash tree adds an extra benefit to store and retrieve the data that are stored in the distributed file system. PHT mainly helps in dealing with lookup, resilient and handling the quires in an efficient manner. However to avoid the delay in getting the most popular files the replicas are created, and also the underutilized replicas are also removed.

REFERENCES

1. D.Saravanan, Clustering of video information using brich methodology in Pakistan journal of biotechnology(2017)
2. Daniel Bauer, Paul Hurley, Marcel Waldvogel "Replica Placement and Location Using Distributed Hash Tables, 32nd IEEE conference on local computer networks,(2007) pp.315-324
3. Haiying Shen, Senior Member, IEEE, Guoxin Liu, Student Member, IEEE and Lee Ward "A Proximity-Aware Interest-Clustered P2P File Sharing System" in ieee transactions on parallel and distributed systems, vol. 26, no. 6, june (2015)
4. I.Karthika ",Safeguarding Scada Network by Captious State Based Filter", International Journal of Innovative Research in Computer and Communication Engineering Volume 2 Special Issue 1 -ISSN(Online): 2320-9801 (2014)

5. I.Stoica, R.Morris, D.Liben-Nowell, Chord:A scalable peer to peer look up protocol for Internet Application" in IEEE/ACm Transactions on Networking, vol.11 Issues: 1, Feb(2003)
6. Liu, X., D. Yuan, G. Zhang, J. Chen and Y. Yang, 2010.SwinDeW-C: a peer-to-peer based cloud workflow system. In Handbook of Cloud Computing, pp: 309-332. Springer US
7. Naresh, " An efficient resource allocation strategy based on Improved Particle Swarm Optimization (IPSO) in Pakistan journal of biotechnology(2016)
8. Sylvia Ratnasamy, Paul Francis, Mark Handley, Richard Karp, "A scalable- Content Addressable Network" in ACM Digital Libarary
9. Sylvia Ratnasamy, Mark Handley, Richard Karp, Scott Shenker, Topologically-Aware Overlay Construction and Server Selection" in IEEE INFOCOM, sep(2002)
10. Vinothan Damodharan and saravanan M,"Institution system analysis by using similarity based clustering on social network access", in Pakistan journal of biotechnology(2016)

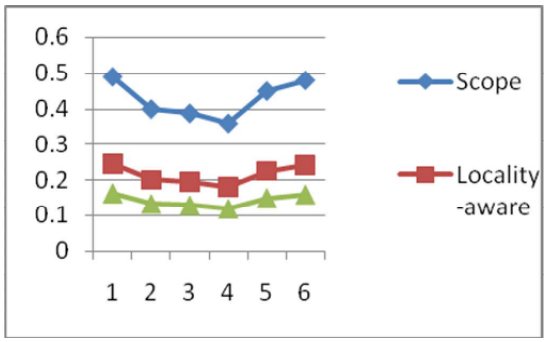


Fig.1. Replication analysis

Indian Journal of Natural Sciences *www.tnsroindia.org.in ©IJONS*

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Innovative Air Pollution Monitoring with Sensor Controlled Wireless Communication

Karthik.V[1]* and Swapna. P[2]

[1]Assistant Professor, M.Kumarasamy College of Engineering (Autonomous), Karur,TamilNadu,India
[2]Assistant Professor, Sethu Institute of Technology, Viruthunagar,TamilNadu,India.

Received: 12 Feb 2018 Revised: 23 Feb 2018 Accepted: 23Mar 2018

***Address for correspondence**
Karthik V
Assistant Professor, M.Kumarasamy College of Engineering (Autonomous),
Karur ,TamilNadu,India

ABSTRACT

Air pollution has been a major threat to the human welfare that too specifically smokes from vehicles. A system composed of wireless sensor network that keeps an eye on the pollution on the surroundings is our main theme. The system has a Mobile Data Acquisition unit backed up by a pollution server that stores the data related. The system also includes GPRS and a GPS for location based and a network based needs. The Mobile DAQ unit collects the pollution details using the sensors and sent to the pollution server through GPRS service along with the GPS co ordinates and after accumulation of huge data the dataset will be taken and a machine learning algorithm is applied for the better results and effective suggestions or the recovery of pure air.

Keywords: WSNs, Mobile Data-Acquisition Unit (Mobile-DAQ), GeneralPacket Radio Service Modem (GPRS-Modem), Global PositioningSystem Module (GPS-Module).

INTRODUCTION

Wireless sensor networks has got a unique place in many areas including military, surveillance cameras and modern day communication which becomes unavoidable but few years back mobile communication is just used rarely. Many areas like soil environment monitoring, that helps in monitoring the temperature humidity of the soil and also the moisture content and few more attributes that led to the healthy growth of the cultivation.

Mobile Cloud Computing

Cloud computing which takes over the problem of losing the data while it is in the local machine has been vanished and now our data is on the air and it can be accessed anywhere you go and just an internet connection is more than enough and a device to access it.

Mobile computing does not need an introduction yet the back bone of our communication networks and has been performing exceptionally well in connecting people. We would like to merge these two and take in the advantages of the above and trying to get the best out of them. Data will be managed on the cloud servers. And the mobile communications will be necessary to transmit the data from the sensors to the cloud servers.

Wireless Air Pollution Monitoring System (WAPMS)

Due to rapid development of the industry vehicles and many more artificial things that are tend to reduce human works are now a days indeed producing more and more pollution and it is indirectly consumed by the human beings which will eventually reduce the life span. People who are working closely with machines produce lot of smokes are more likely get Lung Cancer. We had an idea to monitor the pollution and its severity over different regions and try to give a better suggestion to downgrade the pollution amount on the same area thereby helping mankind and the environment. The database servers deployed on the cloud server will always be ready to receive the data from the mobile networks used by the sensor and the devices. At a later point of time we will take the big data set and try a predictive analytics and find a possible cause and a possible solution for it.

Reading Sensor

A random value is generated on value that holds seriousness.

Reading Transmitter

Value is read and transmitted through communicator.

Power Controller

A method called as "Node On" will be present on every node and it can be called when needed.

Communicator

Interprocess communication with send and receive methods are done by using Socket programming.

Launcher

Data collector is been informed to collect based on delivery mode.

Data Collector

The nodes that are supposed to collect readings will be noted and receives the values.

Aggregator

RCQ algorithm belongs to data aggregator.

Connection Initiator

JDBC that is Java Database Connectivity that is deployed to connect the database with the front end and the parameters like username and password etc.

Connection Destructor

At the end of the code the connection object is supposed to be closed.

Energy Defined Air Pollution Monitoring Scheme for Secure WSN

A passive sampler is selected to sample the gaseous pollutants and then it is sent into the chemically treated surface so that the nature of gas sent through it will have an impact and will have a change that can be recorded and the nature of the gas is found.The above system can be tried in much more efficient method like using the paper tape samplers and passing the gas forcibly into the surface using a pump so that we will get much more accuracy and efficiency. Gases like H2S, HCN and many more things are expected. The gaseous vapours are then collected and sent to the laboratory for various tests.

Instrumental Monitoring Methods

The most common pollution monitoring systems are using the image processing techniques and the data collected based on the color contrast of the image taken. We had given a different approach to it and got some data that gives us faith towards the progress in reducing the pollution.

CONCLUSION

The evolution of Environmental Sensor Networks has given faith in helping the mankind in monitoring and thereby reducing the adverse effects of both the machines that invented by man as well as the waste that machine produces. We largely produce air pollution from vehicles. Our future scope will be immediately after collecting the data the efficient solutions will be searched and we will try our maximum extent to reduce the pollution over the particular area where it is supposed to be smoke clouds may be during the particular time based on various external factors.

REFERENCES

1. C. H. Ettema and D. A. Wardle, "Spatial soil ecology," Trends in ecology & evolution, vol. 17, pp. 177-183, 2002.
2. H.-j. FANG, X.-m. YANG, X.-p. ZHANG, and A.-z. LIANG, "Spatial Heterogeneity and Pattern of Black Soil Organic Carbon of Sloping Field [J]," Bulletin of Soil and Water Conservation, vol. 3, p. 005, 2005.
3. R. B. Franklin and A. L. Mills, "Importance of spatially structured environmental heterogeneity in controlling microbial community composition at small spatial scales in an agricultural field," Soil Biology & Biochemistry, vol. 41, pp. 1833-1840, 2009.
4. P. Gao, T. FU, K. Wang, H. Chen, and F. Zeng, "Spatial Heterogeneity of Surface Soil Nutrients in a Small Catchment in Karst Peak-cluster Depression Area," Reseach of Agricultural Modernization, vol. 34, pp. 362-366, 2013.
5. Hsu-Cheng Lu , Jen-Hao Liu , Joe-Air Jiang ; Tzai- Hung Wen , Chih-Hong Sun , Jehn-Yih Juang ,"Application Of A Reliable MAC Protocol For The Urban Air Quality Monitoring System Based On The Wireless Sensor Network", IEEE Conference, pp. 1-6, (March 2012)
6. R. John, J. W. Dalling, K. E. Harms, J. B. Yavitt, R. F. Stallard, M. Mirabello, S. P. Hubbell, R. Valencia, H. Navarrete, M. Vallejo, and R. B. Foster, "Soil nutrients influence spatial distributions of tropical tree species," Proceedings of the National Academy of Sciences, vol. 104, pp. 864-869, January 16, 2007.
7. Saravanan, D., Information Retrieval using: Hierarchical clustering algorithm. International Journal of Pharmacy and Technology 8(4): 22793-22803 (2016).
8. Saravanan, D., Video content Retrieval using image feature selection. Pak. J. Biotechnol. 13 (3): 215-219 (2016),

9. H. Stoyan, H. De-Polli, S. Böhm, G. P. Robertson, and E. A. Paul, "Spatial heterogeneity of soil respiration and related properties at the plant scale," Plant and Soil, vol. 222,pp. 203-214, 2000.
10. J. Yu, "Spatial variability of soil nutrients and Application Research Based on GIS and geostatistics," PhD, Wuhan: Huazhong Agricultural University, 2007.
11. F. Zhang, J. Shen, J. Jing, L. Li, and X. Chen, "Rhizosphere Processes and Management for Improving Nutrient Use Efficiency and C'rop Productivity," in MolecularEnvironmental Soil Science at the Interfaces in the Earths Critical Zone, J. Xu and P. Huang, Eds., ed: Springer Berlin Heidelberg, 2010, pp. 52-54.
12. J. Zhao, H. Liu, Y. Sui, X. Zhang, and K. Meng, "Analysis for Spatial Heterogeneity of Organic Matter Content and Available Nutrients in Blacksoil Crop Area with Different Scales," Journal of Soil and Water Conservation, vol. 20, pp. 41-44,62, 2006
13. X. Zhang, Y. Sui, X. Zhang, K. Meng, and S. J. Herbert, "Spatial Variability of Nutrient Properties in Black Soil of Northeast China," Pedosphere, vol. 17, pp. 19-29, 2007.

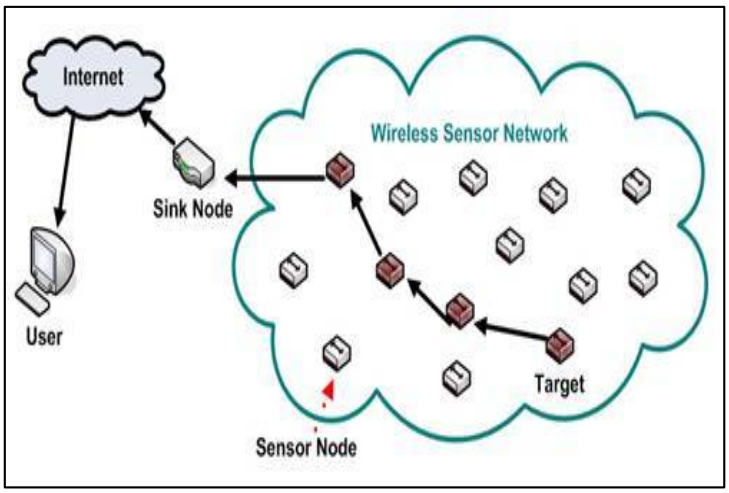


Fig:1.Wireless sensor network

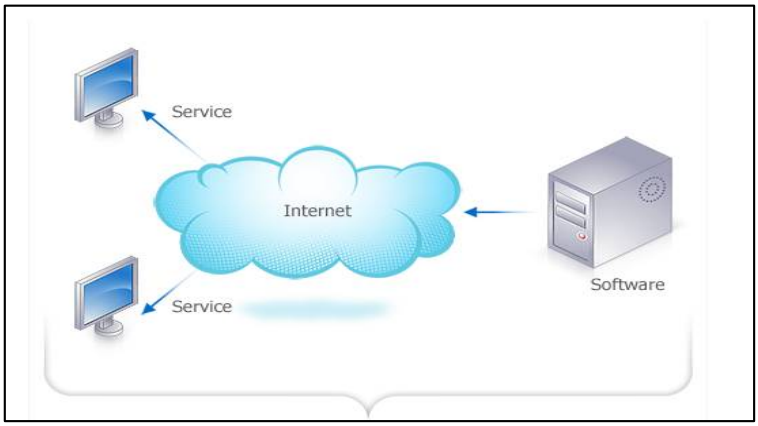
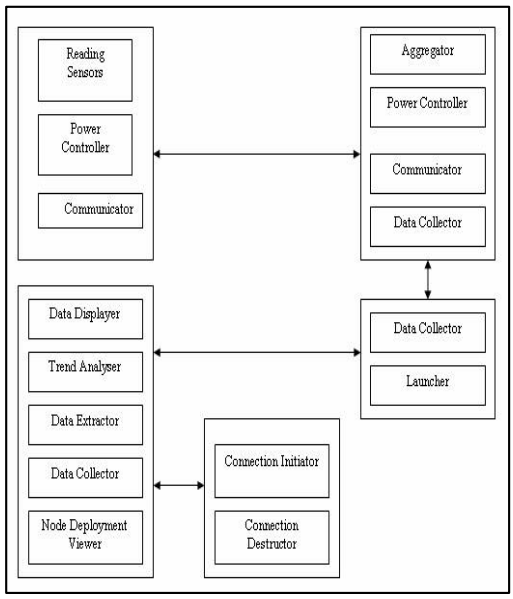


Fig:2.Mobile cloud computing



ARCHITECTURE DIAGRAM OF WAPMS

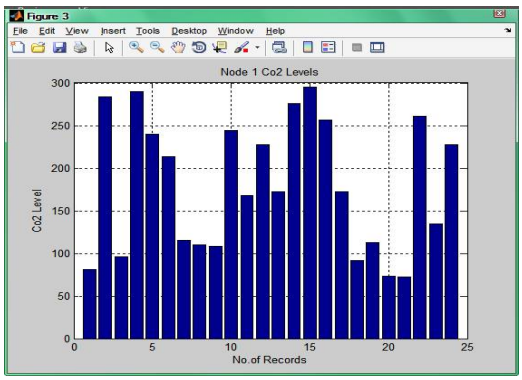


Figure 3:Co2 Levels

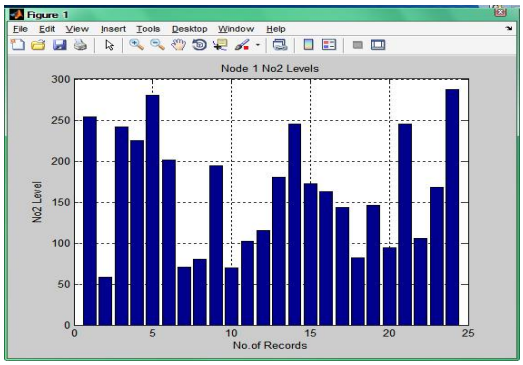
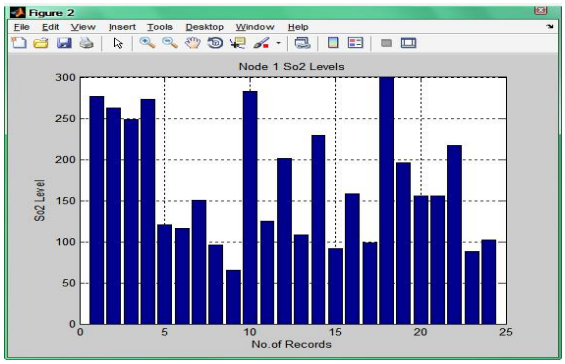


Figure 3:No2 Levels



PERFORMANCE AND SIMULATION STUDY

Indian Journal of Natural Sciences www.tnsroindia.org.in ©IJONS

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

A Data Destruction in Cloud Based Multi-Tenant Database Deduplicatable

S.Keerthi[1]* and P.Rajesh Kanna[2]

[1,2]Assistant Professor, Department of Computer Science and Engineering, M. Kumarasamy College of Engineering, Karur,TamilNadu,India.

Received: 10 Feb 2018 Revised:15Feb 2018 Accepted: 30 Mar 2018

***Address for correspondence**
S.Keerthi
Assistant Professor,
Department of Computer Science and Engineering
M. Kumarasamy College of Engineering,
Karur, TamilNadu, India.
E.mail: keerthisridhar77@gmail.com, mailmeatrajeshkanna@gmail.com

ABSTRACT

Conceptual Applications being convey as administration, equipment and server farms which give administrations. Distributed storage rise up out of a capacity model to another administration demonstrate where information is being overseen, kept up, and put away in various remote servers for move down reasons. Cloud stage server groups are running in system condition and it might contain various client information and the information might be scattered in various virtual server farms. In a multi-client shared distributed computing stage clients are just legitimately disengaged, yet information of various clients might be put away in same physical gear. These types of gear can be expediently provisioned, actualized, scaled up or down and decommissioned. A helpful multi-customer conveyed capacity structure needs the secured client side cross-customer deduplication technique, which allows a customer to skirt the exchanging strategy and achieve the duty regarding records immediately, when diverse proprietors of comparative archives have exchanged them to the cloud server.

Keywords: proof of storage, cross user deduplication technique, Symmetric Data Key.

INTRODUCTION

In the PC business, the word security or the expression PC security hints procedures for guaranteeing that informational collection away in a PC can't be inspected or managed by any people without support. Most PC prosperity tries mean information encryption and passwords. Information encryption is the interpretation of information into an edge that is dubious without a translating instrument. A watchword is a puzzle word or expression that gives a customer access to a foreordained program or system. In1983, Kevin Mitnick did an interruption on a Pentagon's PC Robert Tappan Morris made the fundamental worm and sent it from MIT to the web

and made $50,000 of harms. In 1994, Vladimir Levin bounced in an American bank PC and stole 10 million dollars Jonathan James "confidant", 16 years of age, entered a NASA PC in 1999 and had segment to information worth 1,7 Million dollars Today (CSI Report, 2007) 46% of affiliations have confessed to torment money related incidents by virtue of security.

Domain Introduction

Disseminated figuring is a preparing perspective, where a far reaching pool of structures are related in private or open frameworks, to give dynamically flexible establishment to application, data and record stockpiling. With the presence of this advancement, the cost of computation, application encouraging, content stockpiling and transport is diminished in a general sense. Distributed computing is a pragmatic way to deal with experience coordinate money saving advantages and it can possibly change a server farm from a capital-escalated set up to a variable estimated condition. Cloud registering depends on an extremely crucial foremost of reusability of IT capacities. The distinction that distributed computing conveys contrasted with conventional ideas of "matrix processing", "disseminated figuring", "utility registering", or "autonomic processing" is to widen skylines crosswise over authoritative limits. Forrester characterizes distributed computing as: "A pool of disconnected, very versatile, and oversaw register framework equipped for facilitating end client applications and charged by utilization." Distributed computing is an advancement that uses the web and central remote servers to keep up data and applications. Conveyed processing grants buyers and associations to use applications without foundation and get to their own particular records at any PC with web get to. This innovation takes into account significantly more productive figuring by unifying information stockpiling, preparing and transfer speed. A basic case of distributed computing is Yahoo email, Gmail, or Hotmail and so on.

Cloud Computing Services

Cloud Provider offer administrations are classified into three categories.

Software as a provider (SaaS)

In this model, whole software is obtainable to the customer, as an affiliation on request. A solitary event of the affiliation maintains going for walks on the cloud and unique end clients are adjusted. On the consumer side, there's no important for sincere eagerness for servers or programming licenses, whilst for the supplier, the costs are chopped down, considering simply a utility need to be empowered &maintained. Today saas is obtainable with the aid of dating, for example, google, income oblige, Microsoft, zoho, and so forth.

Platform as a carrier (PaaS)

Here, a layer of programming or exchange situation is typified &provided as an organisation; where upon other extra lifted measures of business enterprise may be made.The purchaser has the opportunity to keep up his very own applications, which maintain running at the provider's shape. to meet sensibility and adaptability requirements of the applications, PaaS suppliers offer a predefined mix of OS and application servers, for example, LAMP mastermind (Linux, Apache, MySql and php), obliged J2EE, Ruby and whatnot.Google's App Engine, force. com and so forth are some of understood PaaS diagrams.

Framework as a carrier (IaaS)

IaaS provider's key stockpiling and looking after cut-off points as institutionalized associations over the machine. Servers, stockpiling frameworks, masterminding tools, server develop space and whatnot. They're pooled and made open to supervise workloads. patron would with the aid of and large ship his very own particular programming in this foundation.

LITERATURE SURVEY

Compact proofs of retrievability

In a proof-of-retrievability framework, an information stockpiling focus must display to a verifier that he is genuinely securing the majority of a customer's information. The focal test is to assemble structures that are both competent and provably secure that is, it ought to be conceivable to remove the customer's information from any prover that passes an attestation check.

Practical dynamic proofs of retrievability

Confirmations of Retrievability (PoR), proposed by Juels and Kaliski in 2007, engage a client to store n archive hinders with a cloud server so that later the server can show responsibility for significant number of data in a to a great degree successful route (i.e., with relentless computation and information exchange limit). Though various capable PoR gets ready for static data have been produced, only two components PoR arranges exist**.**

Proofs of ownership in remote storage systems

Dispersed capacity structures are winding up observably continuously predominant. A promising improvement that holds their cost down is deduplication, which stores just a solitary duplicate of rehashing information. Customer side deduplication endeavours to perceive deduplication openings beginning at now at the customer and additional the data trade breaking point of trading duplicates of existing reports to the server. In this work we perceive strikes that endeavour customer side deduplication, enabling an assailant to get to discretionary size records of different clients in context of a little hash indications of these reports. All the more particularly, an assailant who knows the hash indication of a record can impact the point of confinement preferred standpoint that it ensures that report, from this time forward the server allows the aggressor to download the entire record.

Towards efficient proofs of retrievability in cloud storage

Confirmations of Retrievability (POR) is a cryptographic technique for remotely evaluating the respectability of documents put away in the cloud, without keeping a duplicate of the first records in nearby stockpiling. In a POR plot, a client Alice reinforcements her information document together with some confirmation information to a possibly deceptive distributed storage server Bob. Afterward, Alice can intermittently and remotely check the trustworthiness of her information put away with Bob utilizing the validation information, without recovering back the information document amid a confirmation

A scalable cloud files system with efficient integrity checks

We show Iris, a handy, validated record framework intended to bolster workloads from extensive ventures putting away information in the cloud and be versatile against possibly deceitful specialist organizations. As a straightforward layer implementing solid uprightness ensures, Iris gives an undertaking a chance to occupant keep

up a huge record framework in the cloud. In Iris, occupants get solid confirmation on information honesty, as well as on information freshness, and also information retrievability in the event of inadvertent or ill-disposed cloud disappointments. Iris offers an engineering adaptable to numerous customers (on the request of hundreds or even thousands) issuing operations on the document framework in parallel. Iris incorporates new improvement and endeavor side reserving methods particularly intended to beat the high system dormancy normally experienced when getting to distributed storage. Iris additionally incorporates novel eradication coding procedures for effective support of element Proofs of Retrievability (PoR) conventions over the document framework.

Existing System

Data reliability is a champion among the most key properties when a customer outsources its records to disseminated stockpiling. Customers should be influenced that the records set away in the server are not modified. Regular techniques for guaranteeing data respectability, for instance, message affirmation codes (MACs) and propelled marks, oblige customers to download the archives from the cloud server for check, which realizes a staggering correspondence cost. These strategies are not suitable for conveyed stockpiling organizations where customers may check the respectability as regularly as would be prudent, for example, consistently. Accordingly, scientists presented Proof of Storage (PoS) for checking the honesty without downloading documents from the cloud server. Besides, clients may likewise require a few element operations, for example, change, addition, and erasure, to refresh their documents, while keeping up the capacity of PoS. Traditional methodology for securing data respectability, for instance, message approval codes (MACs) and electronic imprints, oblige customers to download most of the records from the cloud server for check, which achieves a significant correspondence cost. These frameworks are not suitable for conveyed stockpiling organizations where customers may check. The trust worthiness as often as possible, for example, consistently. Consequently, specialists presented Proof of Storage (PoS) for checking the respectability without downloading documents from the cloud server. Besides, clients may likewise require a few element operations, for example, alteration, inclusion, and erasure, to refresh their documents, while keeping up the capacity of PoS. Dynamic PoS is proposed for such element operations.

Disadvantage:

- Message verification codes brings about an overwhelming correspondence cost
- Dynamic PoS can't encode the piece records into names, since the dynamic operations may change many records of non-strengthened squares.
- There is no dynamic PoS that can support secure cross-customer deduplication

Proposed system

To the best of our knowledge, this is the primary work to display a primitive called deduplicatable component Proof of Storage (deduplicatable component PoS), which comprehends the structure arranged qualities and private mark period challenges.. The proposed technique distinguishes approach to track individual clients' information and their encryption keys and gives answer for totally erase the information from the cloud supplier's multi-occupant stockpiling engineering. It likewise guarantees cancellation of information duplicates as there are dependably potential outcomes of more than one duplicate of information being kept up for move down purposes. The information obliteration confirmation should likewise be given to client ensuring that the proprietor's information is totally evacuated. Erasing information does not erase it from the plate. For the most part, erasing implies that lone the pointers to the information are erased and the space can be reused by simply overwriting on a similar stockpiling range. Until the space is overwritten, the information is not expelled from the plate and can be perused by different projects. We propose and execute the primary productive development of deduplicatable element PoS called DeyPoS, which bolsters boundless number of check and refresh operations. The security of this development is demonstrated in the arbitrary prophet display, and the execution is dissected hypothetically and tentatively.

The System (CPOD: Cloud-Proof of Deletion) that we created on top of the current API gave by the supplier that stores and recover documents from cloud information stockpiling. Any document stockpiling in the cloud is related with set of arrangement records. The arrangement record indicates the imperatives under which particular activity can be performed on specific touchy assets. The documents must be gotten to just when the certifications or keys created by trusted experts are satisfied according to approach.

Modules

Execution is the time of the meander when the hypothetical course of action is changed out into a working structure. In this way it can be thought to be the most basic stage in accomplishing a beneficial new framework and in giving the client, sureness that the new structure will work and be persuading. The utilization sort out incorporates careful masterminding, examination of the present system and it's prerequisites on execution, laying out of techniques to achieve changeover and appraisal of changeover techniques. A module is a piece of a program. Projects are made out of at least one unreservedly made module that is not joined until the program is associated. A lone module can contain one or a couple plans.

Our project modules are given below:

- USER
- ADMIN
- DEDUPLICATABLE DYNAMIC

User

Customers should be induced that the records set away in the server are not modified. Traditional strategies for guaranteeing data trustworthiness, for instance, message affirmation codes and propelled marks, oblige customers to download most of the records from the cloud server for check, which achieves a staggering correspondence cost.

Admin

The cloud server gives back the relating obstructs alongside their labels. The verifier checks the square trustworthiness and file accuracy. The previous can be specifically ensured by cryptographic labels. Step by step instructions to manage the last is the significant contrast amongst PoS and element PoS. In a large portion of the PoS schemes , the square rundown is "encoded" into its mark, which suggests the verifier can check the piece trustworthiness ..

Deduplicatable Dynamic

There are five stages in a deduplicatable dynamicPoS framework: pre-handle, transfer, deduplication, refresh, andproof of capacity. In the pre-prepare stage, clients mean to transfer their localfiles. The cloud server chooses whether these documents shouldbe transferred. In the event that the transfer procedure is without a doubt, go into theupload stage; generally, go into the deduplication phase.In the transfer stage, the documents to be transferred don't existin the cloud server. The first clients encodes the neighborhood filesand transfer them to the cloud server.In the deduplication stage, the documents to be transferred alreadyexist in the cloudserver**.**

CONCLUSION

In this paper, we display the outline and execution of customer, a framework that gives secure and practical document guaranteed cancellation on the distributed storage we proposed the exhaustive prerequisites in multi-client distributed storage frameworks and presented the model of deduplicatable element POS. The execution result demonstrates that the circumstances taken for the total record transfer handle increments exponentially when number of documents and their record measure increments. Our future work will be to decrease the time taken for

key era and approach record the hypothetical and exploratory outcomes demonstrate that our DeyPoS execution is proficient, particularly when the document estimate and the quantity of the tested pieces are vast

REFERENCES

[1]M. Armbrust, A. Fox, Griffith et al., —Above the clouds: A berkeley view of cloud computing,‖ EECS Department,University of California, Berkeley, Tech. Rep. UCB/EECS-2009-28, 2009.

[2]S. Aulbach, T. Grust, D. Jacobs, A. Kemper, and J. Rittinger. Multi-tenant databases for software as aservice:schema-mapping techniques. In SIGMOD, pages 1195–1206, 2008.

[3]F.Chong, G. Carraro, and R. Wolter, —Multi-Tenant Data Architecture,‖ MSDN Library,Microsoft Corporation, 2006.

[4]P. A. Bernstein, I. Cseri, N. Dani, N. Ellis, A. Kalhan, G. Kakivaya, D. B. Lomet, R. Manner, L. Novik, and T. Talius. Adapting Microsoft SQL Server for Cloud Computing. In ICDE, pages 1255–1263, 2011.

[5]D. Jacobs and S. Aulbach, —Ruminations on multi-tenant databases,‖ BTW Proceedings, 2007.

[6]N. Chohan, C. Bunch, S. Pang, C. Krintz, N. Mostafa, S. Soman, and R. Wolski. AppScale: Scalable and OpennAppEngine Application Development and Deployment. In CloudComp, b2009.

[7]C. Curino, E. Jones, R. Popa, N. Malviya, E. Wu, S. Madden, H. Balakrishnan, and N. Zeldovich. Relational Cloud: A Database Service for the Cloud. In CIDR, pages 235–240,b 2011.

[8]S. Das, S. Agarwal, D. Agrawal, and A. El Abbadi. ElasTraS: An Elastic, Scalable, and Self Managing Transactional Database for the Cloud. Technical Report 2010-04, CS, UCSB, 2010.[9]B. Reinwald. Database support for multi-tenant applications. In IEEE Workshop on Information and Software as Services, 2010.

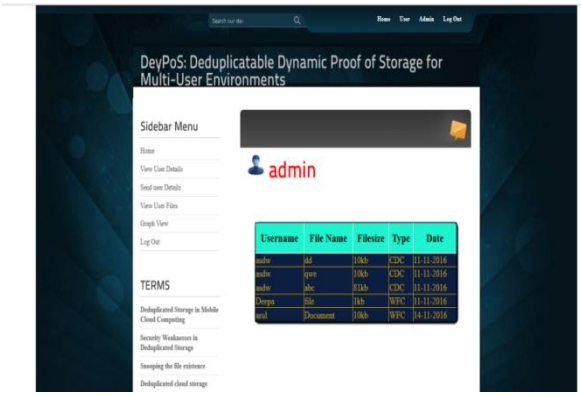


Fig.1: Modules

Fig.2: Login

Fig.3: Upload

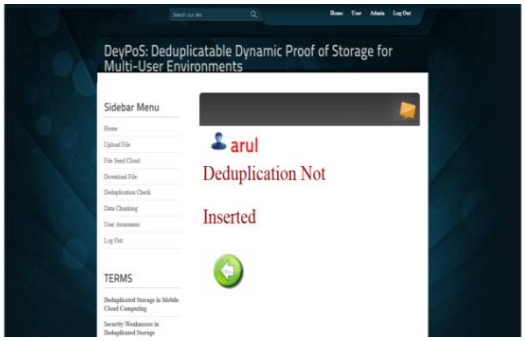


Fig.4: Deduplicatable Dynamic

Indian Journal of Natural Sciences www.tnsroindia.org.in ©IJONS

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

A New Displaying Technology for Information Hiding Using Temporally Brightness Modulated Pattern

Deepa.K*, Prakatheesh.TJ, Rajalingam.P and Ram kumar.M

Department of Computer science and Engineering, M.Kumarasamy College of Engineering, Karur, TamilNadu,India

Received: 10 Feb 2018 Revised:17 Feb 2018 Accepted: 30 Mar 2018

***Address for correspondence**
Deepa. K,
Assistant Professor,
Deaprtment of Computer Science and Engineering,
M.Kumarasamy College of Engineering,Karur,
TamilNadu,India.
E.Mail: deepak.cse@mkce.ac.in

ABSTRACT

This proposes another show framework innovation that can conceal mystery data behind a showed picture, all the while fulfilling both high imperceptibility and comprehensibility prerequisites. The concealed data is a sort of twofold picture including characters, and different sorts of examples, i.e., Quick Response (QR) codes. This method makes utilization of transiently brilliant balanced imperceptible example in a moving picture or video. Likewise this paper examine about the security examination of QR (Quick Response) code scanners on Android.The exploration was made as of late and the outcome demonstrates that some of those QR code scanners were not ready to identify assaults misusing pernicious URLs implanted in QR codes, particularly phishing and malware assaults. This paper investigates on the security issues and the best approach to conquer it. This likewise assesses QR Codes and how they can be utilized for both human cooperation and robotized frameworks. As the encoded data is expected to be machine intelligible just, a human can't perceive between a legitimate and a malevolently controlled QR code. This task depends on Mobile Application, through which the client can filter QR Code utilizing their portable camera. This Project utilizes Java, XML as front end and SQL Server as back end. XML is for outlining the Application and Java is utilized for summoning the application. This will give the great UI

Keywords: Quick Response, Android, legitimate.

INTRODUCTION

Strategies for disguising information in pictures have been ahead of time all through the past two decades by various researchers. They have developed essentially to find the movement respond in due order regarding progressed watermarking. There are distinctive courses for scrutinizing out hid information from cutting edge pictures. In any case, some new applications use modernized cameras, i.e., the propelled picture that imperceptibly contains information inside itself is first gotten with an electronic camera, and after that the information is scrutinized out from the gotten picture by automated taking care of. Previously another application using a modernized camera that got a photo expected onto a real issue by a projector was proposed anyway it has some security issues. This paper gives another technique that examines out covered information in an indicated picture on a sweeping level board appear (FPD, for instance, that for cutting edge signage. This system moreover uses a passing balanced imperceptible case for moving pictures and when understanding them out, diagram pictures over a couple of periods are shown raise the distinction of the vague case so as to make it unmistakable. We endeavoured a system for time move testing to handle this issue. This paper moreover portrays the appraisal of the method we assisted through trials. In like manner this paper discusses the security issues and the ways to deal with recover it. This paper mainly excited about phishing and malware ambushes in light of QR codes. Phishing strikes trap customers to uncover their unstable information by going up against the presence of a solid substance. There are diverse sorts of strikes, for instance, social planning and cross-site ambushes. Regardless of the way that fascinating, a whole report for these strikes is out of area of this paper. The security examination of existing QR code scanners on Android was coordinated starting late and the result shows that some of those QR code scanners were not capable distinguish strikes abusing malicious URLs introduced in QR codes, especially phishing and malware ambushes. In the examination, QR code scanners open at the Google Play were destitute down the extent that their security features and security forewarning capacities, and two awesome QR code scanners, Norton Snap and QR Pal, were picked and attempted against phishing ambushes manhandling poisonous URLs introduced in QR codes. In our past examination, an ensured QR code scanner course of action called Safe QR was proposed to better perceive phishing and malware strikes in light of malicious introduced URLs. Safe QR uses two security APIs, one is Google Safe Browsing API and another Phish tank API, with a particular true objective to upgrade the revelation rate for the attacks.

Related Works

The 10 Commandments or Codes

QR Code is that little square, routinely will be in high difference that one can find progressively constantly now day by day's including film tickets. It can contain distinctive sorts of accommodating information for clients. To outline, QR Codes have ability to extend the information present on a physical help and give a propelled estimation to the publicizing exertion that you are driving. Its impact on buyers is unequivocally extended. Also the visual piece of QR Codes is progressing and getting the opportunity to appeal. For they are used as specific gadgets, QR Codes must intrigue and striking to draw the thought of your customers and propel them to check. QR Codes are examined by implied "scrutinizing applications", which are basically institutionalized ID scanners. There are numerous them that you can clearly download on your Smartphone. The QR Code can be printed at any sizes in any case you ought not to print it under 3cm for each 3cm (1,2 inch).

Qr Code Security: A Survey of Attacks and Challenges for Usable Security

QR (Quick Response) codes are two-dimensional institutionalized recognizable pieces of proof which is pushed variation of Barcode with the ability to encode particular sorts of information. Despite when extent of purposes of intrigue, QR codes poses critical security risks. Aggressors can encode pernicious associations that lead e.g. to phishing regions. Such harmful QR codes can be engraved on little stickers and supplant considerate ones on declaration plugs. Since various certifiable instances of QR code based strikes have been represented in the media, simply little research has been driven in this field and no thought has been paid on the trade of security and human-

PC coordinated effort. In this Paper, we delineate the complex use occasions of QR codes. Additionally, we research the most basic ambush circumstances with respect to the specific use cases.

Quick Response (or) Codes in Libraries: Case Study on the Use of Qr Codes in the Central Library, Nitk

QR codes are 2D organize scanner labels. Heretofore a continuing improvement in business and business arranged usage of these QR codes with the happening to sagacious and web talented PDAs was seen. In any case, using QR codes in Library and Information field is still in its Initial stage. This paper discuss the utilization of QR Codes in Library resource and organization care works out. The present examination bases on point by guide preamble toward QR codes close by how it capacities and its essential features. It in like manner tries to clear up the use of QR codes in libraries, with remarkable reference to QR code executions in Central Library, NITK. This paper fills in as an escort to every last one of the people who need to execute QR Codes in their Library.

Qr Codes in Education

Quick Response (QR) codes are versatile. A long substance, a URL, a SMS message, a business card or essentially any information can be stow away into the 2D scanner tag. Partner with arranged mobile phones, QR Codes would interface be able to the customers to the information quickly and in straightforward way. In this paper, we research how QR codes can be used as a piece of side of preparing. The low specific limit of making and scrutinizing QR codes empowers creative educators to merge them into their educational endeavors. The technique to Store and recoup QR codes are to an awesome degree straightforward and rapid, and with PDAs, make them the ideal informative contraptions for teaching and learning. This paper is dealt with as takes after. In Japan, QR codes are everywhere and considerable number individuals have PDAs furnished with QR code perusers. Disregarding the way that QR codes existed for more than couple of years, there are not too various examination applications here.

CONCLUSION

This proposed another show framework that can hide the information in a demonstrated picture. This uses a transient adjusted imperceptible case for moving pictures and moreover when they are examined out, the edge pictures over a couple of periods are investigated to raise the intricacy of the impalpable case keeping in mind the end goal to make it self-evident.This also proposed a strategy that deals with theissue of asynchrony between the showand camcorder.This drove tests and the results got from these revealed that hid combined pictures could be examined out. Moreover, we also certified that the illustrations used as a piece of this framework were vague. In this manner, we displayed the feasibility of the methodology we propose. In like manner this paper considers the ambushes in QR Codes and ways to deal with recover it. Furthermore this designs the way that it will be useful for trainings additionally.

REFERENCES

1. XiongZou, Guo-dong Liu, Jian-min Wang, “Study on the Sequence of Steps in the QR Code Image Preprocessing,” 2nd International Conference on Future Computer and Communication, 2010.
2. YueLiu,JYang,Mingjun Liu, “Recognition of QR Code with Mobile Phones,” Chinese Control and Decision Conference, 2008, pp. 204 .
3. R.C. Gonzalez and R.E. Woods, DigitalImageProcessing, Prentice Hall, Inc., U.S.A., 2002
4. Denso wave incorporated,” http://www.denso-wave.com/qrcode/index-e.html.
5. Ching-yin Law Simon So, "QR Code in Education" 2010.

6. Iranna M Shettar, " Quick Response (QR) Codes in Libraries: Case study on the use of QR codes in the Central Library, NITK "
7. Peter Kieseberg, Manuel Leithner, Martin Mulazzani, Lindsay Munroe, Sebastian Schrittwieser, Mayank Sinha, Edgar Weippl, " QR Code Security".

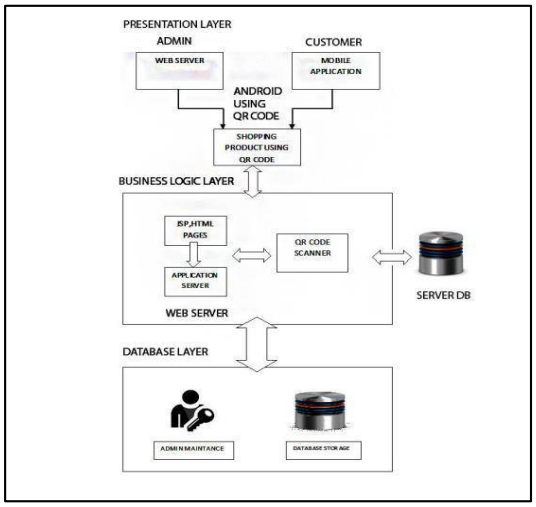


Fig.1: Working process flow diagram

Indian Journal of Natural Sciences www.tnsroindia.org.in ©IJONS

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

A Survey on Data Security Using DNA Cryptographic Techniques

K.Prem Kumar[1*] and Soundarya.S[2]

[1]Assistant Professor, Department of CSE, M. Kumarasamy College of Engineering, Karur, Tamil Nadu, India
[2]Student, Master of Engineering, Department of CSE, M. Kumarasamy College of Engineering, Karur, Tamil Nadu, India

Received: 11 Feb 2018 Revised:18Feb 2018 Accepted: 30 Mar 2018

[*]**Address for correspondence**
K.Prem Kumar
Assistant Professor, Department of CSE,
M. Kumarasamy College of Engineering, Karur,
Tamil Nadu, India
E mail: soundaryashanme@gmil.com

ABSTRACT

Cryptography plays a vital role in the field of information security. It is a study of different kinds of procedures that applies for encryption and decryption process. Cryptography has been used for various purposes especially to maintain data confidentiality. DNA Cryptography is one of major field which is used in the cryptographic techniques. The aim of the DNA Cryptography is to attain high level of confidentiality and integrity. It is implemented by using various arithmetic and biological techniques for authentication purpose.This paper focus on various schemes and algorithms in DNA Cryptography.

Keywords: Cryptography, Confidentiality, Integrity, Encryption, Decryption, Authentication, DNA Cryptography..

INTRODUCTION

Crypto is derived from Greek word kryptos which is said to be hidden information, graphen is used to write the information. Cryptography is used for conversion of plaintext to cipher text. It is very essential in the field of information and network security. Data confidentiality, integrity are the main goals of data security [2]. Nowadays, the internet plays major role in exchanging of information in the major aspects like banking, commerce transaction etc [6]. DNA Cryptography is used to hide data in the form of DNA sequence. a gram of DNA includes 1021 bases equivalent to 108 terra-bytes. DNA Encryption describes about combining the DNA technique with cryptology to give a efficient cipher services by producing different cryptography. The cryptography is classified into two they are Symmetric cryptography and Asymmetric cryptography. They are clearly made understand in means of five alphabets. They are M, C, K, E, D where, M is nothing but a message space which is also said to be plaintext space, C is a cipher text space, K is a key space. E and D is nothing but the Encryption algorithm and Decryption algorithm.

Encryption E of the plaintext message M is to be converted into cipher text C using the key value K is expressed as EKe[P]= C. Likewise for the decryption algorithm D, cipher text C is converted into plain text message M along with the key value K. They are represented as DKd[C] =P [4]. The DNA synthesis is one of the techniques which is carried out in cryptographic concept. The DNA synthesis is designing and reconstruction of information in the DNA form [7].

Biological background on DNA cryptography

DNA is Deoxyribonucleic Acid. It is s self replicating material which is present in every living organism. DNA holds genetic information to create molecules and proteins. DNA is a double helical structure called nucleotide which consists of base pair and sugar phosphate backbone. There are four nitrogen bases present over here. They are Adenine (A), Thymine (T), Guanine (G) and Cytosine (C). In common Adenine (A) and Thymine (T) form a base pair and other pair is said to be Guanine (G) and Cytosine(C) which is nothing but the pairs are said to be the hydrogen bonds. In common Adenine (A) and Thymine (T) form a base pair and other pair is said to be Guanine (G) and Cytosine (C). When Thymine (T) is replaced by Uracil (U) then that particular DNA is now RNA (Ribonucleic Acid) [1]. If all the hydrogen bonds are gets separated between the nucleotide, then two long single sided strands occur. That structure is said to be single sided DNA (ssDNA) or Oligonucleotides [5]. The structure of DNA and the DNA digital code table is shown below.

Hybridization is a process of forming a double standard DNA molecule. In the hybridization process, it is mandatory that the hydrogen bond is formed only between the Watson-Crick pairs. The length of the ssDNA must be of same length in the hybridization. If not then that leads to fragmentation [5]. The diagrammatic representation of hybridization is shown below

The amplification of single or multiple strands of DNA to the multiple copies of certain DNA is called the Polymerase Chain Reaction (PCR) [2]. During the process of anneal, the two single strand of DNA forms a double strand. Transcription is a process of conversion over DNA to mRNA. In the process of transcription the nucleotide starts creating protein. In RNA stop codons are UAG (Amber), UGA (opal), UAA (Ochre). Mutation is change in DNA, which is hereditary material of life. They are the raw material for genetic variation. Mutations are classified into two different types. They are Non-sense mutation and missense mutation. Non sense mutation used for stop codon that terminates the process of protein generation. Missense mutations are used for to generate different proteins for the different codon process [1].

Survey on DNA cryptography

Mazhar Karimi et al presented a paper on **Cryptography using DNA Nucleotides** [1]. In their paper they focus on conversion of plaintext into cipher text using new symmetric key generation model in the field of DNA cryptography. The biological framework includes transcription, replication, anneal, marking and mutation. The plaintext is given in the form of bits. They can be 0's and 1's. The first step is, the key bits are encoded with the nucleotides (A, T, G, C). Then annealing is applied which is nothing but formation of molecular pair with that each encoded bits of the previous process. The transcription is started with this step which is conversion of DNA into RNA. These are the process involved in encryption of the plain text and the decryption is just the reverse process of encryption. This algorithm shows highly secured way of communication.

Mansi Rathi et al presented their paper on **Data Security using DNA Cryptography** [3]. In their paper they focus on Symmetric Key Block Cipher Algorithm with DNA sequencing. The plain text can be given as a symbols or a character. Each block and key matrix in this algorithm is of 16 bytes. The secure data communication is done with the conversion of plaintext to ASCII by using matrix form. By using a transpose and reversing the column and the row concepts concept they are focusing more confidentiality on sending the plaintext. A special key is used for

decryption process. There they applied a DNA sequencing concept for the decryption of the plaintext. Though this technique is very complex with the mathematical calculations, the result of this algorithm is very efficient with the high level of accuracy.

Niktia S.Kolte et al, presented their views on **DNA Cryptography using Index-Based Symmetric DNA Encryption Algorithm** [4]. Here they proposed an algorithm named new index-based Symmetric DNA encryption algorithm. The method endorse here is using of DNA chromosomes as One Time Pad (OTP), also used FASTA format and MATLAB functions for the encryption and decryption process. The encoding process undergoes conversion of each character of plaintext into ASCII with their equivalent binary format then to DNA sequence which is said to be symmetric key. They used a FASTA format for the DNA sequence. In the second phase they used a MATLAB function to generate an index for each character. Each index indicates the original plain text characters. The OTP which is used for decryption. The data received will be in the form of corresponding index positions. This is one of the complex and secure algorithm for transmitting the messages from sender to receiver. It is very useful for study of applicability for DNA cryptographic technology.

Shreyas Chavan presented his paper on **DNA Cryptography Based on DNA Hybridization and One Time pad scheme** [5]. The algorithm used here is DNA cryptography using a combination of Hybridization of DNA Oligonucleotides and the generic binary one time pad technique. This algorithm focuses on two keys. One key is for encryption and the other is for decryption. The OTP is a key which is used for decryption. This paper fixes a scale value for Oligonucleotides at the sender side. In the sending phase, the creation of ssDNA sequence is done here. By using the below calculations, the sending process continues for the secure way of communication.

Length of ssDNA = Length of binary plaintext * n.
Length of binary sequence = Length of ssDNA * 2.

The generated keys, for Watson-Crick is related to all nucleotides that are corresponding to bit 1 in the plaintext of binary sequence. This is said to be the cipher ssDNA. XOR operation is done with the DNA sequence. Now send this cipher binary sequence to the receiver. The cipher binary sequence and the binary is of same length. The second phase is said to be the receiving side perform XOR key to binary key. Then convert resultant binary sequence into ssDNA grouping pairs. Here this paper goes with hybridization. Then group the sequence into 7 bits to its corresponding decimal number which is nothing but the ASCII values. Finally convert ASCII character to the original plaintext. This algorithm gives high level of security with the DNA hybridization which is highly scalable and reusable.

Noorul Hussain et al, presented their views on **A Novel DNA Computing based Encryption and Decryption Algorithm** [6]. They proposed a new algorithm called novel, secure, unique and dynamic DNA based encryption and decryption. The first phase focus on DNA computing based encoding algorithm. The sample output of the DNA encoding table is shown[6]. The DNA encoding table is generated after every pre-defined session intervals and hereinafter the DNA sequences and the assignment of alphabets to them would be different across different sessions. The second phase of the algorithm focus on DNA computing based encryption and decryption. The given plain text is divided into two halves, the encryption process carried out in the form of applying XNOR operation at the stage of initial process for both half of the data simultaneously. There are many small conversions involved here, they are like XNOR to mRNA to tRNA then right shifted by using the reverse transcription method. By applying the amino acid sequence, the resultant of the cipher text is obtained. The decryption process is also a vast process same as an encryption process. The receiver uses the clue to find the plaintext. The received cipher text which is said to be the protein sequences that are converted into tRNA using the amino acid sequence. The multiple rounds of functions are applied over here to decrypt the given text. This is one of the best encryption and decryption process. The strength of the algorithm fulfills the functional and non-functional attributes based on DNA sequence.

Bibhash Roy et al, presented their views on **Enhanced key Generation Scheme based Cryptography with DNA Logic** [7]. The proposed concept used here is key generation based idea using DNA synthesis. The strength of the key is increased with three levels, in that one key will be strongest. Initially they encrypted the plaintext with the first level of key. The second levels of keys are the private keys which are obtained using CTB's. There are two computations undergone here, the sender computation and the receiver computation. Both these computations uses array concept to encrypt the data. The sender and the receiver prepare the substituent list which is nothing but an array. The key1 is size of the array. The sender uses key2 to encrypt the data. The same key2 is used for the decryption also. This is one of the strongest key based cryptography. The decryption process is more powerful which provides high level of confidentiality.

Summary of survey

The survey can be summarized in the following table2.

CONCLUSION

In this survey paper we are discussed about many concepts in DNA cryptography and the algorithms overviewed by many researches till today. It is one of the emerging technologies in the field of cryptography. The study of different methods gives us depth knowledge about the DNA cryptography. This survey also explains about various arithmetic and biological operation used in DNA cryptography. The upcoming technologies in this technique give highly confidential way of communication.

REFERENCES

1. Mazhar Karimi, Waleej Haider, "Cryptography using DNA Nucleotides", International Journal of Computer Applications (0975-8887), June 2017.
2. Nileena Ouseph, "A Survey on Diverse DNA Cryptographic Techniques", International Journal of Innovative Research in Computer and Communication Engineering, September 2016.
3. Mansi Rathi, Shreyas Bhaskare, Tejas Kale, Niral Shah, Naveen Vaswani "Data Security Using DNA Cryptography", International Journal of Computer Science and Mobile Computing, Vol.5 Issue.10, October-2016, pg. 123-129.
4. Niktia S.Kolte, Prof. Dr. K.V. Kulhalli, Samrat C. Shinde , "DNA Cryptography using Index-Based Symmetric DNA Encryption Algorithm", International Journal of Engineering Research and Technology. ISSN 0974-3154 Volume 10, Number 1 2017.
5. Shreyas Chavan, "DNA Cryptography Based on DNA Hybridization and One Time pad scheme", International Journal of Engineering Research & Technology (IJERT) ISSN 2278-181,Volume.2 Issue 10, October 2013.
6. Noorul Hussain, UbaidurRahman, Chithralekha Balamurugan, Rajapandian Mariappan, "A Novel DNA Computing based Encryption and Decryption Algorithm", International Conference on Information and Communication Technologies (ICICT 2014). Procedia Computer Science 46 (2015) 463 – 475.
7. Bibhash Roy, Gautam Rakshit,Ritwik Chakraborty, "Enhanced key Generation Scheme based Cryptography with DNA Logic", International Journal of Information and Communication Technology.

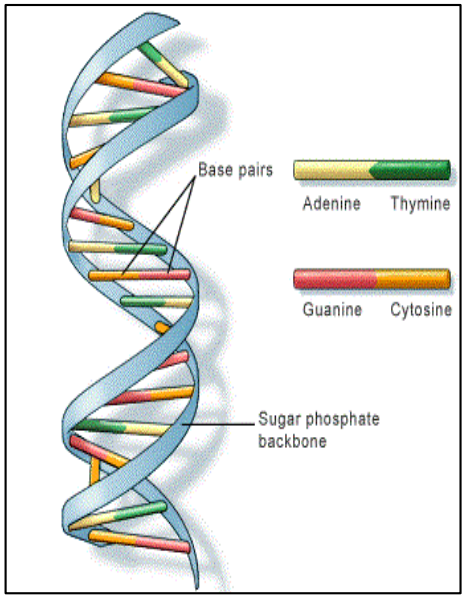


Fig.1 DNA Structure

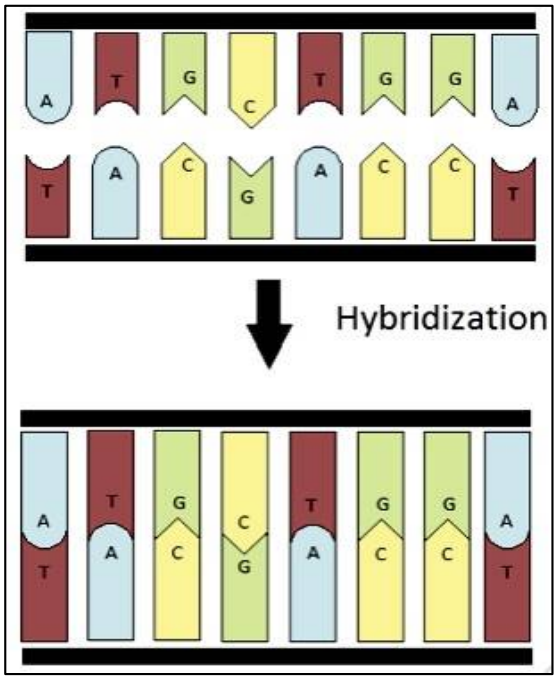


Fig.2 DNA Hybridization

Table.1 DNA Digital Code

DNA Nucleotide	Binary Sequence
A	00
C	01
G	10
T	11

Table2. Summary of all survey

Author	Title	Algorithm	Year
Mazhar Karimi, Waleej Haider [1]	Cryptography using DNA Nucleotides	New Symmetric Key Generation	2017
Mansi Rathi, Shreyas Bhaskare, Tejas Kale, Niral Shah, Naveen Vaswani [3]	Data Security Using DNA Cryptography	New Index-Based Symmetric DNA Encryption Algorithm	2016
Niktia S.Kolte, Prof. Dr. K.V. Kulhalli, Samrat C. Shinde [4]	DNA Cryptography using Index-Based Symmetric DNA Encryption Algorithm	Symmetric Key Block Cipher Algorithm	2017
Shreyas Chavan [5]	DNA Cryptography Based on DNA Hybridization and One Time pad scheme	DNA cryptography using a combination of Hybridization of DNA oligonucleotides and the generic binary one time pad technique	2013
Noorul Hussain, UbaidurRahman, Chithralekha Balamurugan, Rajapandian Mariappan	A Novel DNA Computing based Encryption and Decryption Algorithm	Novel, Secure, Unique And Dynamic DNA Based Encryption And Decryption	2015
Bibhash Roy, Gautam Rakshit, Ritwik Chakraborty	Enhanced key Generation Scheme based Cryptography with DNA Logic	Key Generation Using DNA Synthesis	2011

Indian Journal of Natural Sciences www.tnsroindia.org.in ©IJONS

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Alumni Re-Union

P.Rajesh Kanna[1*], B.Vishali[2], A.S.Sree Vidya[2], V.Priyanka[2] and T.Priyanka[2]

[1]Assistant Professor, Department of Computer Science and Engineering, M.Kumarasamy College of Engineering, Karur,TamilNadu,India.
[2]Final Year Student, Department of Computer Science and Engineering, M.Kumarasamy College of Engineering, Karur,TamilNadu,India.

Received: 11 Feb 2018 Revised:19 Feb 2018 Accepted: 30 Mar 2018

[*]**Address for correspondence**
P.Rajesh Kanna,
Assistant Professor,
Department of Computer Science and Engineering,
M.Kumarasamy College of Engineering, Karur, TamilNadu,India.
E mail:rajeshkanna@gmail.com

ABSTRACT

Many people are using mobile phones but non-communication between their school and college friends and staffs because they forget their contacts. To overcome the above issues this android application is used for the reunion of alumnus and connect alumni members regularly, to displays, share recent news and photos, offline alert, inaugurate forum, video streaming, online seminar talk etc.

Keywords: Reunion, alumnus, offline alert, video streaming, online seminar connect regularly.

INTRODUCTION

In Mobile computing, transmission occurs not through physical connectivity but mostly through satellite or wireless communication enabled networks. The interconnected components are Mobile hardware, Communication devices including transmission towers, Mobile software and apps. Mobile computing is a generic term that allows people to access and share data and information.

Mobile Communication

Mobile communication includes vast array of protocols, transmission towers, band width, satellite systems and networks. Data systems are defined so that there is no collision with existing system sending and receiving similar kinds of data. In mobile computing, transmission occurs not through physical connectivity but mostly through satellite or wireless communication enabled networks.

Alumnus

An alumnus can also be a former member, employee, contributor or inmate as well as former student. Alumni portal is also a communication, innovation and cooperation instrument for sustainable international co-operation in the fields of culture, education, research and science. That is a global network of alumni from a wide range of disciplines. Alumni portal wants to enable and inform and inspire partnerships and cooperation. Project examples, tips, hands-on activities and events show that every single person including alumni is in demand and can contribute something.

Alumni Portal Mobile App

Mobile app is an excellent information hub and communication tool. They serve up only the most relevant content and can be easily updated in real-time. A mobile app encourages alumni to stay connected and reach recent graduates and donors with mobile apps for home coming reunions, regional networking events and fundraising programs. Some of Mobile Operating Systems are Android is a Linux-based used in tablets and smart phones; IOS developed by Apple Inc used in iPod, iPod and Apple TV; Symbian used to design smart phones like Nokia and other; Mobile IP is communication protocol. Alumni Portal Mobile App runs on any Android operating system. This is simple and effective platform to connect our alumni globally. Alumni Directory app can create a location for alumni to add their contact information.

Connected Work

The foundation of any successful alumni relations program is an accurate alumni record database. Alumni associations developed programming around fostering the reciprocal relationship with their alumni and communicating value-driven messages. Networking and social events around the country allow alumni to maintain and build relationships with peers, students and alumni officers.

Our vision is to create an active alumni networking point for the world wide exchange of knowledge through the Alumni portal mobile app in order to make international cooperation efficient and sustainable. This mobile app protecting the personal data of alumni is important to us. The data entered during registration and after that remain under your own control. You determine which information is visible to who search engines and non-registered users have no access to the pages with personal data.

This online app can be accessed the organization and outside customers as well with proper login provided, which will give better service to the alumni's.

Alumni app organizes Alumni Reunions. Build Alumni Database. Engage Alumni. Raise Alumni Funds. Our alumni mobile app solution empowers alumni members and streamlines the alumni engagement activities. This app-based exclusive alumni network will help you promote a meaningful engagement with your alumni and leverage the power of alumni relationships.

Existing System

Many institutions and universities maintain the information manually about present and past students. This does not allow efficient data management and retrieval process. Most of the apps to accomplish a complex and time consuming task. It is developed to an efficient mobile user interface design to connect and collaborate former students. It developed on Android platform and it facilitates various services to connect with each other. This allows the user to post messges, displays recent news and photos, live conversation with their friends.

Disadvantages

- Only an option of sending text messages to the friends.
- No video streaming.
- No offline alert.
- Alumnus can't post any seminar talks and other useful social information.
- No Dynamic Grouping.

Proposed System

In this proposed system, the app allows students to register and then search the database on different criteria. Then included video streaming, online seminar talk, offline alert, event management, auto archive gallery, membership card management and dynamic email group and user permissions for the alumni members in the institution to get connected and communicated easily.

Advantages

- The main advantage is to find friends and also to be contact with old friends.
- News regarding alumni meet can be easily meet can be easily known since this system Conway such information to the passed students once posted.
- This can post the various information regarding job vacancies, competitive exams and so on. So one's circle becomes big.

METHODOLOGIES

Admin

The admin will helps to register user. After registration to add the result of the search engines. Manages annual and gift activities to support strategic plan. Implements and monitors comprehensive resource development program, makes recommendations to improve department policies and procedures. Serves as liaison to alumni groups to foster activities and maintain strong relationships. Works closely with alumni in planning activities and programs. Maintain communications with alumni, encouraging long-term relations.

Student

Most communication between alumni and students is done through e-mail. Networking is beneficial to all students, as you explore, define and finally pursue your career interests. The students know the information about upcoming events, any cultural programs, sports and about the examination details. Also having recently conducted event images, videos.

Alumini

Alumni are very important part of universities creating the best possible experience for students. Alumni meet can be easily known since this system conveys such information to the passed students once posted an alumnus can post any information, so one's circle becomes big.

CONCLUSION

The analysis work Alumni Re-union using Android Application is used for the reunion of alumnus and also to meet their old friends contact and communicate regularly. This app is one such attempt to bring the passed out students of

the college closer to their junior, helps to connect and get information and guidance for their career such communications are video streaming, online seminar talk. It has been developed using java and SQLite server compact, the complete system is thoroughly tested with the availability of data and throughput reports which are prepared manually.

REFERENCES

1. M.A.Norasiah and A. Norhayati, A Survey on "SMART CONNECT" an Android and Web Based Application for College Management System, Vol 8, Issue 3, Pub year 2003.
2. D.C.Hillman, D.J.Willis, &C.N.Gunawardena, Alumni Portal Mobile App. Vol 7 Issue 2, Pub year 2005.
3. Prof. SeemaVanjire, UnmeshKanchan, Java Project On Alumni Information Database, Vol 6, Issue 3, Pub year 2014.
4. Jyothi T.S., Catherine Mathew, Irene George, Android Based College Notification System, Vol 16, Issue 2, Pub year 2013.
5. Mahesh Kadibagil, Dr.H.S.Guruprasad, College Phone Book Android Application, Vol 4, Issue 3, Pub year 2014.
6. DeepaV.Jose, Lakshmi Priya C.G. Priyadars Job Portal Android Application, Vol 5, Issue 1, Pub year 2015.
7. KomwitSurachat, SupasitKajkamhaengAlumni Tracking Using Google Map Api And Social Media Based On Gps AndLbs, Vol 6, Issue 5, Pub year 2010.
8. ZigurdMednieks, Laird Domin, Eliz-Instant Messanger For Android Platform, Vol 5, Issue 3, Pub year 2009.
9. T.Govindasamy, An Innovation System That Can Quickly Responces To The Needs Of Students And Alumni, Vol 3, Issue 4, Pub year 2008.
10. Mahfudzah Othman, Zainab Othman, Contacts And Image Uploading Android Project Vol 4, Issue 5, Pub year 2007.

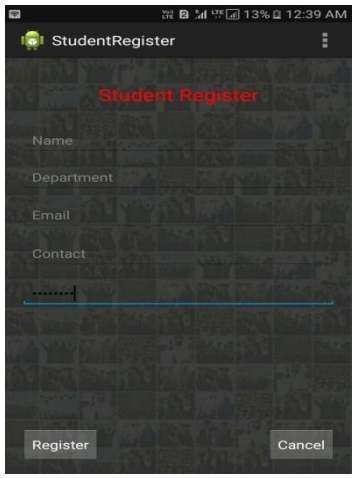


Fig.1.Student Register

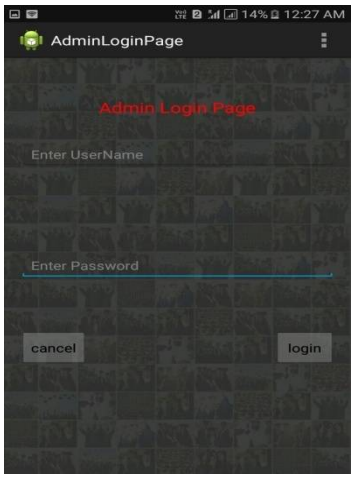


Fig.2.Admin Login Page

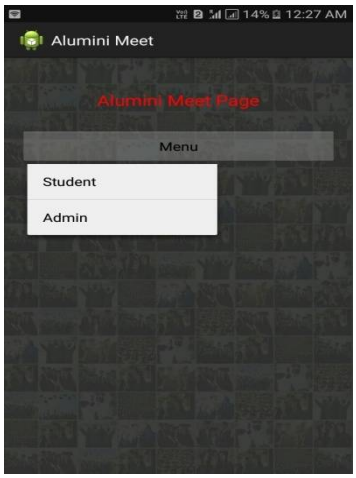


Fig.3.Alumni Meet Page

Indian Journal of Natural Sciences www.tnsroindia.org.in ©IJONS

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Analyzing Water Quality under IoT Environment

S.Saravanan* and Ranjitha K

Department of Computer Science and Engineering, M.Kumarasamay College of Engineering, Karur, Tamilnadu, India.

Received: 10 Feb 2018 Revised:17 Feb 2018 Accepted: 30 Mar 2018

***Address for correspondence**
S.Saravanan
Department of Computer Science and Engineering,
M.Kumarasamay College of Engineering,
Karur, Tamilnadu, India.
E mail: Saravanans.cse@mkce.ac.in

ABSTRACT

Now a day water is getting dirtied by a few ways and checking it's quality is a biggest arrangement for globalization. This issue makes to screen persistently the nature of the water. With the goal that water Parameter like its temperature, PH, turbidity and level of the water ought to be checked at which rate of water gets flushed out from industry. By physically checking the water quality is exceptionally tedious and in addition to get water Parameter continuously. To beat this testing gear are to be set in the repository, with the goal that the varieties happens in these Parameter will discover the nearness of toxins. In this paper another approach is utilized to investigate the nature of the water under IOT condition. The proposed framework contains information acquisitions hub, controlling units as ARM, base station and checking unit, this units are associated with each other. Information can be gathered from different sensor and every sensor has its own specific manner to peruse its information about physical marvel. So as to control and to do some preprocessing work MCU units has been utilized after information accumulation from different sensor with the end goal that temperature ,PH, turbidity, and so on at base station is send to show on observing hub. By persistently observing water quality individuals gets mindfulness about water condition so we decrease the impact on human life, recorded, and so forth.

Keywords: IOT , Temperature sensor , level sensor , turbidity sensor.

INTRODUCTION

At present situation we all are facing the global warming due to the pollution which may be a water pollution or environment pollution. These issues can be avoided by protecting environment. Meanwhile water pollution is the major task in nowadays. Water gets polluted due to industrial wastage or by human or due to some uncontrollable disaster. This paper presents different technologies to analyze the quality of water. Wireless sensor networks are

widely used in various kinds of application to sense and collect data about physical things for better understanding. By developing new design method for multisensory data obtaining interface device for plug and play for different types of sensor. So that new design adopted MCU controller has advantages of low cost and easy implement with high performance.

LITERATURE REVIEW

As a happened innovation realized fast advances in late remote media transmission, Internet of Things (IoT) has connected with a ton of consideration and furthermore anticipated that would bring favorable position to endless application regions including modern WSN frameworks and medicinal services frameworks manufacturing[1]. Sensor interface apparatus is fundamental for recognize different sorts of sensor information of modern WSN in IoT conditions. It empowers us to obtain sensor information. Along these lines, we can successfully comprehend the outside condition information[2]. With brisk advancement of IoT, real producers are focused on the examination of multi sensor obtaining interface apparatus[3]. There are a great deal of measurements procurement a few interface contraption with trustworthy advancements available. Be that as it may, these interface apparatuses are extremely spent significant time in working style, so they are not independently flexible to the changing IoT environment[4].Wireless sensor systems (WSN) are appropriate for long haul natural insights procurement for IoT portrayal. This paper introduces the useful outline and usage of an entire WSN stage that can be utilized for a scope of long haul natural observing IoT applications. The application prerequisites for minimal effort, high number of sensors, quick arrangement, long lifetime, low support, and high caliber of administration are considered in the particular and plan of the stage and of every one of its segments. Low-exertion stage reuse is additionally viewed as beginning from the determinations and at all plan levels for a wide cluster of related checking applications[5]. Remote sensor systems (WSNs) and radio-frequency identification(RFID) frameworks give a great foundation to information obtaining, dissemination, and preparing. In this paper, some key difficulties identified with the incorporation of WSN and RFID advancements are discussed[6]. A five-layer framework design has been proposed to accomplish synergistic execution. For the mix of WSN and RFID, one of the basic issues is the low effectiveness of correspondence because of repetitive information as excess information expands vitality utilization and causes time delay[7]. To address it, an enhanced information cleaning calculation has been proposed; its plausibility and viability have been confirmed by means of reproduction and a correlation with a distributed algorithm[8].

METHODOLOGY

The designed system was adopted two hardware topologies intelligent sensor module like IMS and remote data acquisition unit for data storing, acquisition and for transmission. Dynamic power management which scheduled switching mode used to switch a sensor node to sleep mode after transmitting a data packet. The presents of five layer system architecture to integrated WSN and RFID technologies for interactive action for manufacturing atmosphere. They adopted data cleaning algorithm to prevent redundant data which causes due to duplicate reading and due to time delay between each sensor nodes. Data fusion framework comprises of hypothesis, methods and algorithm. Data collection is the essential applicationof WSN and more importantly it is the foundation of otheradvanced applications in IoT environment. The tool for data manipulating, managing to increase the efficiency of processing as well as to achieve advanced intelligence in IOT.

Proposed System

For current scenarios monitoring the polluted water resources is essential need to avoid major effect on human life, filed, industrial application to mitigate and monitor critical situation from contaminated water we propose a new design which is includes interface device to collect data effectively from diverse sensor and MCU for controlling the entire operation and for some preprocessing Fig1.

RESULTS AND DISCUSSION

Presently a day's water get dirtied by numerous ways as being human, debacle or might be expected industry so dissecting the nature of sullied water is pivotal part for modern application, home, condition application and so forth. By checking the nature of the water assets is to maintain a strategic distance from real impact on both human lives and in addition on recorded, which causes dieses through defiled water. With the goal that individuals need to get mindfulness region water condition. The water contamination can be controlled by persistently checking the water Parameter like its temperature, PH, turbidity and level of the water ought to be checked at which rate of water gets flushed out from industry. So we can decrease the impact on both person and recorded.

Performance Analysis

Thus the performance analysis has been conducted and adequate effects is carried out on water reservoir monitoring for industrial application by detecting the value from temperature sensor, level sensor turbidity sensor, we can minimize the hypercriticalstages for industrial water reservoir monitoring for industrial application using IoT .Water gets polluted in many ways like industrial wastages, due to disaster or human being.

The above graph represents the effect due to industrial wastage on water resource fig 2. Before the industry gets started the water resource was pure and environment was healthy. After the industry gets started, the wastage in industry was added in water resource and made them impure.

This graph represents the purity of water gets polluted day by day and increases the impurity of water fig 3. Because of adding the wastage into the water resource and also effect the living begins. After the industry gets started the purity of water reduced and impurity of water gets raises.

CONCLUSION

Water gets dirtied by numerous ways as being human, debacle or might be expected mechanical waste. So examining the nature of debased water is urgent part for mechanical application, home, condition application and so on. By checking the nature of the water assets to dodge real impact on both human lives and also on documented and furthermore which causes infections through polluted water. With the goal that individuals need to get mindfulness environs water condition. To settle this basic issues the framework has been proposed for water repository checking utilizing IoT under this framework MCU is utilized as a controller to do some particular and in addition preprocessing undertaking, information procurement should be possible in parallel route by identifying the for temperatures sensor esteem, turbidity sensor and level sensor esteems. Along these lines increment the nature of water and lessen its terrible effects.

REFERENCES

1. Qing ping Chi, Hairong Yan, Chuan Zhang, Zhibo Pang, and Li Da Xu, “A Reconfigurable Smart Sensor Interface for Industrial WSN in IoT Environment,” IEEE Transactions on Industrial Informatics, vol. 10, no. 2,pp.1417-1425, May 2014
2. L. Wang, L. D. Xu, Z. Bi, and Y. Xu, “Data cleaning for RFID and WSN integration,” IEEE Transactions on Industrial Informatics, vol. 10, no. 1, pp. 408-418, February 2014.
3. M. T. Lazarescu, “Design of a WSN platform for long-term environmental monitoring for IoTapplications,”IEEE Journal On Emerging And Selected Topics In Circuits And Systems, Vol. 3, No. 1, pp. 45-55,March 2013.
4. F.Salvadori et.al. “Monitoring in industrial system using wireless sensor network with dynamic power management.” IEEE Transactions on Instrumentation and Measurement, vol. 58, no. 9, pp. 3104-3111, September 2009.
5. Y. Fan, Y. Yin, L. Xu, Y. Zeng, and F. Wu, “IoT based smart rehabilitation system,” IEEE Trans. Ind. Informat., vol. 10, no. 2, pp. 1568–1577, 2014.

6. S Saravanan, V Venkatachalam ,“ Improving map reduce task scheduling and micro-partitioning mechanism for mobile cloud multimedia services” International Journal of Advanced Intelligence Paradigms ,Vol 8(2),pp157-167,2016.
7. S.Saravanan, V Venkatachalam ,“ Advance Map Reduce Task Scheduling algorithm using mobile cloud multimedia services architecture” IEEE Digital Explore,pp21-25,2014.

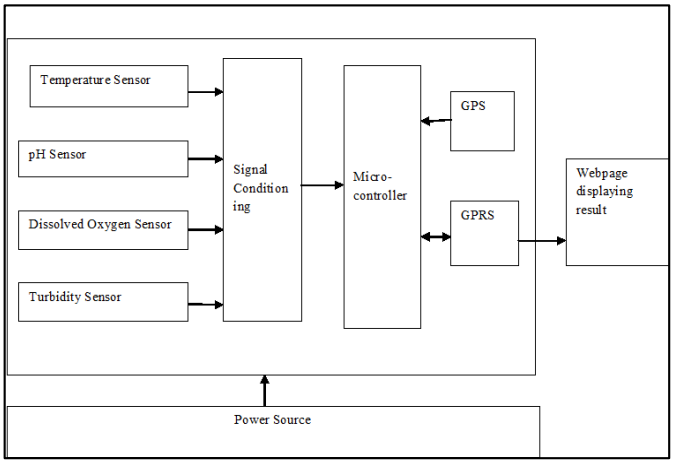


Fig.1.Block Diagram of Water Monitoring Hardware System

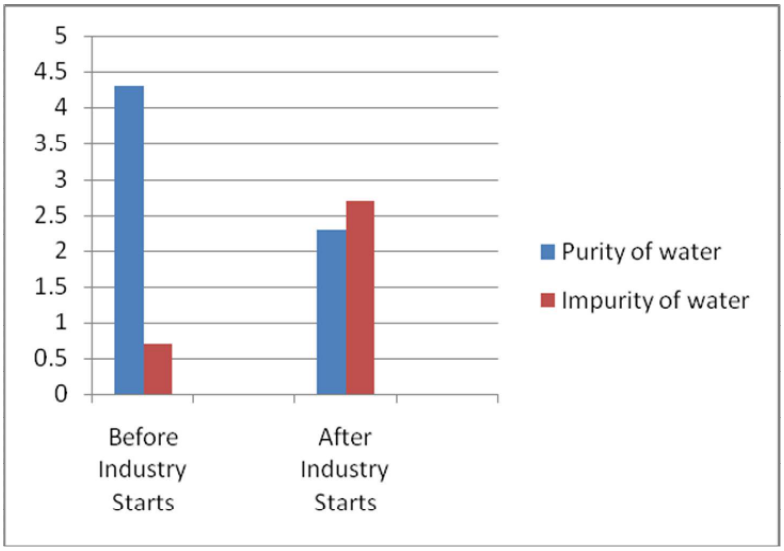


Fig.2.Effects Due To Industrial Wastages

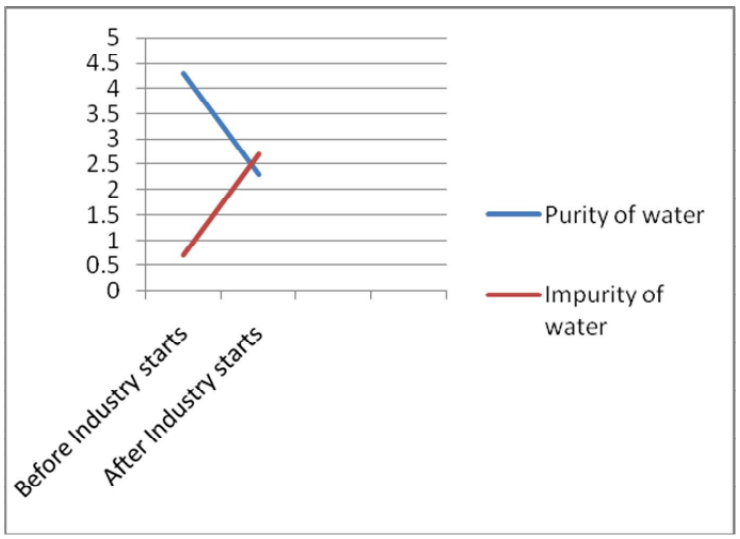


Fig.3.Graphical Representation of Increasing Impurities in Water

Table:1. Difference between purity and impurity of water

Starts	Purity of Water	Impurity of Water
Before Industry Starts	4.3	0.7
After Industry Starts	2.3	2.7

Indian Journal of Natural Sciences *www.tnsroindia.org.in ©IJONS*

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Improving Association Rule Mining by Using Data Structure

R.Bharathi[1]* and Ranjitha K[2]

[1]Assistant Professor, Department of Computer Science and Engineering, M.Kumarasamay College of Engineering,Karur, Tamilnadu,India.
[2]P.G. Student, Department of Computer Science and Engineering, M.Kumarasamay College of Engineering, Karur,Tamilnadu,India.

Received:10 Feb 2018 Revised:18 Feb 2018 Accepted: 30 Mar 2018

***Address for correspondence**
R.Bharathi
Assistant Professor,
Department of Computer Science and Engineering,
M.Kumarasamay College of Engineering,
Karur, Tamilnadu.India.
E mail: Bharathir.cse@mkce.ac.in / ranjikumarasamy@gmail.com

ABSTRACT

In now a days, the storage of data in different fields are made in digital format and it becomes easy to use and the storage media also available at lower coast. The data that we use are highly relating to measurements and miscellaneous in nature. The knowledge discovery process is hard due to large data set with highly relating to measurements and miscellaneous data. This type of process can be condensed as association rule mining (ARM). In nowadays many algorithm has been put forward for association rule mining and the mining process will be difficult when we assign huge amount of data. When the size of data is extremely huge in terms of documents. Therefore the aim is to create a new data structure for storing the data in accurate way but not to create a new algorithm for mining. The actual data is understandable, identifiable and the entry time is enlarged for that data, to encounter the effectiveness such as time and storage space are required in main memory. By using the functions such as shuffling , inverted index mapping and run length encoding, the storage in primary memory access time will be less and data usage are achieved. As a result the existing association rule mining algorithms are accessed along with data structure for quick process of mining and decreasing the requirement made in primary memory, does not modifies the actual algorithms. Thus it can further upgraded by replacing Run Length Encoding by Modified Run Length Encoding Algorithm to get superior memory access and effectiveness to mining process.

Keywords: Association Rule Mining(ARM), Knowledge Discovery, Data Compression, Index Compression, Data Structure.

INTRODUCTION

In early days, all the fields such as technical and non-technical fields are made as digitization to increased the production of digital data in larger amount by every day[1]. It is efficient to store the large data since the cost has been reduced when compared to early storage medias. Hence the storage cost plays an insignificant association with larger amount and heterogeneity of the data. But the process of mining is affected tremendously. The process of mining has become more interesting but it takes long time, because of vast amount of data. Association rule mining[2] is well known and data is collected by mining rules from researched field. Data analysis are done by introducing many algorithms to increase the speed.

Disparate schemes are used to decrease the number of items sets and transaction or both the number of compared data sets. Yet the process of mining with huge dataset will have very highly relating to measurements data[1] will be a difficult process. Thus alternatively without modifying the current time algorithms otherwise creating a new algorithm for quick accesses of mining process, it will be good to establish a new data structure[3] for storing larger amount of data in compressed format. Here the new compressed data structure is applied and that data are stored in compressed format of actual data in data structure without any change in actual content. The three sequential techniques namely Shuffling, Inverted index mapping and Run length encoding are applied on actual data to obtain this new compressed data structure.

In further by using Modified Run Length Encoding we can improved the data structure. When compared to any of the existing mining algorithm, this new data structure helps in managing dataset, quick access to data and high speed of computing.

LITERATURE REVIEW

Association rule mining [2] is used as the major techniques for designing the mining rules. The main aim is the get the input data set as strong and correlative items. Mainly this mining rule helps to improve the business and this rule was designed for shopkeepers to sort the sale items through market basket analysis.

Apriori[4] is built for association rule mining and is the first algorithm. It is mostly depends on complete search. It operators with two condition such as I. all the frequent item sets are found from data set and II. Those frequent item sets are used for deriving the association rules. The [2^k-1] item sets can be created using the data set of k single items and N number of transactions. The processing will be hard due to huge size of data set, because of the difficulty in above working will have O(N X M X K) impartial to acquire the item sets.

The frequent pattern FP-growth[5] algorithm was initiated for decreasing the number of transactions and related differentiation. Information which is stored should scanned at least once because FP-growth will store the frequent items sets in a tree structure. This type of storage is required because it suffers with huge size if I/O and also with huge size of requirement for memory to store all data sets. To overpower the existing pitfall, like large requirement of memory space and it takes long times to response, many evolutionary association rule mining algorithms were proposed [6],[7],[8],[9]. Even though after viewing dissimilar data sets in mining rules and it is also hard to work with huge dimensional data.

As a result, the advantageous is found by creating a new data structure so that it decreases the size of original data set and it helps to quickly access the mining, and uses the current algorithms of original data set without any changes in its functioning.

In this paper data structure is introduced in order to decrease the size of the data set and also to improve the data access speed. It gives assurance that the data will be valid and also will keep the original data set values without any change. The data can be converted into new data structure which follows three steps such as 1.Shuffling, 2. Inverted Index Mapping and 3.Data Compression. The Original Run Length Encoding algorithm input data can be more compressed with help of [10] Modified Run Length techniques and it finally helps to improve the performance of ARMs

System Architecture / System Overview

Thus the conclusive goal of work is to create a new data structure in order to decrease size of data set and to improve the operation speed.As shown in the above figure the original data helps in creating the new data structure for performing three function on actual data such as shuffling, inverted index mapping and run length encoding. Each functions will follow this work.

Shuffling

The characteristic of specific data set has a difference between documentations that is two or more files allows same feature values. The document can be sorted among data by using this characteristics with the help of Hamming Distance(HD). HD is used in different metric in sequences to reduce number of modification in the feature values from document to document. Compression is made in sorted data by merging the characteristic of subsequent documents. To the lower of hamming distance between two documents which implies higher chance in clustering into same group. For this Function original data set will be the input and which chance the sorted data after the completion of this function.

Inverted Index Mapping

In this input will be taken as sorted data and its aim is to built the efficient structure such as index based structure. Index has the attribute values formed on self satisfied transaction and then consecutive transaction are used to join indexed attribute values to share similar values. Each list of attributes has key value pair which is assigned by inverted index methodology and transaction index are satisfied through each attribute which has a pointer to access the attribute value.

RLE encoding

This is the major step for data compression without any chance to original data and quick access of data. In this the input is takes as inverted index structure and attributes are group as sequence of index into 2-records that will explain the being index and all transaction as number of sequence that carries the similar feature values{index, displacement}.

Modified RLE

If the compression ratio for any type of data can mostly improved by modified run length encoding scheme. First the input data is analysed and is the fundamental step. Analysis the data to high point and if any huge numbers of consecutive then it will enlarge the number of bits for representing each run length. This process has input data that holds close value with its neighbouring data and have same data for both values. In this result compressed data does not lose any of the actual data and forms data structure.

System Analysis

Dataset Information: In this we use dissimilar datasets like Milk and Egg containing various attributes including instances.

Parameter features

Size of data: The compressed data structure is represented by the proposed data structure based on the survey of zip which is one the universal compressors that produces the result based on the corresponding file size.

Load Time: The original data is now compressed and there is a variation in the loading time where the loading time for actual data is higher than the loading time for compressed data.

Memory Size: Only 53 KB of memory is required for analyzing milk data set out of 465 KB. Hence, the compressed information ultimately requires less memory space after the data is processed.

RESULTS

By using shuffling, mapping and RLE and then Modified RLE compression techniques for compressing the data sets of Milk and Egg. After applying Apriori algorithm on compressed datasets only the result is obtained. The graph is plotted in order to compare data size and time required to store and process respectively for both original and compressed data structures by using the observations and results.

From the table and graph it is noted that, the compressed data set decrease the size and time required to store and process respectively.

CONCLUSION

As there is a vast increase in the data sets the use of Association Rule Mining is more fascinating and user-friendly. Where the mining process will take huge amount of time such data using traditional mining algorithms. Hence the performance of existing algorithms can improved by data compression techniques without changing their original data set schema. The mining algorithms uses the above compressed data from data compression techniques and it increases the speed based on size of data, load time and memory size. In future, the mining process can be improved by using FP-Growth.

REFERENCES

[1]V. Marx, "Biology: The big challenges of big data," Nature, vol. 498, no. 7453, pp. 255–260, 2013.
[2]R. Agrawal and R. Srikant, "Fast algorithms for mining association rules in large databases," in Proc. 20th Int. Conf. Very Large Data Bases (VLDB), Santiago, Chile, 1994, pp. 487–499.
[3]José María Luna, Alberto Cano, Mykola Pechenizkiy, " Speeding-Up Association Rule Mining With Inverted Index Compression", IEEE TRANSACTIONS ON CYBERNETICS, 2016.
[4]R. Agrawal, H. Mannila, R. Srikant, H. Toivonen, and A. I. Verkamo, "Fast discovery of association rules," in Advances in Knowledge Discovery and Data Mining. Menlo Park, CA, USA: Amer. Assoc. Artif. Intell., 1996, pp. 307–328.
[5]J. Han, J. Pei, Y. Yin, and R. Mao, "Mining frequent patterns without candidate generation: A frequent-pattern tree approach," Data Min. Knowl. Disc., vol. 8, no. 1, pp. 53–87, 2004.

[6]B. Goethals and M. J. Zaki, "Advances in frequent itemset mining implementations: Report on FIMI'03," ACM SIGKDD Explor. Newslett., vol. 6, no. 1, pp. 109–117, 2004.

[7]C. Lucchese, S. Orlando, and R. Perego, "DCI closed: A fast and memory efficient algorithm to mine frequent closed itemsets," in Proc. IEEE ICDM Workshop Frequent Itemset Min. Implement. (FIMI), Brighton, U.K., 2004, pp. 20–28.

[8]J. M. Luna, J. R. Romero, C. Romero, and S. Ventura, "On the use of genetic programming for mining comprehensible rules in subgroup discovery," IEEE Trans. Cybern., vol. 44, no. 12, pp. 2329–2341, Dec. 2014.

[9]C.Borgelt,"Efficient Impleme -ntations of Apriori and Eclat," in Proc. 1st IEEE ICDM Workshop Frequent Item Set Min. Implement. (FIMI), Melbourne, FL, USA, 2003, pp. 90–99.

[10] S. Joseph, N. Srikanth, J. E. N. Abhilash, "A Novel Approach of Modified Run Length Encoding Scheme for High Speed Data Communication Application ," International Journal of Science and Research Volume, 2 Issue 12, December.

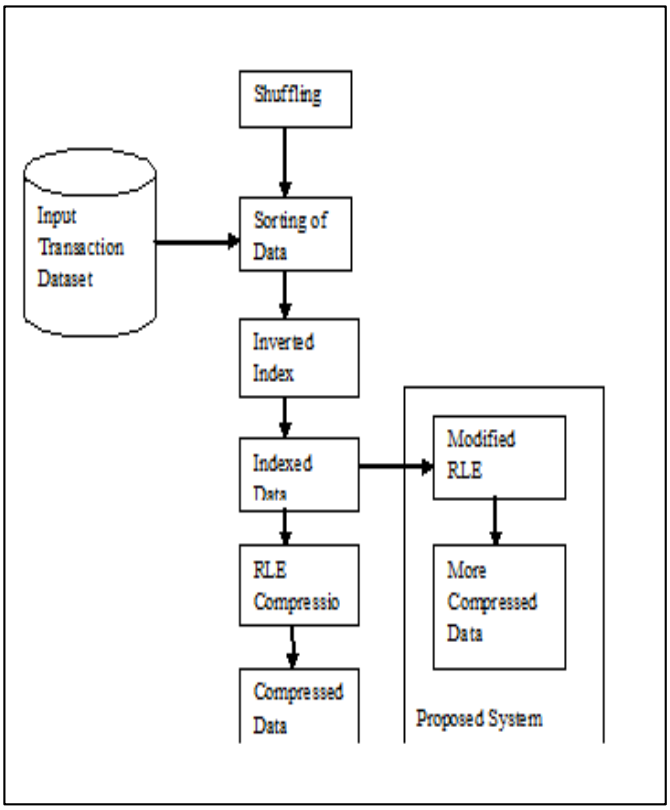
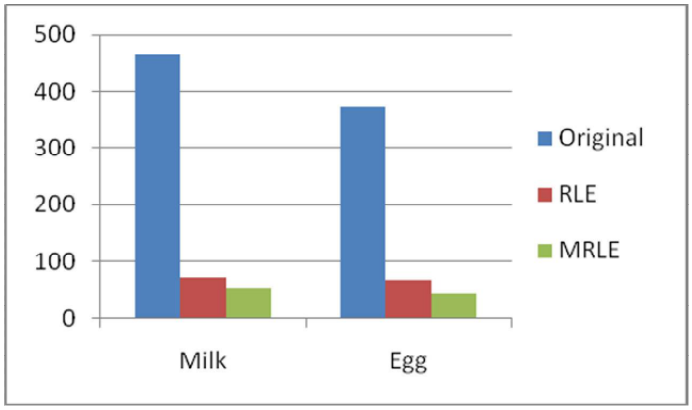
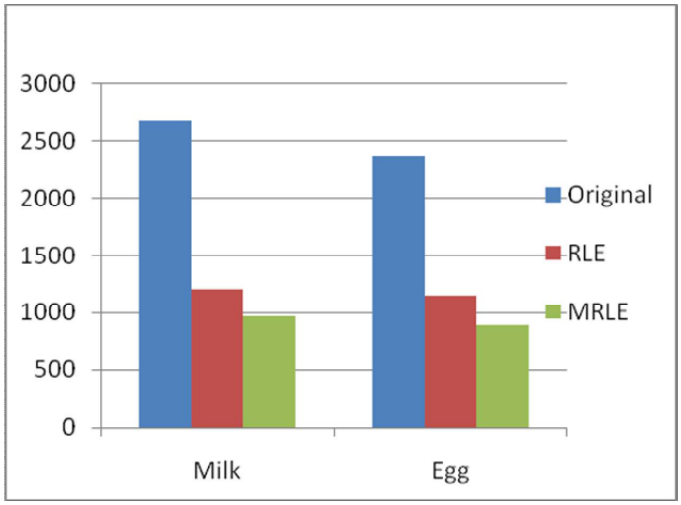


Fig.1: System Architecture / System Overview



Graph: 1



Graph: 2

Table -1: Data Compression

Data Structure	Milk	Egg
Original	465	372
RLE	72	68
MRLE	53	46

Table-2: Time required processing Apriori Algorithm

Data set	Original	RLE	MRLE
Milk	2663	1204	964
Egg	2358	1146	893

Indian Journal of Natural Sciences www.tnsroindia.org.in ©IJONS

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Efficient BWT and SVM Based Multilevel Image Segmentation for Brain Tumor Detection in Mrimages

P.Santhi[1]*, V.Nivetha[2], R.Ponmozhi[2] and L.Ruban[2]

[1] Associate Professor Department of CSE,M.Kumarasamy College of Engineering,Karur,TamilNadu,India
[2]Final Year UG Student, Department of CSE,M.Kumarasamy College of Engineering, Karur, TamilNadu

Received: 05 Feb 2018 Revised: 25 Feb 2018 Accepted: 30 Mar 2018

***Address for correspondence**
P.Santhi, Associate Professor,
Department of CSE,
M.Kumarasamy College of Engineering,Karur, TamilNadu,India
E mail: santhip.cse@mkce.ac.in

ABSTRACT

Picture division using Magnetic Resonance pictures gives striking information and has an indispensable impact in recognizing the different sorts of tumor. Diverse learning techniques have been utilized for tumor acknowledgment by taking a gander at removed incorporate reasons for the photo under examination and reference picture. It is a trying task to develop and assemble a trustworthy data for cerebrum tumor acknowledgment by means of getting ready on account of tremendous assortments of mind picture perfectly healthy and control. The division, acknowledgment of the affected region, and feature extraction of impacted tumor area from MR pictures are a fundamental concern however a troublesome and time taking errand performed by radiologists or clinical experts, and their precision depends upon their experience so to speak. Thusly, the use of PC helped development ends up being uncommonly vital to crush these limitations. In our wander, Berkeley wavelet change (BWT) is used to upgrade the execution of tumor distinguishing proof and decline the inconveniences show in the therapeutic picture division process. Also, moreover improve the exactness and quality distinguishing proof of the assistance vector machine (SVM) based classifier; noteworthy features are removed from each divided tissue.

Keywords: Magnetic Resonance, Diverse learning techniques, Berkeley wavelet change.

INTRODUCTION

Appealing Resonance Imaging (MRI) is comprehensively supported for analyzingthe structure of psyche and diagnosing cerebrum related contaminations. Not with standing the way that enlisted tomography (CT) uncovers the anatomical information by applying X-bars, the patients may be affected by radioactivity. In MRI, it is possible to isolate preferable tissue separate contemplated over CT without the risk of radiation. In light of MRI system, white issue and diminish issue are unmistakably perceived in thinks about while analyzing mind structures and cerebrum

related information. White issue and dull issue are the major cerebral tissues as a main priority structures. MR pictures outfit awesome spatial assurance with less affectability for perceiving low riches particles. The correct information extraction of cerebrum structure isn't possible in MRI pictures in view of its weakness to quantify and confine mind tumor tissues definitely. The above issues drive the need for MRI picture division in MRI based finding. Barely any case pictures with irregular personality tissue are showed up in figure 1.

Solid automated picture division methodology are fundamental for MRI pictures since the manual division techniques exhausts some portion of time and the division comes to fruition differed from one ace to other ace. In electronic division approaches, the specialist can get more information related to the size, territory and energy of cerebrum tissues and tumor. Regardless of the way that modernized division techniques are important from numerous perspectives at the top of the priority list tumor assurance, it is a troublesome task to secure exact information in light of various shapes, foggy points of confinement, nearby cerebrum tissues and unmistakable zones. A couple of MRI division approaches are open in the writing in perspective of edge, edge acknowledgment, watershed, gathering, histogram, graph cut and real getting ready. The wide collection of division approaches are required on account of the diserse quality related with the sound personality tissues, tumor and edema.[3].

LITERATURE REVIEW

Different picture division approaches are open in composing for isolating basic features in test pictures, remedial pictures and satellite pictures. Thresholding based approach is a fundamental sort of disconnecting picture into different division in light of certain edge estimation of diminish level. Wang et al (2011) showed a multilevel thresholding to isolate the photo into different classes to research the distinctive pieces in a photo. The decision of breaking point regards are urgent and requires some entropy based bundle computations, for instance, cushy c-portion counts. Avci et al. presented fuzzy3-partitioncriterion where the perfect decision of edge is changed over into entropy maximization of soft interest work. Kalra& Kumar showed fuzzy4-section theory that surveyed the joining among establishment and articles to pick perfect farthest point an impetus for division. Edges characterize boundaries and are therefore a problem of fundamental importance in image processing. Image Edge detection significantly reduces the amount of data and filters out useless information, while preserving the important structural properties in an image.[17]. Driven primarily by the widespread availability of various small-animal models of human diseases replicating accurately biological and biochemical processes.[2].[5].

GC is a standard chart based division for recognizing mind tumor in MR pictures where edges are addressed as centers to find the closeness. Boykov& Jolly showed a GC estimation for PET and MRI picture division where the contracting issue was not portrayed properly.[1]. Feathery c-suggests gathering (FCM) gets a larger number of information from the given picture than other hard batching approaches, however the FCM without considering the spatial information is more fragile to hullabaloo. partition based Fuzzy C-Means gathering (KFCM) count vanquishes the issues related with FCM by mapping data into higher dimensional Hilbert space to upgrade the batching methodology.[10]. Cluster analysis is unsupervised learning method that constitutes a cornerstone of an intelligent data analysis process. It is useful for the exploration of inter-relationships among a collection of patterns, by organizing into homogeneous clusters.[6]. The detection of edema is done simultaneously with tumor segmentation, as the knowledge of the extent of edema is important for diagnosis, planning, and treatment. Whereas many other tumor segmentation methods rely on the intensity enhancement produced by the gadolinium contrast agent in the T1-weighted image, the method proposed here does not require contrast enhanced image channels.[16]

System Architecture

Picture division is one of the indispensable methodologies for acknowledgment of tumor from the MR pictures or from other restorative imaging modalities pictures for picking right remedial guide at the ideal time. In proposed

structure new and beneficial tumor area is realizing with the help of BWT and SVM figurings. The extraction of the tumor needs the division of the mind pictures into two zones. One area contains the tumor cells of the psyche and moreover the second region contains the common cerebrum cells. To construct the execution level and decreases the inconveniences show inside the helpful picture division system, we have investigated Berkeley Wavelet Transformation (BWT) based tumor division. The image is changed to binary to ease the computing process and maximize the speed of calculating due to the rapid Boolean operators. By changing to binary, the femur shaft can be separated from the soft tissue shade which can be considered noise during the bone shaft image processing[11]. Support Vector Machine (SVM) classifier is used to remove the related features from each partitioned tissue, to improve the precision and quality rate tumor recognizable proof, the assistance vector machine. The proposed methodology performs division, feature extraction, and request as is done in human vision wisdom, that perceives exceptionally shocking things, altogether sudden surfaces, contrast, sparkle, and significance of the photo. The problem of segmenting the object from the background is addressed in the proposed Gaussian and Gabor Filter Approach (GGFA) for object segmentation. An improved and efficient approach based on Gaussian and Gabor Filter reads the given input image and performs filtering and smoothing operation[14]. Several screening approaches are used now to detect suspicious lung cancer lesions. Computed tomography (CT) is especially sensitive to hard-to-detect nodules and enhances radiologists' diagnosis accuracy. X-ray is often thought as an obsolete medical imaging method, but usage of digital technologies and machine learning now revives the significance of X-ray in medical imaging diagnosis[8].

SVM Classification

The SVM computation relies upon the examination of a regulated learning technique and is associated with one-class gathering issue to n-class portrayal issues. The rule purpose of the SVM computation is to change a nonlinear segregating focus into an immediate change using a limit called SVM's bit work. In this examination, we used the Gaussian piece work for change. By using a bit work, the nonlinear cases can be changed into a high-dimensional future space where the parcel of nonlinear cases or data may wind up possible, making the portrayal invaluable. The SVM count portrays a hyper plane that is divided into two instructional courses as described in underneath condition

$$f(y) = Z^T \phi\ (y) + b$$

RESULT ANALYSIS

The ampleness of our proposed system is viewed as and the obtained results are discussed around there. Test pictures are obtained from both the online web database and constant checked MR pictures. Progressing MR image securing is performed by Siemens 48 channel MRI furnished with Magnetom Avanto-Tim headways. At first stage, the recorded MR check picture is preprocessed using picture histogram and thresholding for artifact clearing, trailed by fundamental center filtering to remove some other fuss. For standard MRI database pictures, the antiquated irregularity removal step isn't essential. After removal of noise from input picture division was performed. Affectability, specificity and division precision are processed for the database pictures and ceaseless inspected pictures. By then SVM arrange is used to recognize the tumor absolutely. To show the sufficiency of the proposed procedure, estimation or execution time and required number of cycles are figured and differentiated and practically identical grouping based approaches.

CONCLUSION

In our endeavor use a skull stripping computation in light of constrain procedure to improve the skull stripping execution. Berkeley wavelet change is used to section the photos and support vector machine to arrange the tumor

orchestrate by analyzing feature vectors and area of the tumor. In our proposed structure surface and histogram based features are amassed and seen for the gathering of cerebrum tumor from MR mind pictures. From the testing comes to fruition performed on the particular pictures, clearly the examination for the mind tumor acknowledgment is fast and correct when differentiated and the manual revelation performed by radiologists or clinical experts.

REFERENCES

[1] S. Leibfarth, F. Eckert, S. Welz, C. Siegel, H. Schmidt, N. Schwenzer,D. Zips, D. Thorwarth, Automatic delineation of tumor volumes by co segmentation of combined PET/MR data, Physics in Medicine and Biology,60(14) (2015) 5399-5412.

[2] H. Zaidi, Molecular imaging of small animals, New York, NY, USA: Springer, (2014).

[3] Sikka, K., Sinha, N., Singh, P. K., & Mishra, A. K. (2009). A fully automated algorithm under modified FCM framework for improved brain MR image segmentation. Magnetic Resonance Imaging, 27(7), 994-1004.

[4] J. J. Corso, E. Sharon, S. Dube, S. El-Saden, U. Sinha, and A. Yuille, "Efficient multilevel brain tumor segmentation with integrated Bayesian model classification," IEEE Trans. Med. Imag., vol. 27, no. 5, pp. 629-640, May 2008.

[5] J. R. Jiménez-Alaniz, V. Medina-Bañuelos, and O. Yáñez-Suárez, "Data-driven brain MRI segmentation supported on edge confidence and a priori tissue information," IEEE Trans. Med. Imag., vol. 25, no. 1, pp. 74-83, Jan 2006.

[6] S.Thilagamani, N.Shanthi," Literature survey on enhancing cluster quality", International Journal On Computer Science and Engineering, Vol.2, pp. 1-10. No.6, 2010.

[7] K. M. Iftekharuddin, J. Zheng, M. A. Islam, R. J. Ogg, "Fractal-based brain tumor detection in multimodal MRI," Applied Mathematics and Computation, vol. 207, pp. 23–41, 2009.

[8] Dr.P.Santhi, S.Kiruthika," Lung Based Disease prediction Using Lobe Segementation Based on Neural Networks" International Journal of Pure and Applied Mathematics,Vol.118,pp.2-10 No.8,2018.

[9] S Thilagamani, N Shanthi," A novel recursive clustering algorithm for image oversegmentation", European Journal of Scientific Research,Vol.52, no.3, 2011.

[10] W. Dou, S. Ruan, Y. Chen, D. Bloyet, J.-M. Constans, "A framework of fuzzy information fusion for the segmentation of brain tumor tissues on MR images," Image and Vision Computing, vol. 25, pp. 164–171, 2007.

[11]P.Santhi,R.Vikram, "Implementation of classification System Using Density Clustering Based Cray Level Co-Occurrence Matrix (DGLCM) for Green Bio Technology," International Journal of Pure and Applied Mathematics Vol.118,pp. 3-5.No.8,2018.

[12]S.Thilagamani, N.Shanthi," A survey on image segmentation through clustering", International Journal of Research and Reviews in Information Sciences,Vol.1,Issue 1,2011.

[13] N. Zhang, S. Ruan, S. Lebonvallet, Q. Liao, Y. Zhu, "Kernel feature selection to fuse multi-spectral MRI images for brain tumor segmentation," Computer Vision and Image Understanding, vol. 115, pp. 256–269, 2011.

[14] S.Thilagamani, N.Shanthi," Gaussian and Gabor Filter Approach for Object Segmentation" Journal of Computing and Information Science in Engineering, Vol.14, pp.10-11 No.2, 2013.

[15] H. A. Vrooman, C. A. Cocosco, F. Lijn, R. Stokking, M. A. Ikram, M. W. Vernooij, M. M. B. Breteler, and W. J. Niessen, "Multi-spectral brain tissue segmentation using automatically trained k-nearest-neighbor classification," NeuroImage, vol. 37, pp. 71–81, 2007.

[16] M. Prastawa, E. Bullitt, S. Ho, G. Gerig, "A brain tumor segmentation framework based on outlier detection", Medical Image Analysis, vol. 8, pp. 275–283, Jul 2004.

[17]Dr.S.Thilagamani, V.Manochitra, "An Intelligent Region-Based Method for Detecting Objects from Natural Images", International Journal of Pure and Applied Mathematics,Vol.118,pp 1-15 .

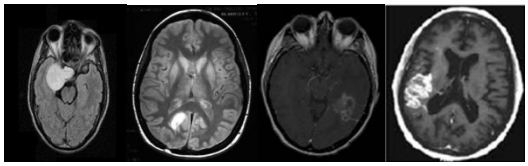


Fig 1: Abnormal Brain Images

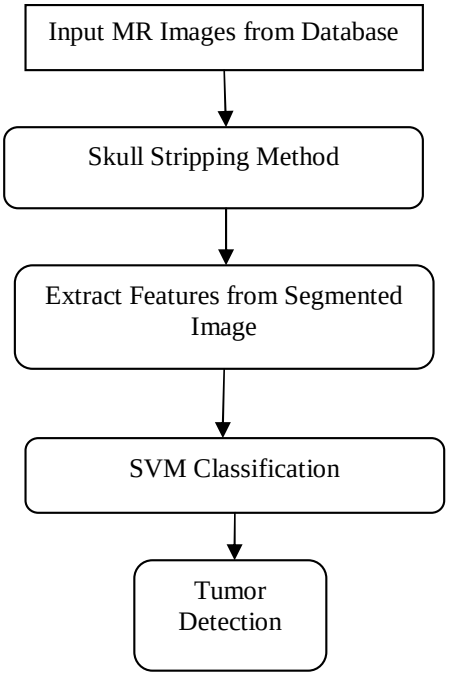


Fig.2 - Architecture

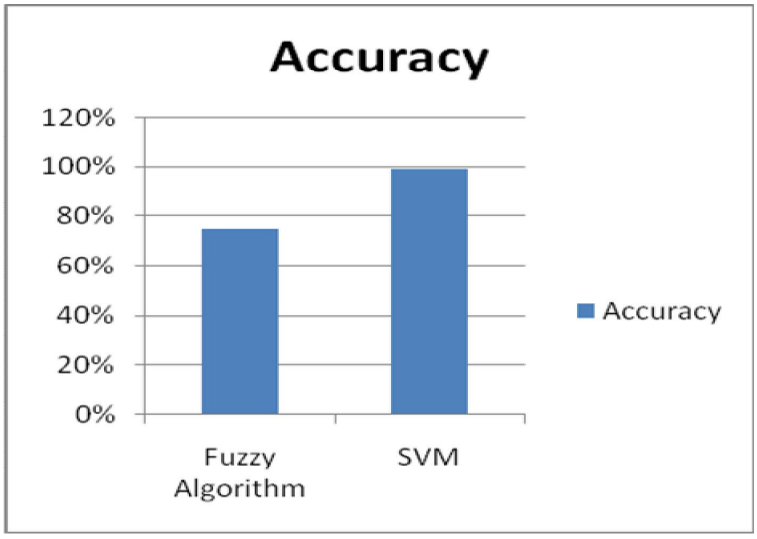


Fig.3- Accuracy (Fuzzy Vs SVM)

Indian Journal of Natural Sciences www.tnsroindia.org.in ©IJONS

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Cloud Based Privacy Preserving System for Removing Tantamount Data

S.Santhiya[1]*, K.Vishalini[2], K.Tharani[2], B.Praveena[2] and R.Swathi[2]

[1]Assistant Professor, Department of CSE,M.Kumarasamy College of Engineering, Karur, TamilNadu
[2]Student, Department of CSE, M.Kumarasamy College of Engineering, Karur,TamilNadu,India.

Received: 12 Feb 2018 Revised:16 Feb 2018 Accepted: 30 Mar 2018

*Address for correspondence
S.Santhiya
Assistant Professor, Department of CSE,
M.Kumarasamy College of Engineering, Karur, TamilNadu,India.
E mail: santhiyas.cse@mkce.ac.in

ABSTRACT

With the rapidly growing amounts of facts produced global, networked and multi-consumer storage systems have become very famous. However, worries over information safety still prevent many customers from migrating facts to far flung garage. The conventional solution is to encrypt the information earlier than it leaves the owner's premises. While sound from a safety angle, this method prevents the garage issuer from effectively making use of storage efficiency capabilities, which includes compression and deduplication, which could permit best utilization of the resources and consequently lower carrier fee. Client-aspect data deduplication particularly ensures that more than one uploads of the equal content handiest devour community bandwidth and garage area of a unmarried upload. Deduplication is actively used by a number of cloud backup providers as well as various cloud services. Unfortunately, encrypted facts are pseudorandom and as a consequence cannot be deduplicated: therefore, cutting-edge schemes need to completely sacrifice both security and garage performance. In this paper, we present schemes that permit a greater quality-grained change-off in records chunk similarity. The instinct is that outsourced records may additionally require exceptional degrees of safety, relying on how popular it's miles: content material shared through many users. Various deduplication schemes are analyze and provide experimental outcomes that suggests proposed cozy facts bite similarity provide improved effects in real time cloud environments.

Keywords: Data chunks, Similarity matching, Parallel processing, Data security, Data compression

INTRODUCTION

Now a day there is boom in records. With infinite storage area offer by means of cloud carrier provider customers generally tend to apply as much area as they are able to and vendors constantly search for techniques aimed to limit redundant records and maximize area financial savings. Users will get admission to statistics in step with their desires and maximum customer's access identical facts time and again, the fee of computation, software hosting, content material storage and transport is reduced drastically. The cloud makes it possible to be able to get admission to your information from anywhere at any time. Cloud provides advantages consisting of, flexibility, disastracter,

restoration, software program updates routinely, pay- according to-use version and cost discount.[3] While a regular PC setup calls for you to be inside the indistinguishable area as your data stockpiling device, the cloud makes away that stride. The cloud disposes of the need with an end goal to be inside the indistinguishable substantial area as the equipment that shops your records. Each issuer serves a particular characteristic, giving customers extra or much less manipulate over their cloud depending on the type. Your cloud wishes will change depending on how you intend to utilize the hole and resources related with the cloud. Distributed computing alludes to the utilization of PCs which get to Internet areas for registering force, stockpiling and projects, without a need for the character get right of section to elements to hold any of the framework. Information deduplication is a procedure for decreasing the measure of carport space an association needs to shop its insights. In many associations, the carport structures incorporate generation duplicates of numerous bits of insights. For instance, the indistinguishable document can be put away in various particular areas by method for selective clients, or more noteworthy archives that are not same may likewise still comprise of a dreadful parcel of similar records. Along with low ownership prices and flexibility, users require the protection in their statistics and confidentiality ensures via encryption. To make statistics control scalable deduplication we're use Encryption for relaxed deduplication offerings.[1] Unfortunately, deduplication and encryption are conflicting technology. While the aim of deduplication is to locate identical records segments and shop them best as soon as, the end result of encryption is to make two same records segments in distinguishable after being encrypted. This manner that if information is encrypted through customers in a fashionable way as like shared authority, the cloud storage company cannot follow deduplication for the reason that identical data segments could be one of a kind after encryption. On the other hand, if statistics aren't encrypted by customers, confidentiality through can't be guaranteed and records aren't blanketed against curious cloud garage providers. There are styles of deduplication as far as the scale: (I) document organize deduplication, which finds redundancies among elite records and evacuates these redundancies to diminish ability needs, and (ii) block level deduplication, which finds and disposes of redundancies between records squares. The archive can be partitioned into littler consistent length or variable-estimate pieces. Utilizing fixed size squares improves the calculations of piece constraints, while the utilization of variable-length squares. An approach which has been proposed to meet those two conflicting necessities is Tag era and AES Scheme wherein the encryption secret is commonly the end result of the hash of the facts segment. Although encryption seems to be an amazing candidate to attain confidentiality and deduplication at the equal time, it sadly suffers from numerous well-known weaknesses. The confidentiality issue can be treated through encrypting touchy information earlier than outsourcing to remote servers. Along with low ownership charges and adaptability, clients require the wellbeing in their actualities and classification guarantees through encryption. In this paper, we adapt to the expressed constrainment issue to propose a common specialist to the documents which Deduplicted principally based protection saving confirmation for the cloud insights carport, which acknowledges verification and approval without bargaining a man's non-open records. The basic data chunk similarity is shown in fig 1

Related Work

Wang,et.al,... [1] Proposed a dynamic open cloud use show for little to medium scale logical organizations to influence utilization of versatile sources on an open cloud web to webpage on the web while keeping their adaptable device controls, i.e., make, enact, hunch, continue, deactivate, and demolish their over the top confirmation control elements—supplier control layers without understanding the data of control. Second, we layout and implement an progressive device—Dawning Cloud, at the middle of which might be mild-weight carrier manipulate layers strolling on top of a commonplace control provider framework. The typical oversee supplier system of Dawning Cloud never again fine permits building light-weight backer oversee layers for heterogeneous workloads, however also makes their control obligations simple. Third, we look at the frameworks exhaustively the utilization of both copying and genuine tests.

Li,et.al,... [2] Took a bit nearer to bringing the different endowments of the MapReduce model to incremental one-pass examination. In the fresh out of the plastic new model, the MapReduce framework peruses input data most straightforward when, plays incremental handling as additional certainties is perused, and makes utilization of

framework sources viably to accomplish intemperate general execution and adaptability The objective is to format a stage to help such versatile, incremental one-sidestep examination. This stage might be utilized to help intelligent actualities investigation, which may likewise contain online accumulation with early surmised arrangements, and, inside the future, development inquiry handling, which offers near constant bits of knowledge as new information arrives We contend that, with the goal that it will bolster incremental one-skip examination, a MapReduce framework must avoid any closing off activities and also computational and I/O bottlenecks that keep measurements from "effortlessly" streaming through guide and lessen stages at the preparing pipeline.

R. Kienzler,et.al,... [3] Propose an incremental actualities get right of section to and handling method for insights escalated cloud programs that can camouflage measurements switch latencies while saving straight versatility. Comparative in soul to pipelined question assessment in traditional database frameworks, data is gotten to and prepared in little additions, in this manner engendering insights pieces from one level of the measurements investigation dare to another as fast as they're accessible as opposed to holding up till the entire dataset transforms into accessible. This way we can process information typically in memory (therefore, decrease tedious I/O to adjacent plate and cloud carport, and avoid stockpiling charges) and additionally achieving pipelined parallelism (further to the current apportioned parallelism), primary to a lessening in normal mission last touch time.

C. Olston,et.al,...[4] Proposed a device for Building and updating a search index from a movement of crawled net pages. Some of the numerous steps are deduplication, link analysis for unsolicited mail and exceptional classification, joining with click on-based totally recognition measurements, and file inversion. Processing semi-structured records feeds, e.g. Information and (micro-)blogs. Steps consist of de-duplication, geographic region decision, and named entity reputation. Processing alongside these strains is an increasing number of completed on a new technology of bendy and scalable facts management systems, inclusive of Pig/Hadoop. Hadoop is a scalable, fault-tolerant machine for strolling character map-reduce processing operations over unstructured information documents. Pig adds higher-stage, based abstractions for statistics and processing. Despite the success of Pig/Hadoop, it's far becoming apparent that a new, higher, layer is wanted: a workow supervisor that offers with a graph of interconnected Pig Latin applications, with statistics handed among them in a non-stop style. Given that Pig itself offers with graphs of interconnected data processing steps, it's far natural to ask why one would layer any other graph abstraction on pinnacle of Pig.

K.H. Lee, et.al,... [5] Carried out The programming rendition is invigorated by the guide and diminishes natives found in Lisp and other viable dialects. Before building up the MapReduce system, Google utilized masses of isolated usage to strategy and register huge datasets. A large portion of the calculations were fairly simple, yet the information data transformed into frequently exceptionally gigantic. Consequently the calculations should have been circulated all through heaps of PCs to finish counts in a moderate time. MapReduce is detectably proficient and versatile, and thus might be utilized to framework huge datasets. At the point when the MapReduce structure moved toward becoming conveyed, Google completely changed its net inquiry ordering contraption to apply the new programming model. The ordering contraption delivers the records structures utilized by Google web look. The parallelization doesn't really need to be finished over numerous machines in a system. There are unprecedented executions of MapReduce for parallelizing figuring in particular situations. Phoenix is a usage of MapReduce, which is pointed toward shared-memory, multi-center and multiprocessor frameworks, i.e. Unmarried PC frameworks with numerous processor centers.

Data Deduplication Types

In this chapter, we can discuss the various data deduplication types as follows

Document level de-duplication

It is typically called single-case carport, document level data de-duplication looks at a record that must be chronicled or reinforcement that has just been spared by method for checking every one of its traits against the list. The list is refreshed and spared just if the document is specific, if not than best a pointer to the current record this is put away references. Just the unmarried case of record is spared inside the final product and material duplicates are changed through "stub" which focuses to the remarkable record.

Square level de-duplication

Square stage measurements deduplication works based on sub-record level. As the call infers that the record is being broken into fragments pieces or lumps with a specific end goal to be tried for already put away information versus repetition. The mainstream strategy to choose excess certainties is by appointing identifier to bite of insights, by the utilization of hash calculation for instance it creates a one of a kind ID to that novel square. The exact particular Id can be in correlation with the basic file. On the off chance that the ID is as of now present, at that point it speaks to that before best the actualities is prepared and put away sooner than .Due to this exclusive a pointer reference is put away inside the region of beforehand put away information. On the off chance that the ID is new and does never again exist, at that point that piece is specific. Subsequent to putting away the one of a kind bite the extraordinary ID is refreshed into the Index. There is change in size of bite as predictable with the vender. Some will have settled square sizes, in the meantime as a couple of others utilize variable piece sizes

Variable square level de-duplication

It looks at different sizes of information obstructs that can diminish the odds of crash, expressed Data joins Orlandini. The difference between deduplication schemes are shown in fig 2.

Secure Deduplication Algorithms

The main objective of this paper to analyze various encryption algorithms with deduplication schemes. The basic algorithms are shown as follows:

Traditional Encryption algorithm

Although it is recognized that statistics deduplication gives extra blessings, protection and confinementconcerns get up due to the fact the customers touchy records is vulnerable to both the outsider in addition to insider assaults. So, whilst thinking about the traditional encryption techniques to secure the customers touchy statistics there are many issues are related. Traditional encryption offers records confidentiality but it is not well suited with deduplication. As in regular encryption elite clients scramble their information with their own keys. Along these lines, similar information of the uncommon clients will prompt unmistakable figure message that is making the data deduplication almost outlandish on this conventional strategy. The fundamental advance of the calculation as shows:

KeyGenSE: k is the key age calculation that produces κ utilizing security parameter I

EncSE (k, M): C is the symmetric encryption calculation that takes the mystery κ and message M and after that yields the ciphertext C;

DecSE (k, C): M is the symmetric decoding calculation that takes the mystery κ and ciphertext C and afterward yields the first message M.

Merged Encryption calculation

The joined encryption systems are those which give the information secrecy to the clients outsourced information put away on general society mists.These techniques while providing the confidentiality to the data are also compatible with the data deduplication process. In this algorithm the encryption key is itself derived from the message. So it supports data deduplication also, because the same file will give the same encryption key so it will generate the same cipher text irrespective of users which makes data deduplication possible.

KeyGenCE(M) → K is the key age calculation that maps an information duplicate M to a merged key K;
EncCE(K,M) → C is the symmetric encryption calculation that takes both the merged key K and the information duplicate M as sources of info and after that yields a ciphertext C;
DecCE(K,C) → M is the decoding calculation that takes both the ciphertext C and the merged key K as sources of info and after that yields the first information duplicate M; and
TagGen (M) → T (M) is the label age calculation that maps the first information duplicate M and yields a label T (M).

Block cipher algorithm

In cryptography, a square figure is a deterministic course of action of signs walking around standard length associations of bits, suggested as pieces, with an unvarying change that is particular through strategy for a symmetric key. Square figures execute as key smooth included substances in the course of action of various cryptographic traditions, and are fundamentally used to region into affect encryption of mass records. Iterated thing figures finish encryption in a couple of changes, everything about usages a greatly saw as one in each stand-out sub key got from the certified key. One colossal execution of such figures is profoundly completed inside the DES figure. Various particular recognize of square figures, all in all with the AES, are set apart as substitution-change frameworks. The guide of the DES consider balanced along with essential in the all inclusive community records of bleeding edge piece figure outline. It other than facilitated the instructional difference in cryptanalytic strikes. Both differential and direct cryptanalysis developed out of research on the DES outline. There is a palette of strike techniques inside the course of which a piece figure should be secure, despite being strong closer to creature strain ambushes. Without a doubt, even a quiet piece figure is proper mind boggling for the encryption of an unmarried square under an outrageous and brisk key. Countless of action were planned to allow their reiterated use safy, consistently to get the protection yearns for mystery and authenticity. Regardless, piece figures can similarly function as creating squares especially cryptographic traditions, which merge noble hash aptitudes and pseudo-unpredictable combination turbines.

One key sort of iterated piece figure known as a substitution-change organize (SPN) takes a square of the plaintext and the indispensable detail as wellsprings of information, and applies different pivoting rounds which join a substitution degree found through a phase degree—to offer each square of figure abstract substance yield. The non-straight substitution degree mixes the urgent thing bits with those of the plaintext, building up Shannon's confuse. The immediate stage affirmation by then scatters redundancies, making scattering. One key sort of iterated piece figure known as a substitution-change orchestrate (SPN) takes a square of the plaintext and the crucial detail as wellsprings of information, and applies different turning rounds which join a substitution degree found through a phase degree—to offer each square of figure artistic substance yield. The non-straight substitution degree mixes the significant thing bits with those of the plaintext, building up Shannon's confuse. The immediate stage affirmation by then scatters redundancies, making scattering.

Variable chunk similarity

It calls for greater processing energy than the file deduplication, seeing that the amount of identifiers that want to be processed increases considerably. Correspondingly, its index for tracking the person iterations gets additionally a whole lot massive. Using of variable period blocks is even more source-big. Moreover, once in a while the equal hash variety may be generated for 2 exquisite records fragments that are referred to as hash collisions. If that occurs, the system will now not keep the contemporary statistics as it sees that the hash wide variety already exists inside the index. The algorithm steps as follows BlockTag(FileBlock) - It registers hash of the File obstruct as document square Tag;

DupCheckReq(Token) - It asks for the Storage Server for Duplicate Check of the document piece. FileUploadReq(FileBlockID, FileBlock, Token) – It transfers the File Data to the Storage Server if the document piece is Unique and updates the record square Token put away.

FileBlockEncrypt(Fileblock) - It scrambles the document hinder with Convergent Encryption, where the concurrent key is from SHA Hashing of the record square; TokenGen(File Block, UserID) – the procedure stacks the related benefit keys of the client and create token. FileBlockStore(FileBlockID, FileBlock, Token) - It stores the FileBlock on Disk and updates the Mapping. The variable lump closeness level deduplication is appeared in fig 3.

The mathematical model as follows

Give S a chance to be the framework protest.
It comprise of following S={U,F,CSP} U= no of clients U={u1,u2,u3,… ..un} F= no of documents
F={f1,f2,f3,… ..fn} B=no of pieces. B{B1,B2,… ,Bn} CSP={C,PF,V,POW} C=challenge
PF =proof by CSP
V= confirmation by TPA
POW= verification of proprietorship CSP= Cloud Service supplier CSP={PF,F} PF=proof
F=files
The proposed architecture is shown in fig 4

EXPERIMENTAL RESULTS

The proposed algorithm is analyzed in terms of storage preserving and implemented in real time environments. The proposed result is shown in fig 4.The proposed system preserves the storage up to 80% storage with security.

CONCLUSION

In this paper investigated dispensed deduplication frameworks to upgrade the unwavering quality of records while accomplishing the privacy of the clients and also shared expert outsourced records with an encryption system. The structures have been proposed to help report-degree and piece organize information deduplication. The security of label consistency and trustworthiness has been done. We completed our deduplication frameworks the utilization of the square figure plot with variable nibble likeness and tried that it brings about little encoding/interpreting overhead when contrasted with the system transmission overhead in common include/down load tasks. Information secrecy is proficient for the reason that wrapped esteems are traded all through transmission. Client imprisonment is prevalent with the guide of motivate passage to solicitations to secretly inform the cloud server regarding the client's entrance dreams.

REFERENCES

1 L. Wang, J. Zhan, W. Shi and Y. Liang, "In cloud, can scientific communities benefit from the economies of scale?" IEEE Transactions on Parallel and Distributed Systems 23(2): 296-303, 2012.

2 Dr.S.Thilagamani, Dr.P.Santhi,"A Survey on Audit Free Cloud Storage Via Deniable Attribute Based Encryption" IRA-International Journal of Technology & Engineering: Vol.05,No.1 2016

3 B. Li, E. Mazur, Y. Diao, A. McGregor and P. Shenoy, "A platform for scalable one-pass analytics using mapreduce," in: Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD'11), 2011, pp. 985-996.

4 R. Kienzler, R. Bruggmann, A. Ranganathan and N.Tatbul, "Stream as you go: The case for incremental data access and processing in the cloud," IEEE ICDE International Workshop on Data Management in the Cloud (DMC'12), 2012

5 K.Sangeetha, S.Santhiya,"Hospital Record Search using RDF based Information Retrieval",International Journal of Software Engineering and its Application,Vol.10,No.8,2016.

6 C. Olston, G. Chiou, L. Chitnis, F. Liu, Y. Han, M.Larsson, A. Neumann, V.B.N. Rao, V. Sankarasubramanian, S. Seth, C. Tian, T. ZiCornell and X.Wang, "Nova: Continuous pig/hadoop workflows,"Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD'11), pp. 1081-1090, 2011.

7 S.Santhiya, K.Sangeetha,"Image based classification scheme to mine the opinion in social network for exactness",Journal of Chemical and Pharmaceutial Sciences, ISSN:0974-2115,2017.

8 K.H. Lee, Y.J. Lee, H. Choi, Y.D. Chung and B. Moon,"Parallel data processing with mapreduce: A survey," ACM SIGMOD Record 40(4): 11-20, 2012.

9 C.Selvarathi,K.Sindhanaiselvan,"Object Segmentation from Movable frames using key segments",Journal of Chemical and Pharmaceutical Sciences.ISSN:0974-2115,2017.

10 X. Zhang, C. Liu, S. Nepal and J. Chen, "An Efficient Quasiidentifier Index based Approach for Privacy Preservation over Incremental Data Sets on Cloud," Journal of Computer and System Sciences (JCSS), 79(5): 542-555, 2013.

11 X. Zhang, T. Yang, C. Liu and J. Chen, "A Scalable Two-Phase Top-Down Specialization Approach for Data Anonymization using Systems, in MapReduce on Cloud," IEEE Transactions on Parallel and Distributed, 25(2): 363-373, 2014.

12 S.Santhiya, S.Sarvanan, "Intrusion Detection & Prevention System in cloud using Decision Tree Model",International Journal of Advanced Research in Computer and Communication Engineering,Vol.06,No.1,2017.

13 N. Laptev, K. Zeng and C. Zaniolo, "Very fast estimation for result and accuracy of big data analytics: The EARL system," Proceedings of the 29th IEEE International Conference on Data Engineering (ICDE), pp. 1296-1299, 2013.

14 T. Condie, P. Mineiro, N. Polyzotis and M. Weimer,"Machine learning on Big Data," Proceedings of the 29th IEEE International Conference on Data Engineering (ICDE), pp. 1242-1244, 2013.

15 Aboulnaga and S. Babu, "Workload management for Big Data analytics," Proceedings of the 29th IEEE International Conference on Data Engineering (ICDE), pp. 1249, 2013.

16 M. Abadi, D. Boneh, I. Mironov, A. Raghunathan, and G.Segev,"Message-locked encryption for lock-dependent messages," in Advancesin Cryptology –

17 CRYPTO 2013 - 33rd Annual Cryptology Conference,Santa Barbara, CA, USA, August 18-22, 2013. Proceedings, PartI, 2013, pp. 374–391.

18 T. Jiang, X. Chen, Q. Wu, J. Ma, W. Susilo, and W. Lou, "Secure andefficient cloud data Deduplication with randomized tag," IEEE Trans.Information Forensics and Security, vol. PP, no. 99

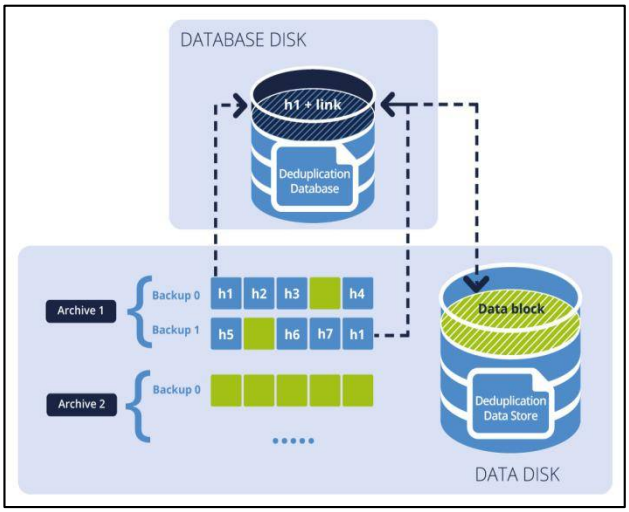


Fig 1: Data chunk similarity

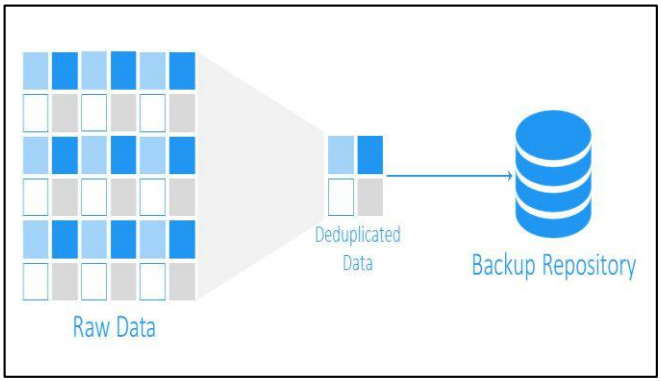


Fig 2: Deduplication schemes

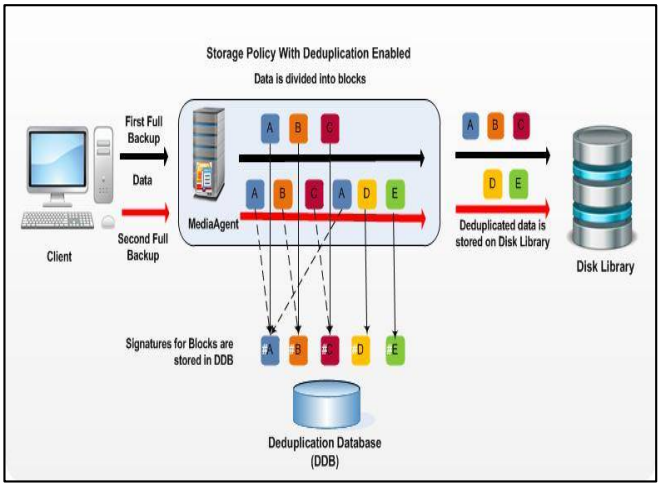


Fig 3: Variable chunk similarity backup server

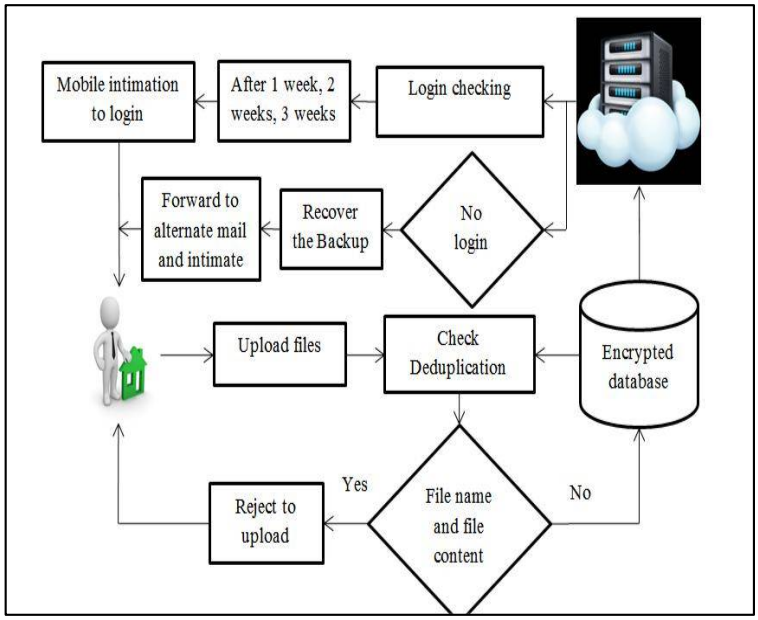


Fig 4: Proposed framework

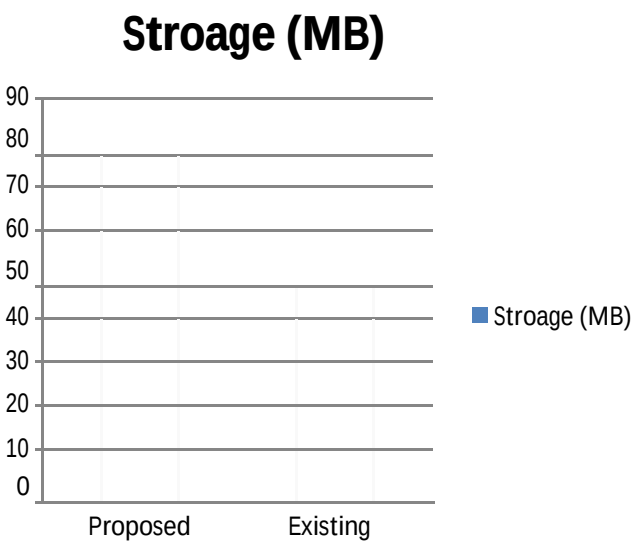


Fig 5: Experimental Results

Indian Journal of Natural Sciences *www.tnsroindia.org.in ©IJONS*

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Defect Finding Using Honeypot in Wireless Networks

S.Prabavathi[1]*, P.Mythili[2], P.Nandhini[2], K.Indira[2] and G.Naresh[2]

[1]Assistant Professor, Department of CSE,M.Kumarasamy College of Engineering, Karur, TamilNadu,
[2]Final year student, Department of CSE,M.Kumarasamy College of Engineering, Karur,TamilNadu,India

Received: 11 Feb 2018 Revised:17 Feb 2018 Accepted: 30 Mar 2018

***Address for correspondence**
S.Prabavathi
Assistant Professor,Department of CSE,
M.Kumarasamy College of Engineering,
Karur,TamilNadu,India

ABSTRACT

There is some requirement for security against digital assaults as essentially any remote association gadgets might be powerless against the malevolent hacking endeavors. A honeypot is a PC framework which is having a set up to go about as an imitation to draw in digital assailants, and furthermore to discover the examination an endeavor which is utilized to increase unapproved access to the data frameworks. Here we can break down the issue of protecting against assaults, that is available in honeypot-empowered systems. In misleading the association of an aggressor and furthermore a protector is available. The aggressor dependably attempt to cheat the preventer by doing some various types of assaults which are running from an untrutsable to an apparently ordinary movement, here the preventer thus can utilize honeypots as an apparatus, to discover assailants. These sort of issue is displayed as a Bayesian round of untrusted information's. Where evening out are distinguished here for both the various shot amusement and also the rehashed diversion adaptations. Here Our outcomes demonstrate that there is an untrusted recurrence of dynamic aggressors is available, or more which the two players can take the beguiling activities and afterward underneath the preventer can stir up his or her system utilized for keeping the assailant's prosperity rate low. Keywords: cyber-attacks, defender, deception.
Keywords: cyber-attacks, defender, deception.

INTRODUCTION

A honeypot as characterized a data framework asset whose esteem exists in unapproved access to utilization of that resource.A honeypot is a host that is unprotected and furthermore accessible to the assailants. Consider a honeynet is an arrangement of interconnected honeypots. These frameworks are observed with a specific end goal to learn new assaults, devices and a few strategies used to obtain entrance. Honeynets are helpful for examining the vast scale assaults like worms, infections and additionally botnets and furthermore have a little esteem utilized for breaking down particular assaults focused to a particular host To enhancing the recurrence of the Internet assaults some sort of research gatherings and associations were created and furthermore sent in honeynet frameworks. Honey nets is

the utilization of a trickery, a traditional counter insight strategy which is as yet utilized broadly by knowledge associations. In light of the level of the interchanges between the assailant and checked framework, honeynets are splitted into two classifications: low-collaboration honeynets and high-association honeynets. Low-communication honeynets recreate a low arrangement of framework functionalities like the system stack. The upside of low-association honeynets is adaptability, permitting than checking of several IP addresses. The real disadvantage is that can't screen in detail the aggressors' activities once the host is bargained. High-association honeynets are mimic every one of the functionalities of genuine framework which is permitting the social occasion of substantially greater data to the assailant's activities. The downside of utilizing high communication honeynets is their high cost and troublesome administration.

At the point when the honeypot is in genuine machine with its IP address we call it as a physical honeypot. Assailants have some approaches to test the system endeavoring to distinguish a honeypots. Numerous honeynet organizations not utilize particular procedures to secure the mapping of honeypots by aggressors. In the event that a host Is tested, it give reaction. This is profitable to aggressor as she get data about honeynet. A safeguard is dodge the mapping of honeypots to cunningly control the reaction to beguile the assailants.

The physical honeypots are costly to introduce and furthermore keep up and they are not adaptable, if the virtual honeypots are versatile and after that reasonable. The enhancing electronic correspondence and assembling in different associations and furthermore financial substances on a worldwide scale brings about the data administration frameworks, which are more mind boggling and they are advanced than harbingers. This systems need to guarantee the respectability, characterization of set away information. We are engaging particular customers to get to the commendable data and organizations. Grabbing these goals is possible with the standard check system for right now recognizing clients, which is a piece of the entrance control instrument to figuring out what are data clients and qualified for get to. The security of an entrance control framework is of the prime significance and is likewise essential for a fruitful activity. Be that as it may, static defensive measures are not adequate it is utilized to secure a complex organized framework. Soaccess control frameworks require interruption location (ID) as a vital piece of their task. Interruption discovery frameworks (IDSs) can expand assurance by checking their occasions in the systems administration framework, and dissecting them for indications of security issues by cautioning the framework overseers as appropriate.MANETs, bargained hubs may give potential Byzantine disappointments in a steering conventions. In a Byzantine disappointment, set of hubs could be bargained .Incorrect and malevolent conduct can't be distinguished. Malignant hubs can perpetrate By zantine disappointment on a framework, by making new steering messages, publicizing non-existent connections and furthermore giving mistaken connection state data. It can seen that interruption aversion measures, as same as intrusion discovery methods were generally utilized as a part of wired systems keeping in mind the end goal to secure arranged frameworks. Interruption recognition systems outfitted towards wired systems. This is a direct result of the last's absence of a settled framework, versatility and the powerlessness of remote transmissions to overhang dropping and the trouble of clear partition amongst ordinary and unusual conduct of the hubs. The specially appointed systems administration worldview not considers the nearness of the activity focus focuses in arrange, so most customary interruption recognition frameworks (IDS) outfitted towards a wired systems in light of such design.

Where we give a short prologue to the ideas of the current strategies. In Section 4, we show our model for beguiling assault and safeguard diversion in honeypot-empowered systems for the remote systems .

Related Work

In [1],AnimeshPatcha presents an amusement theoretic model to break down interruption recognition in portable impromptu systems. We utilize amusement hypothesis to show the associations between the hubs of an impromptu system. We see the communication between an assailant and an individual hub as a two player non-agreeable diversion, and develop models for such an amusement. The goal of the aggressor is to send a noxious message from

some assault hub with the intension of assaulting the objective hub. The interruption is esteemed effective when the malevolent message achieves the objective machine without being recognized by the host IDS. We accept that an interruption is distinguished and the barging in hub is blocked when a message sent by a plausible gate crasher is caught and the host IDS can state with assurance that the message is malevolent in nature.

In [2],NandanGarg proposes a diversion theoretic system for demonstrating misdirection in honeynets. The system depends on broad recreations of defective data. We think about the harmony arrangements of these diversions and show how they are utilized to decide the methodologies of the assailant and the honeynet framework. The assailant must assault one of these hosts. The objective of the assailant is to assault a host that isn't a honeypot while the objective of the protector is to have a honeypot assaulted by the aggressor. The aggressor tests the hosts and breaks down their reactions to distinguish honeypots. A smart safeguard will react to the tests in a way that the aggressor does not pick up data on whether the host is a customary one or a honeypot. The reaction of the host relies upon whether the host is a honeypot or a customary host. The reaction of the honeypot might be created with the end goal that to cover the character of the honeypot.

In [3],QuangDuy La considers the issue of guarding against assaults in honeypot-empowered systems by taking a gander at a diversion theoretic model of trickery including an aggressor and a safeguard. The aggressor may endeavor to bamboozle the safeguard by utilizing diverse kinds of assaults extending from a suspicious to an apparently typical movement, while the protector thusly can make utilization of honeypots as an apparatus of misleading to trap assailants. The issue is demonstrated as a Bayesian round of deficient data, where equilibria are distinguished for both the oneshot diversion and the rehashed amusement adaptations. Our outcomes demonstrated that there is a limit for the recurrence of dynamic aggressors, above which the two players will take tricky activities and underneath which the safeguard can stir up his/her procedure while keeping the assailants achievement rate low.

In [4],SurabhiThukral presents an intrusion detection module based on honeypot technology, which utilizes IPTrace back technique and mobile agents. By using the mobile agents, this module has the capability of distributed detection and response. By the use of honeypot technology the intrusion source can be traced to the farthest. Honeypot intruder obtained by information security experts can make a better understanding of the various attacks, security experts to provide a learning platform for all kinds of attacks and better protect the system should be protected. IP Trace back is a critical ability for identifying sources of attacks and instituting protection measures for the Internet.

In [5], TansuAlpcan presents an amusement theoretic investigation of interruption recognition in get to control frameworks. A security diversion between the assailant and the interruption identification framework is explored both in limited and nonstop bit renditions, where in the last case players are related with particular cost capacities. The circulated virtual sensor arrange in view of programming specialists with blemished identification capacities is additionally caught inside the model presented. This model is then stretched out to consider the dynamic qualities of the sensor arrange. Properties of the subsequent dynamic framework and rehashed diversions between the players are talked about both logically and numerically. A virtual sensor is characterized as a self-governing programming operator that screens the framework and gathers information for recognition purposes. These sensors report conceivable interruptions or irregularities happening in a subsystem of an expansive system utilizing a particular strategy like mark correlation, design identification, factual investigation, and so forth.

In [6], Weizhe Zhang examines every conceivable assault and security dangers to the IoT and proposes related security advancements. This examination displays an outline of an exceptionally refined security subject and general two-dimensional security engineering incorporated with related wellbeing innovations. the model incorporates a total security work and applies the SOA design. The plan possibly builds up the establishment of future security engineering, thinking about its extraordinary self-sufficiency and customization. This model can be effortlessly

outlined into security as an administration and incorporated into a wide range of programming as an administration, stage as an administration, and foundation as an administration SOA.

In [7], urgenMarkert breaks down techniques for following interruptions and unapproved access by an assailant and additionally recognizing endeavored assaults against the steadiness of remote sensor systems, by utilizing interruption identification frameworks. Up until now, almost no examination has been done in the improvement of honeypots. Such frameworks speak to the ebb and flow cutting edge in the field of research on recognizing assaults on remote sensor systems. It is viewed as that inadequate research has been led in the advancement of remote sensor organize honeypots, in this way offering insufficient open doors for particular examination and catching assaults.

Existing System

The current framework centers around one shot assault and safeguard. It additionally clarifies rehashed assault and barrier game.In this situation, the assailant makes the move first while the protector watches the activity andresponds as needs be. Subsequently, the amusement is a dynamic successive diversion with broad frame portrayal. The diverse methods of an assailant are separated by the sorts,A (dynamic) or P (detached). Aggressor's write is private data and not known to the protector, henceforth the amusement is of inadequate data. The genuine sort of an aggressor toward the beginning of the amusement is displayed by the result of an irregular occasion, i.e., a possibility move to randomize over the assailant's composes. The assailant's activities are either N (ordinary) or(suspicious).Equivalently, the aggressor is said to send two signs (N and S) to the safeguard. After watching a flag, the protector must choose how to respond and he/she likewise has two recognizable activities: R (general) for enabling the movement to the normal frameworks and H (honeypot) for rerouting the movement to honeypots. After both have acted, players get diverse settlements relying upon the succession of activities. Different viewpoints are mulled over while assessing the adjustments for the two players. Prizes and punishments are given to the client in light of criteria.

The one-shot assault and safeguard amusement catches one experience between the protector and an obscure aggressor. Over the long haul, the protector faces a substantial number of such experiences autonomously. The circumstance can be demonstrated as limitlessly numerous redundancies of the one-shot amusement after some time, i.e., a rehashed variant of the assault and protection diversion. In diversion hypothesis, this has a place with the class of amusements known as "multi-organize recreations with watched activities and inadequate data" .Accordingly, the diversion's settlements are thought to be unaltered and undiscounted over every reiteration. Likewise, players must receive methodologies determined in view of amusement history, the succession of past activities prompting the present stage. Here, the suspicion of knowing about history is sensible as the assailant and safeguard ceaselessly watch past activities and alter their systems as needs be.

Proposed System

To give insurance against potential assaults, multi-layer safety efforts are proposed for frameworks with remote applications, in which honeypot-empowered interruption location part adds additional profundity to the resistance. Investigate the assault and protection moves making place between the IDS and an assailant. It is to give a basic leadership support to enhance arrange security. We approach the part of conceivable system assaults. The decisions must be made in order to improve network security. First, we need to understand network attacks. For this we have to investigate further what the causes and consequences of network vulnerability are and how the problem has been approached. Second, we need to design a way to find possible network attacks.

Anomaly Detection is used to detect the nodes which are processed are anomaly nodes or normal nodes. Thedataset is loaded and flow based detection is done. Detection is based on the splitting the node on connection. Next graph based detection is used. It is based on the most connected nodes that forms cluster from nodes.Pivotal Node

Identification is based on the anomaly detection result. Load the anomaly detection data and find the pivotal nodes. Pivotal nodes are connected from source to destination nodes many times. File transmission rate will be highfrom source nodesto destination node.Bot Discovery is based on the load pivotal nodes data. Botdiscovery is based on the community detection. First load the pivotal nodes data and find the social correlation graph. That graph is based on the user source node data transmission the destination node.Community Detection is based on the social correlation graph data. Find the modularity measurement based on this technique. Load the social correlation data and modularity measure. That modularity measure based on the source to destination connection and file transmission.Select data from the drive storage to load into database.Load the data into database for analysis. Pre-process the data to remove the irrelevant record from the data. View thepre-processed data for acknowledgement.Get pre-processed data for find out the abnormal events from the whole data records.Analyse user activities using their port address.Find out the abnormal events from the normal events.View collected abnormal data.

Uses game theory to classify the passive users.Here gather information from abnormal events and the original data.Classify the active user and the passive user with their behaviour analysis.View both active user data and passive user data.Get report of active and passive user.If active user, then direct to regular system.If passive user, then they will redirect to honeypot 3.View honeypot user for acknowledgment.Store passive attacker port address to future analysis.

CONCLUSION

The proposed framework is utilized to distinguish different aggressors in a remote system. The system security is kept up at a predictable level.Theconditions underwhich the trickery assaults to distinguish different aggressors are executed and the assailants are recognized accurately.Modeling theinteraction between the attacker(s) and the IDS as bothfinite and continuouswhich gives understanding into and addresses a wide rangeof asset distribution issues in interruption detection.In expansion, the association betweenthe players over a day and age is examined utilizing repeatedgames and a particular dynamic model for the sensor network.Finally, some fundamental procedures for the IDS and the attackerare talked about.

REFERENCES

1. A. Patcha and J.-M. Park, “A game theoretic formulation for intrusion detection in mobile ad hoc networks,” Int’l J. Netw. Security, vol. 2, no. 2, pp. 131–137, Mar. 2006
2. N. Garg and D. Grosu, “Deception in honeynets: A game-theoretic analysis,” in Proc. IEEE Workshop on Information Assurance, 2007, pp. 107–113.
3. QuangDuy La, Tony Q. S. Quek and Jemin Lee,” A Game Theoretic Model for Enabling Honeypots in IoT Networks”, IEEE International Conference on IoT, 2016.
4. SurabhiThukral, RutbaMaqsood and DivyaUpadhyay,” To Design an Intrusion Detection System based on Honeypot using Mobile Agent and IP Traceback Technique”, IJSR, Volume 2 Issue 4, April 2013.
5. T. Alpcan and T. Basar, “A game theoretic analysis of intrusion detection in access control systems,” in Proc. IEEE Conference on Decision and Control, Dec. 2004, pp. 1568–1573.
6. W. Zhang and B. Qu, “Security architecture of the Internet of Things oriented to perceptual layer,” Int. J. Comput. Consum. Control, vol. 2, no. 2, pp. 37–45, 2013.
7. J. Markert and M. Massoth, “Honeypot framework for wireless sensor networks,” in Proc. Intl. Conf. Advances in Mobile Computing & Multimedia, Dec 2013, pp. 217:217–217:223.
8. T.Mekala,”Improved security in adaptive steganography using game theory”,vol .118,No.8,Feb 2018,pp 111-115.
9. T.Manimaran,”A Theoritical analysis of Different Hacking Techniques in Wireless Networks”,IJCSET.vol.6,no.9,2015.

10. T. E. Carroll and D. Grosu, "A game theoretic investigation of deception in network security," Security Comm. Networks, vol. 4, no. 10, pp. 1162– 1172, 2011.

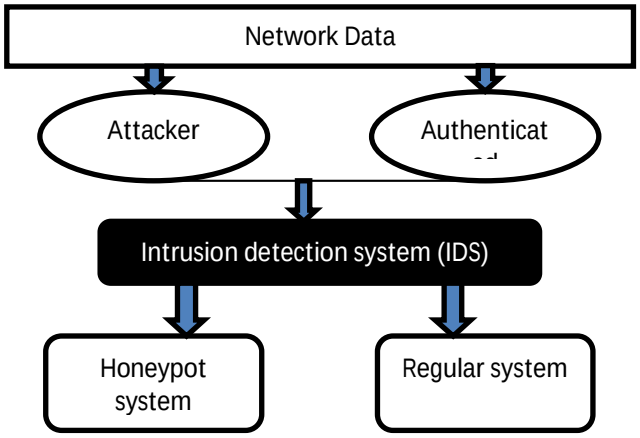


Fig: 1

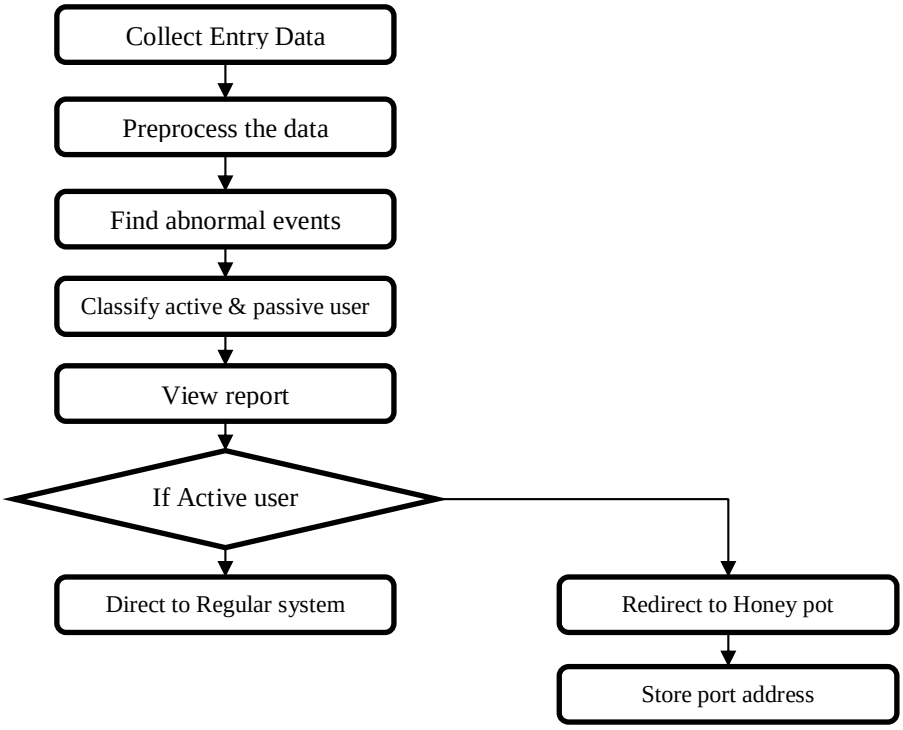


Fig: 2

Indian Journal of Natural Sciences www.tnsroindia.org.in ©IJONS

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Identify Risk Factor and Perform Prediction of Dengue Fever Using Data Mining

Selvi A* and Bharani nayagi S

Department of Computer Science,M. Kumarasamy College of Engineering, Karur, TamilNadu, India

Received: 10 Feb 2018 Revised:18 Feb 2018 Accepted: 30 Mar 2018

***Address for correspondence**
Selvi. A
Assistant Professor,
Department of Computer Science,
M. Kumarasamy College of Engineering,
Karur, TamilNadu , India
E mail: selvia.cse@mkce.ac.in / itbharansiva@gmail.com

ABSTRACT

Arbovirus causes Dengue Fever (DF) and it spread by Aedes mosquitoes. The symptoms of Dengue Fever are high fever, excessive sweating, body pain, rashes, loss in appetite. The infection of Dengue Fever may leads to Dengue Haemorrhagic Fever (DHF) and Dengue Shock Syndrome (DSS). DSS is the severe form of DHF. The World Health Organization (WHO) declares Dengue Fever and Dengue Hemorrhagic Fever are widely spread in all regions. Data mining is most popular to analyze the large number of data and discover algorithm under computational efficiency limitation that produces enumeration of models (or pattern) over the given data. Medical data mining is a significance of Data Mining and also an important research field in healthcare domain. In this paper, the risk factors for having dengue fever are obtained by decision tree algorithm and the predictions are done by regression. The research result shows prediction accuracy of 99%.

Keywords: Medical data mining, Decision Tree Algorithm, Regression, DF

INTRODUCTION

In India, first occurrence of DF is at Vellore district of Tamil nadu in the year of 1956. In 1963, 33% of cases showing hemorrhagic manifestations in Calcutta. Three manifestations of dengue disease are Dengue Fever (DF), Dengue Hemorrhagic Fever(DHF) and Dengue Shock Syndrome Dengue(DSS)[1]. It is caused by four different serotype viruses that belong to the genus Flavivirus and are part of the flaviridae family (DEN I to IV). The female Aedes (Ae) mosquitoes spread the dengue viruses. At present, the first two serotypes (DEN1 and DEN2) are widely spread in

India. DF is also called 'break bone fever' or 'dandy fever'. The spectrum of the dengue virus infection encompasses asymptomatic forms to severe cases with shock, organ dysfunction and relevant bleeding[2].

The development of Information Technology generated and stored huge amount of databases. The information technology manipulates the stored data for further decision making. In this paper, the datasets such as different vector, environmental factor and host factor are generated and classified by the Decision Tree Algorithm.

After classification, we undergo prediction. Prediction construct a model. With the given input, we can predict continuous or ordered value. Classification is to predict categorical class label whereas prediction are models and it is a continuous-valued functions. prediction include accuracy, speed, robustness, scalability, interpretability. The major method for prediction is regression.

Problem Statement

The computer-based patient records minimize medical error, produce confidential patient record, reduce redundant practice variation and enhance patient outcome. This can be obtained by implementing medical data mining. Prediction of Dengue Fever is obtained by medical data mining. It is widely applied for prediction or classification of different types of DF. The key objective of the works are,

Using data set identify key patterns or features.
The relevant in relation to DF attributes are identified and selected.
The domain experts are used to analyze the results of the selected model.

Medical Data Mining

Data mining techniques play a vital role in healthcare domain. The data mining along with healthcare industry provides a survive regarding early detection of disease. The clinical diagnosis in data mining provide various healthcare related systems. Data mining is used to analysis the causes, symptoms, and courses of treatments. Data mining provide a set of tools and techniques to process data and discover hidden patterns. It help healthcare professional by providing additional source of knowledge for making decision. Data mining in healthcare uses different technique to store information such as number of days stay in a hospital, hospitals rank, best effective treatment, fake insurance claimed by patient and provider, readmission of patients, identifies better treatment method for a particular group of patients, effective drug recommendation systems, etc[3]. Healthcare industry generate huge amount of complex data about patient in a daily basis, based on the medical record, doctors provide better medical services to the patients. Medical management decision are provided on the basis of hospital resources, disease diagnosis, electronic patient records, medical devices, administrative reports and other benchmarking finding etc [4]. For decision making, these dataset are key resource to be processed. The key resource are analyzed for knowledge extraction[5].

In healthcare system, the relevant healthcare information are identified by an automated tools in data mining. The relevant information such as process of diagnosis, treatment , prevention of disease, injury and other physical and mental impairment in human. Prediction of numerous diseases are obtained by data mining in healthcare. It is very essential in making clinical decision. The essence of data mining is to identify relationship, pattern and predictive model. It integrates hospital information as a model of decision making. The decision-making process is implemented for diagnosis and treatment planning. Decision making reduces subjectivity as well as reduces time for decision making. The progress of data mining is predicting the comparable patients and classified them based on illness or fitness issues and healthcare system gives them an effectual treatments. Illness is due to the factors that are accountable for diseases such as food type, diverse working situation, learning level, livelihood conditions, accessibility of pure water, health care services, artistic, ecological and farming factors[6].

Dengue Fever (DF)

The bite of infected female Aedes mosquito spread Dengue Fever. A person with a weaker immunity can quickly affected by Dengue Fever(DF). Sometimes the disease may get severe and fatal. The two forms of dengue disease are Dengue Fever(DF) and Dengue Haemorrhagic Fever(DHF) are familiar in India. A more severe form of dengue is Dengue Hemorrhagic Fever, hence it would lead to death. The symptoms of dengue hemorrhagic fever include headache, rashes and fever besides bleeding in the body (haemorrhage), formation of red splotches on the skin, black colored stool, bleeding in gums or nose and a weakened immunity. Dengue hemorrhagic fever is life threatening and causes dengue shock syndrome. The pathway of dengue fever is from 'infected person to Aedes mosquito to another person'. It cannot spread through direct human contact. DF can be prevented by eliminating mosquito breeding places such as Cover water containers. It also includes Removal of rubbish, Biological control, Chemical control[7].

The symptom of DF are:

- High fever(up to 104 degree Fahrenheit)
- Excessive sweating
- Joint Pain
- Nausea and vomiting
- Loss in appetite
- Lower blood pressure
-Low heart rate along with hypotension
- Skin rashes
- Swelling in hands and soles of feet

Causes of Dengue Fever (DF) are[8]:

-Abrupt onset of high fever
-Severe frontal headache
-Pain behind the eyes which worsens with eye movement
-Muscle and joint pains
-Loss of sense of taste and appetite
-Measles-like rash over chest and upper limbs
-Nausea and vomiting

Causes of Dengue Haemorrhagic Fever & Shock (DHF) are[8]:

-Symptoms similar to dengue fever
-Severe and continuous stomach pains
-Pale, cold or clammy skin
-Bleeding from the nose, mouth and gums and skin bruising
-Frequent vomiting with or without blood
-Sleepiness and restlessness
-Persistent vomiting
-Excessive thirst (dry mouth)
-Rapid weak pulse
-Difficulty in breathing
-Fainting

CLASSIFICATION

Classification is a supervised learning. It includes model construction and model usage to estimate accuracy. Classification deals with classifier and predictor accuracy, speed, robustness, scalability, interpretability. Classification under Decision Tress Algorithm provide a high accuracy, sensitivity and specification.

Decision Tree Algorithm

Decision tree is a form of tree structure. In this paper, we built decision tree for classification and regression models for prediction. In Decision tree, the data set has been break down as a smaller subsets. Exactly like the Top-down approach. It consists of two nodes: leaves node and root node. Node with incoming and outgoing edges are called as Internal node. Root node is also known as decision node. Each decision node consist of two or more branches. Leaf node represent a classification or decision. The unambiguous and logarithmic data are handled by decision tree. Using attributes of the logarithmic data, classification are performed. Decision Tree Algorithm is the process of huge amount of data using data mining application. No domain knowledge or parameter settings are required for constructing decision tree.

Algorithm

J.R Quinlan established C4.5 Algorithm. For building decision tree, C4.5 Algorithm play a vital role. It is a top-down approach . Greedy search are carried out through the space of all possible branches with no backtracking. C4.5 Algorithm consist of following two steps,

Entropy.
Information Gain.

Entropy

As decision tree is a top-down approach, it split the dataset as a subset with homogenous values. The homogeneity of a sample are calculated by entropy of C4.5 algorithm. The sample are homogeneous when the entropy value is zero. The sample are equally divided when the entropy value is one. There are two types of entropy to build a decision tree, the corresponding frequency table are generated as follows[9].

a) The frequency table of one attributes:

$$E(S)=\sum_{i=1}^{c} - p_i \log_2 p_i$$

b) The frequency table of two attributes:

$$E(T,X) = \sum_{c \in X} P(c)\, E(c)$$

Information Gain

The information gain is the process of reducing the entropy value after a dataset is split on an attribute. The attributes with highest information gain are used to construct a decision tree.(i.e., the most homogeneous branches). In decision tree, the decision node is the highest information gain. The Information Gain is also known as decrease in entropy. It includes the following steps,

Step 1: Calculate target entropy.
Step 2: The datasets are divided into different attributes. The

value of entropy for each attribute is calculated.Total entropy is the summation of all split entropy.

Step 3: The Information Gain is the subtraction of the result entropy from the entropy before split. The Information Gain is illustrated as below[9],

Gain(T,X) = Entropy(T) - Entropy(T,X)

Step 4: Attribute with the biggest Information Gain is selected as the decision node.

Step 5(a): Leaf node entropy value is 0. It is represented as triangle.

Step 5(b): Splitting are performed for a node with entropy value more than 0. It is a internal node. It is represent by a circle[9].

Decision Rule

Mapping from the root node to a leaf node are carried out to transfer decision tree to a set of rules[9].

Regression Method

Regression is a supervised learning in a data mining technique. Regression technique can be adapted for predication. It consists of two variables namely independent variables and response variables. Independent variables are attributes contains already known variable. Response variables are what we want to predict. In real-world, the problem not only deals with prediction it also depend on complex interactions of multiple predictor variables [10].

Types of regression methods

- Linear Regression.
- Multivariate Linear Regression.
- Nonlinear Regression.
- Multivariate Nonlinear Regression.
- Linear Regression

Visualizing linear regression with a predictor and a prediction is the simplest form of regression. Y axis consist of prediction value. X axis consist of predictor value. The simple linear regression is represented as a straight line. The linear regression model reduces the error rate between the the actual prediction value and the point on the line (the prediction from the model), as shown in the Fig. 2 below.

The following notations are used

x - predictor.

y - prediction

$\theta 1$ - At (x=0), x crosses y axis is the y intercept.

$\theta 2$ - Slope of the line is the angle between data point and the regression line.

The linear regression with a single predictor equation is as follows:

$y = \theta 1 + \theta 2\, x + e$ (2).

This equation is similar to the line equation[10].

Multivariate Nonlinear Regression

In case of multiple predictor, two dimensional space are not used to visualize the regression line. Let x1,x2,.....,xn be a number of predictor in the linear regression and it is referred as multivariate linear regression. The straight line

can be measured by the following equation. The equation is a extension of single-predictor linear regression equation[10].

$y = \theta 1+\theta 2\ x1+ \theta 3\ x2 +....+ \theta n\ x(n-1)+ e$ (3)

Regression Coefficient

The regression coefficient is defined as the regression parameter of the multivariate linear regression. While generating a multivariate linear regression model, the corresponding co-efficient for each of the predictor has been generated by using the algorithm. The smash of the x predictor on the y target is the measurement of the regression coefficient. The regression coefficients can be analyzed by using numerous statistics .These statistics are used to evaluate the regression line that fit the data.

Nonlinear Regression

A nonlinear regression technique is used when x and y co-ordinates are not in a straight line. In this case, to make the relationship linear regression, the data could be preprocessed. The nonlinear relationship between x and y are shown in below figure.

Multivariate Nonlinear Regression

When there are two or more predictors such as x1, x2, ..., xn in a non-linear regression then it is referred as multivariate nonlinear regression. Similarly like the multivariate linear regression, the relationship in the multivariate nonlinear regression cannot be visualized in two-dimensional space [10].

CONCLUSION

In this paper, prediction of Dengue Fever is obtained by implementing data mining techniques. We studied an efficient approach for the extraction of significant patterns from the dataset of the Dengue Fever. It is efficient method for prediction of Dengue Fever. Data mining in medical analysis is expressive in performing a task for the classification and prediction of Dengue Fever. Data mining provide the high quality clinical decision-making. It provide rich environment knowledge to perform classification and prediction. Decision tree technique are used for data mining classification. Regression method for prediction.The further enhancement for the automation of Dengue Fever prediction will be the proposed work. While increasing more number of inputs such as structured and unstructured data. The unstructured data variables in the healthcare industry database are obtained by using text mining technique in data mining.

REFERENCES

1. R. Lall, V. Dhanda, "Dengue haemorrhagic fever and the dengue shock syndrome in India", Review Article, The National Medical Journal Of India Vol. 9, NO. I, 1996D. Kornack and P. Rakic, “Cell Proliferation without Neurogenesis in Adult Primate Neocortex,” Science, vol. 294, Dec. 2001, pp. 2127-2130, doi:10.1126/science.1065467. (Article in a journal).
2. Nivedita Gupta, Sakshi Srivastava*, Amita Jain* & Umesh C. Chaturvedi, "Dengue in India", Centenary Review Article, Indian J Med Res 136, September 2012, pp 373-390 . (Article in a journal).
3. Parvez Ahmad, Saqib Qamar, Syed Qasim Afser Rizvi, Techniques of Data Mining In Healthcare: A Review, International Journal of Computer Applications (0975 – 8887) Volume 120 – No.15, June 2015. (Article in a journal)

4. Neesha Jothia, Nur'Aini Abdul Rashidb, Wahidah Husainc, a*, "Data Mining in Healthcare – A Review", Procedia Computer Science 72 (2015) 306 – 313, Elseiver.
5. M. Durairaj, V. Ranjani, " Data Mining Applications In Healthcare Sector: A Study", International Journal of Scientific & Technology Research Volume 2, Issue 10, October 2013.
6. R. Naveen Kumar 1, M. Anand Kumar 2, "Medical Data Mining Techniques for Health Care Systems", Research Article, DOI 10.4010/2016.811 ISSN 2321 3361 © 2016 IJESC.
7. Bruce R. Guerdan, "Dengue Fever/Dengue Hemorrhagic Fever", American Journal of clinical medicine® • Spring 2010 • Volume Seven, Number Two.
8. https://www.mayoclinic.org/diseases-conditions/dengue-fever/symptoms-causes/syc-20353078Sankari Karthiga1, M. Safish Mary2, M. Yogasini3, " Early Prediction of Heart Disease Using Decision Tree Algorithm", International Journal of Advanced Research in Basic Engineering Sciences and Technology (IJARBEST) Vol.3, Issue.3, March 2017.
9. Nanhay Singhi, Ram Shrinar Raw1 and Chauhan R.K.2, "Data Mining with Regression Technique", Journal of Information Systems and Communication ISSN: 0976-8742 & E-ISSN: 0976-8750, Volume 3, Issue 1, 2012, pp.-199-202.
10. Rosmari Rodriguez-Roche1 and Ernest A. Gould2,3, "Understanding the Dengue Viruses and Progress towards Their Control", Review Article, Hindawi Publishing Corporation BioMed Research International Volume 2013, Article ID 690835K. Elissa, "Title of paper if known," unpublished.
11. DB Kadam1, Sonali Salvi2, Ajay Chandanwale3, "Expanded Dengue ", Review Article, Journal of The Association of Physicians of India Vol. 64.July 2016.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
12. Pinki Sagar, Prinima, Indu, "Analysis of Prediction Techniques based on Classification and Regression", International Journal of Computer Applications (0975 – 8887) Volume 163 – No 7, April 2017.
13. Mr. Sudhir M. Gorade1, Prof. Ankit Deo2, Prof. Preetesh Purohit3, "Early dentification of Disease Based on Responsible Attribute Using Data Mining", International Research Journal of Engineering and Technology (IRJET), Volume: 04 Issue: 07

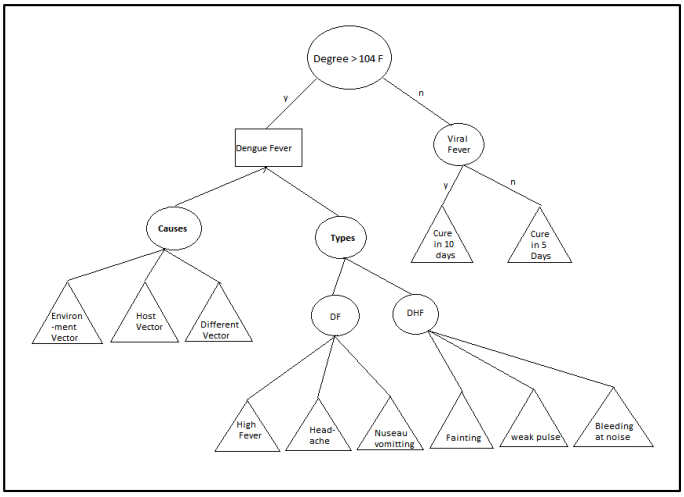


Fig.1-Decision Tree Classification for Dengue Fever.

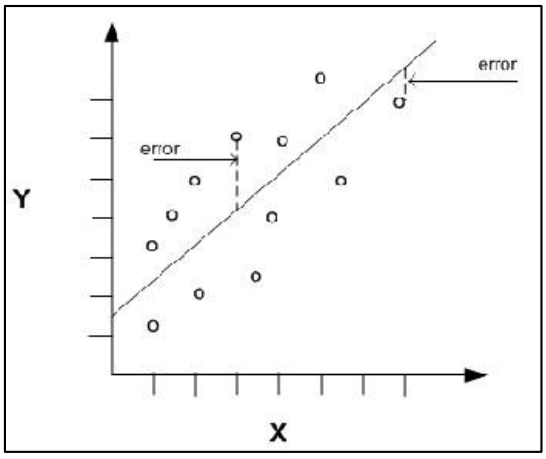


Fig. 2- Linear relationship between x and y

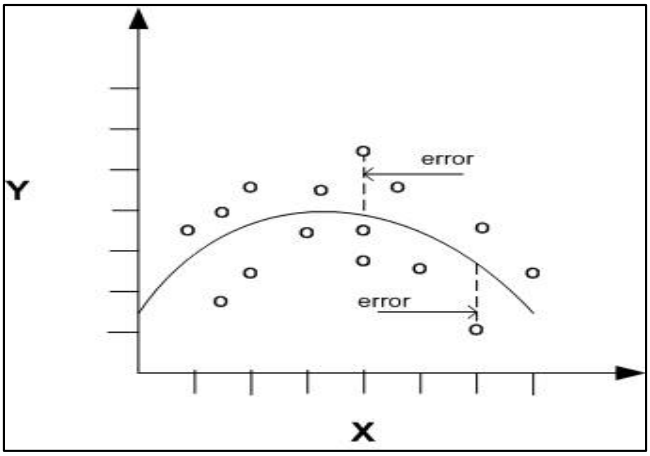


Fig. 3- Non-Linear relationship between x and y

Indian Journal of Natural Sciences www.tnsroindia.org.in ©IJONS

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Criminal Attacks Detection Using Rowhammer Technology

Yazhini S P*[1], Gayathri.R[2], Gowri sankar.K[3] and Nadhaan.R.R[4]

[1]Assistant Professor,Department of CSE, M.Kumarasamy College of Engineering, Karur, TamilNadu,
[2,3,4]Student, M.Kumarasamy College of Engineering, Karur,TamilNadu,India.

Received: 11 Feb 2018 Revised:19 Feb 2018 Accepted: 30 Mar 2018

***Address for correspondence**
Yazhini S. P,
Assistant Professor,Department of CSE,
M.Kumarasamy College of Engineering, Karur, TamilNadu,India.
E mail: gayathrirlgk@gmail.com

ABSTRACT

As an evolution of modern era, people are interlinked with social networks and data are stored in phones. All the information is not completely safe. A mass information gathering and connection structure in light of those data that spilled from the remote gadgets that individuals convey. This project is used to find the criminals and anonymous network users by the data. The device named Snoopy is composed in Python.Capable of working in an appropriated manner.Snoopy apparatus can draw particular and abnormal state conclusions and about people details in light of their computerized remote signalsfrom the remote gadgets that individuals convey. The Row hammer hardware bug allows an attacker to modify memory without accessing it, simply by repeatedly accessing, that is —hammering", a given physical memory location.

Keywords: Snoopy apparatus, Python, Row hammer

INTRODUCTION

Communication devices such as wearable computers, smart phones disclose wireless signals even the devices is not in use on the off chance that the signs uncover by at least one of these gadgets are one of a kind, the gadget can be recognized as being in a specific area at a specific time signalsfrom the remote gadgets that individuals convey.Those signals gather an information and details about the user and users data. The Snoopy tool was originally released for the concept for detecting and tracking.Since then it can rewritten to be secluded with an innovation freethinker structure ready to gather signals disclose from arbitrary technologies as defined by its plugins. The framework is designed to run for long periods and synchronized data to a central server via Wi-Fi.

Drammer is a case of the Flip Shui ex-pion strategy that abuse the physical memory allocator to surgically incite equipment bit ips.Aggressor picked touchy information which out of the blue depends just on dependably on product highlights.Row hammer-based Flip Feng Shui assault to be fruitful, three natives are critical. In the first

place, aggressors should have the capacity to pound adequately hard hitting the memory chips with high frequency.For event, no bits will ip if the memory controller is too moderate. Second, they have to figure out how to message a physical memory with the goal that the right, exploitable information is situated in the avoided physical page.

While attacking the mobile devices,none of those natives can be executed by essentially porting existing strategies. Rowhammer is a product incited equipment that for the most part influences dynamic irregular get to memory chips.In hone, this has the net impact that a bit of software can ips some bits in physical memory by uniquely performing memory read operations. It is imperative to take note of that setting off the Rowhammer bug it is unique in relation to utilizing as a part of a security-significant way. Indeed, an adventure as a rule needs to trap a casualty segment (e.g., another procedure, the OS, or another VM facilitated on the same physical hub) to utilize an unprotected physical memory area to store security-delicate substance. In the general case, programming misuse of this kind ended up being testing.

LITERATURE REVIEW

The communication media like phone callsand online social networks that leavethe digital traces in the form of metadata.In this paper Girvan and Newman algorithm is used to detects communities and removing edges from the original network.it mainly focus on edges that are most likely "between" communities what's more, distinguish those edges that interconnect hubs having a place with different bunches. It is a various level bunching technique which gatherings of hubs are dynamically accumulated to shape bigger groups.The log analysis maintain the group of data for tracking.so we done the tracking by those data. That data is used to track the location of the criminals.[1] The strategy considered in this paper is to executing DoC assault through Wi-Fi get to point without utilizing a Web association it consequently interfaces as web. At the point when an Android cell phone or specialized gadget enters the scope region of a remote access point, it is consequently alloted the identifier and get stacked into the Wi-Fi heap of the cell phone. Gadget interfaces the "TEST_AP" despite the fact that there is a substantial AP with a similar Administration Set Identifier in the region. On the off chance that the phony access point presented while there is continuous association process with another legitimate Wi-Fi get to point their testing demonstrates that the current association won't get interrupted.Tracking calculations is utilized as a part of this venture is to identify the client current position and they have full created and assessed the assaults and in addition the protection models that keep running on Android telephones[2].

In this paper they showed that capable deterministic Rowhammer assaults that give an aggressor root privileges on a given system.The previously mentioned strategy to mount a deterministic Rowhammer assault on portable stages.This method is used to attack the user by an bug it will gather the data and information from the device.The attack mainly damage the hardware in devices main advantage is it slow down the process.Theyexhibited that few gadgets from different sellers are helpless against Rowhammer.Their inquire about demonstrates that handy extensive scale Rowhammer assaults are a genuine risk and keeping in mind that the reaction to the Rowhammer bug has been moderately moderate from merchants[3].

In this paper they gives a review of the current protection arrangements and examines the investigations and measurements that are generally planned and used to assess their execution, which is useful for the future research in the area.pattern matching algorithmisused inthis paper it distinguish the activity stream that is indistinguishable to assault stream and searches for the beginning of the assault and they displayed some cutting edge solutions:some were somewhat simple to fuse in existing Cloud foundations for Cloud suppliers to avoid or decrease DoS and DDoS assaults[4].

In this paper they talked about form a mass information accumulation and connection structure in view of data spilled from the remote gadgets that individuals convey and the one of a kind marks that cell phones emanate were

acquainted and assembling a system with distinguish a client's DTF and exhibit how the client can be followed by the way,And what interface examination could be led against information in both dynamic and latent way.In this paper they talked about form a mass information accumulation and connection structure in view of data spilled from the remote gadgets that individuals convey and the one of a kind marks that cell phones emanate were acquainted and assembling a system withdistinguish a client's DTF and exhibit how the client can be followed by the way,And what interface examination could be led against information in both dynamic and latent way.[5]

Problem Statement

Still we have a problem to identify the activity of the criminals .nowadays the criminals find by their locations. The location identify by their communication signals we have a capable to lock their locations in particular area when their signals only in on state. Here we are going to do that the operation of lock the address of the signals to find the criminals locations when it is off.

Existing System

A large number of data collection framework based on their information that collections of data also called as correlation data.Those data and correlation or collection of data are used to track the wireless devices.The framework is used to identify peoples profiles collected wireless information from user devices.The collected information having more verbose by optionally interrogate devices.These collected data and framework lack from the those wireless devices that people carry.

Disadvantage of Existing System:

- The user has to be within range of the access point.
- Root access cannot be processed i.e. deleting, modifying etc.

Proposed System

The system is used to develop a network connection between the two devices like the system, mobiles. The Rowhammer exploit technology is proposed in this application. In Row hammer technology, snoopy tool is used to provide automatically connect with the target system like android or ARM. It will transfer the necessary information to the hackers from the hacked device. Then the hacker can easily change the content, hacked from the targeted system.

In this system the root access permission will be granted to the hackers by the snoopy tool to read and write from the targeted system. The hacker can easily get the control of the system so the hacker will modify the content in the targeted system.

Advantages of Proposed System

- The target can be accessed even when the target is not inside the coverage area.
- Privilege will be given for Root access and process such as deleting and modifying could be done.
- Log analysis is possible.

CONCLUSION

This paper shows the special marks in the cell phones reveal was presented from this underlying perception a theory were built .Tracking the client in light of their gadgets and distinguish their own data about gadget user.Numerous remote innovations are talked about to recognize their range and conceivable unique mark. The Snoopy system was then presented as a distributed, users profile tracking, data block attempt, and investigation device to discover the criminals.This structure can be utilized to track any signature and clients profile for a reasonable Python module can be composed.

REFERENCES

[1] E Ferrara, S Catanese, P De Meo, and G Fiumara, "Detecting criminal organizations in mobile phone networks", Expert Systems with Applications, Vol.4,Pg.no:5733-5750.

[2] Erich Dondyk, Cliff C. Zou, "Denial of Convenience Attack to Smart phones Using a Fake Wi-Fi Access Point", University of Central Florida."— Presentation transcript, 2012.

[3]Victor van der Veen, Yanick Fratantonio, Martina Lindorfer," Drammer: Deterministic Row hammer Attack on Mobile Platforms", Proceedings of the ACM SIGSAC Conference on Computer and Communications Security, Pages, Pg.No. 1675-1689, 2016.

[4] N,Mahendran, T.Mekala ,"A Survey: Sentiment Analysis Using Machine Learning Techniques for Social Media Analytics ",Journal of Pure and Applied Mathematics"Vol.118, No.8, Pg. No.419-422, 2018.

[5] C.Selvarathi, A.Selvi,"Face Tracking Algorithm for Tracking Target in WSN", International Journal of Pure and Applied Mathematics"Vol.118, No.8, Pg. No.579-584, 201

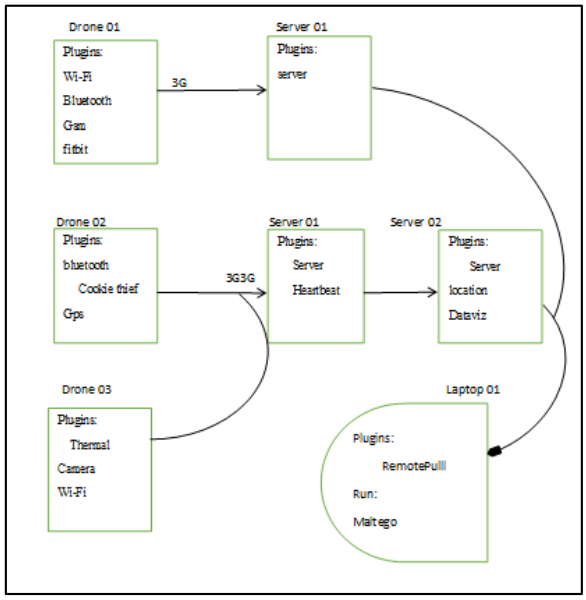


Figure 1.Architecture Design

Indian Journal of Natural Sciences www.tnsroindia.org.in ©IJONS

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

A Dynamic Preserving and Secured Multi-Keyword Ranked Search for Encryption Scheme in Cloud Data Storage

S.Saravanan[1]*, R.Dhanalakshmi[2], R.Ishwarya[2], M.Janani[2] and S.Karthika[2]

[1]Assistant Professor,Department of Computer Science and Engineering, M.Kumarasamy College of Engineering ,Karur, TamilNadu, India.
[2]Final year student,Department of Computer Science and Engineering, M.Kumarasamy College of Engineering ,Karur, TamilNadu, India.

Received: 07 Feb 2018 Revised:19 Feb 2018 Accepted: 30 Mar 2018

***Address for correspondence**
S.Saravanan
Assistant Professor,Department of Computer Science and Engineering
M.Kumarasamy College of Engineering (Autonomous),
Karur, TamilNadu, India.
E mail: saravanans.cse@mkce.ac.in / dhanalakshmiraja1997@gmail.com

ABSTRACT

Due to the greater ubiquity of distributed computing number of information proprietors are inspired to store information in cloud. But the information must be encoded before outsourcing for privacy. This paper underpins multikeyword positioned seek and furthermore bolsters dynamic tasks like inclusion and cancellation of documents. To give multikeyword look through a "Voracious Profundity First Inquiry" calculation is utilized which build a tree based list structure. To look through the closest information the KNN calculation is used. Due to the utilization of our unique tree based structure the pursuit time will be lessened and manages erasure and addition of archives more efficiently. Even however the information is put away in the cloud it isn't that much secured. Thus we are giving more insurance by utilizing Auto key generation. In this procedure the key plays an essential role. The key utilized is mystery so the unapproved individuals can't get to the data. Due to the decreased cost more individuals are surging towards cloud storage.so the put away information must be secured. With the assistance of previously mentioned process we are giving greater security to the information put away in the cloud.

Keywords: Searchable encryption, multi-keyword ranked search, dynamic update, cloud computing.

INTRODUCTION

Circulated figuring is a creating development which provider a significant measure of chances for web sharing of benefits or organizations. One of the key good conditions of CC is pay-as you-go esteeming model, where clients pay

similarly according to their custom of the organizations. Conveyed figuring is a web slanting a handling. All the service in the internet will deliver dynamically. The user depends on net, hardware, software etc.

Benefits

There islittle application in the cloud computing that extends engineering model. Some of the advantages are:

- Reduced Cost
- Increased Storage
- Flexibility

Reduced Cost

Cloud provides information with less cost. The bill pay as indicated by usage; Interchanges will not pay for consequently bring down upkeep. Unique cost and repeating settled cost are much mediocre compared to normal registering.

Increased

With the colossal correspondence that is offered by Cloud provider today, storing and constancy of broad volumes of data is a reality. Rash workload spike are furthermore administered sufficiently and gainfully, since the cloud can scale vivaciously.

Flexibility

This is a basic class. With enormous business need to alter, impressively more rapidly, to related business setting, speed to pass on is basic. Circulated figuring weight on tolerating application to bazaar rapidly, by using the larger part appropriate building squares fundamental for sending.

Related works

1. Q. Shen…,[1]complex get to control on encoded data that we call Figure content Arrangement Characteristic Based Encryption. By using our techniques mixed data can be kept intrinsic paying little heed to whether the limit server is unfrosted; what's more, our procedures are security against complicity ambushes. Going before Quality Based Encryption plot second-hand credit to delineate the encoded data and consolidated system with customer's keys while in our structure attributes are used to depict a customer's respect, and a festival scrambling data chooses a represent for who can unravel. Along these lines, our procedures are theoretically closer to ordinary access direct systems, for instance, Part Based Access Control (RBAC). In addition, we give an utilization of our system and give execution restrict

2.VipulGoyal… ,[2]In an ABE structure, a customer's keys and figure compositions are stamp with sets of illuminating property and a fastidious key can unscramble a particular figure message only if there is a resistance between the characteristic of the figure content and the customer's key. The cryptosystem of Shay and Waters considered unscrambling when in any occasion k attributes cover between figure content and a private key. While this unrefined was offered away to be conventional for botch open encryption with biometrics, the nonattendance of impressibility seems to limit its tangibility to dominating systems. In our cryptosystem, figure content are name with set of property and private keys are connected with contact structure that supervise which figure messages a customer is sagacious to decipher. We make known the tangibility of our agree to dissemination of audit log information and impart encryption. Our structure reinforces task of private keys which subsume Progressive Character Based Encryption (HIBE).

3.Joseph K. Liu… ,[3]In a fine-grained two-factor get the opportunity to control tradition for electronic cloud enroll organizations, using a unimportant security device. The contraption has the going with properties:

(1) It can figure some light computations, e.g. hash and exponentiation;
(2) It is modifying safe, i.e., it is normal that no one can pulverize into it to get the secret information set away inside. With this contraption, our tradition gives 2FA assurance. To begin with the customer puzzle key (which is when in doubt set away inside the PC) is required. Likewise, the security contraption should be furthermore coupled to the PC (e.g. through USB) remembering the true objective to check the customer for get to the cloud. The customer can be yielded get too simply in case he has the two things. Moreover, the customer can't use his riddle key with another contraption have a place with others for the passage. Our tradition chains fine-grained trademark based access which gives a gigantic versatility to the structure to set diverse access plan as showed by disparate circumstances. Meanwhile, the security of the customer is similarly ensured. The cloud structure just understands that the customer has some required attribute, however not the certified identity of the customer.

4.H. T. Dinh… ,[4]ABE just oversees confirmed access on mixed data in circulated capacity advantage. It isn't reasonable to sent by virtue of access control to conveyed figuring organization: The cloud server may scramble a shot message using the passage course of action and demands that the customer unscramble it. If the customer can adequately disentangle the figure content, it is allowed to get to the appropriated registering organization. Regardless of the way that this approach can fulfill the need, it is especially inefficient. In this new idea, a customer can support him/herself to the cloud determine server in mystery. The server just knows the customer secure some essential quality, yet it doesn't be ordinary with the identity of this customer. In supply a k-times edge for obscure access control. That is, the person from staff serving at table may confine a particular course of action of customer (i.e., person's customers with a comparative plan of significant worth) to get to the system for a most outrageous k-time inside a period or an event. Help distinctive access will be denied. We in like manner show the security of our instantiation. Our thriving outcome shows that our arrangement is realistic.

5.B.Wang… ,[5]To achieve more noteworthy flexibility on re-encryption, various varieties of PRE have been proposed, for instance, brief PRE (CPRE), Character Based PRE (IBPRE), and Trait Based PRE (ABPRE). CPRE grants an encryption associated with a state to be changed to another figure content tag with another shape. The advances of IBPRE and ABPRE are to some degree comparative, and a key partition flanked by them is ABPRE acknowledges more expressiveness in data sharing. Besides, the above encryption is sensible to be disfigured to an additional figure content related with another string by a semi trust delegate to whom a re-encryption key is settled. Regardless, the mediator can't get to the essential plaintext. This new obsolete can fabricate the versatility of customers to assign their disentangling rights to others. We furthermore show it as totally picked figure content secure in the standard model.

Existing system

The present frameworks on watchword based information recuperation, which are comprehensively used on the plaintext data, can't be particularly associated on the encoded data. Downloading each one of the data from the cloud and unravel locally is plainly nonsensical. This multi catchphrase looks for plans recuperate filed records in light of the nearness of watchwords, which can't give sufficient result situating helpfulness. In any case, fragile data should be mixed before outsourcing for assurance necessities, which obsoletes data utilization like watchword based file recuperation.

Proposed system

A Protected and Dynamic Multi-catchphrase Positioned Hunt Plan over Scrambled Cloud Information We construct an extraordinary tree-based record structure and propose a "Voracious Profundity first Inquiry" estimation to give

gainful multi-watchword situated look. The proposed plan can achieve sub-coordinate interest time and deal with the deletion and incorporation of records adaptably. Wide examinations are coordinated to demonstrate the capability of the proposed contrive.

1) Abundant works have been proposed under different hazard models to achieve diverse interest helpfulness,
2) Recently, some capable plans have been proposed to help inserting's and deleting assignments on report collection.
3) This paper proposes a secured tree-based chase scheme over the encoded cloud data, which supports multi catchphrase situated interest and dynamic assignment on the file amassing.

CONCLUSION & FUTURE WORK

In this paper, a protected, capable and dynamic chase plot is proposed, which reinforces the exact multi-catchphrase situated look for and the dynamic eradication and incorporation of files. We assemble special watchword balanced twofold tree as the record, and propose an "Insatiable Profundity first Pursuit" estimation to gain ideal efficiency over direct request. Likewise, the parallel chase process should be possible to moreover diminish the time cost. The security of the arrangement I guaranteed against two risk models by using the safe kNN estimation. Exploratory results demonstrate the productivity of our proposed plot. There is as yet various test issues in symmetric SE designs. In the proposed plot, the data proprietor's responsible for making invigorating data and sending them to the cloud server. Along these lines, the data proprietor needs to store the decoded document tree and the information that are essential to recalculate the IDF regards. Such a dynamic data proprietor may not be to a great degree sensible for the dispersed figuring model.In future we will research supporting other multi catchphrase semantics (e.g., weighted inquiry) over mixed data, dependability check of rank demand in question thing and security guarantees in the more grounded hazard show. Enabling to take a gander at secured internal things, predicate encryption we use for cloud server thinks about some establishment information on the outsourced dataset.

REFERENCES

[1] H. Liang, L. X. Cai, D. Huang, X. Shen, and D. Peng, “An SMDP-basedservice model for interdomain resource allocation in mobile cloud networks,"IEEE Trans. Veh. Technol., vol. 61, no. 5, pp. 2222_2232,Jun. 2012.
[2] M. M. E. A. Mahmoud and X. Shen, “A cloud-based scheme for protectingsource-location privacy against hotspot-locating attack in wirelesssensor networks," IEEE Trans. Parallel Distrib. Syst., vol. 23, no. 10,pp. 1805_1818, Oct. 2012.
[3] Q. Shen, X. Liang, X.Shen, X. Lin, and H. Y. Luo, “Exploiting geodistributedclouds for a e-health monitoring system with minimum servicedelay and privacy preservation," IEEE J. Biomed. Health Inform., vol. 18,no. 2, pp. 430_439, Mar. 2014.
[4] H. T. Dinh, C. Lee, D. Niyato, and P. Wang, “A survey of mobilecloud computing: Architecture, applications, and approaches," WirelessCommun. Mobile Comput., vol. 13, no. 18, pp. 1587_1611, Dec. 2013.
[5] W. Sun, et al., ``Privacy-preserving multi-keyword text search in the cloudsupporting similarity-based ranking," in Proc. 8th ACM SIGSACSymp.Inf., Comput.Commun.Secur., 2013, pp. 71_82.
[7] B.Wang, S.Yu,W. Lou, andY. T. Hou, “Privacy-preserving multi-keywordfuzzy search over encrypted data in the cloud," in Proc. IEEE INFOCOM,Apr./May 2014, pp. 2112_2120.
[8] E. Stefano, C. Papamanthou, and E. Shi, ``Practical dynamic searchableencryption with small leakage," in Proc. NDSS, Feb. 2014.
[9] Y. Yang, H. Li, W. Liu, H. Yang, and M. Wen,``Secure dynamic searchablesymmetric encryption with constant document update cost," in Proc.GLOBECOM, Anaheim, CA, USA, 2014.
[10]S.Saravanan,V.Venkatachalam,Improving map reduce task scheduling and micro-partitioning mechanism for mobile cloud multimedia services [J]. Int J of AdvIntell Paradigms ***8***(2), 156–167 (2016*)*

[11]S.Saravanan,V.Venkatachalam,"Advance Map Reduce Task Scheduling algorithm using mobile cloud multimedia services architecture" IEEE Digital Explore,pp21-25,2014.

[12]M.Murugesan,"Secure Data compression scheme in Cloud Environment with Backup Recovery Scheme"International Journal of Pure and Applied Mathematics,Feb 2018

[13]S. P. Yazhini, S. Santhiya, "Reliability and Confidentiality Based Data Storage in Cloud Using Merkle Hash Technique"International Journal of Pure and Applied Mathematics,Feb 2018

[14] G.Ranjith, M. Vinoth, "An Unified Approach for Effective Use of Cloud Metering Service"International Journal of Pure and Applied Mathematics,Feb 2018

[15] S.Saravanan, R. Bharathi, "Improved Management of Big Data Analytics in Map Reduce Data for Cloud Services"International Journal of Pure and Applied Mathematics,Feb 2018

[16]A.Selvanayagi,M.Sindhuja "Optimizing Cloud Gaming Experience through Map Reducing"International Journal of Pure and Applied Mathematics.

Fig: 1.Cloud Computing architecture

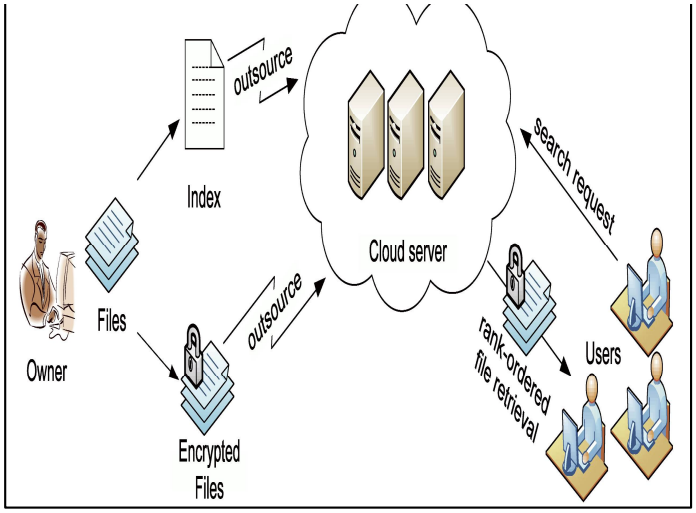


Fig: 2.Proposed system Architecture

Indian Journal of Natural Sciences www.tnsroindia.org.in ©IJONS

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Efficient Data Access Authority for Multi Data User Authentication Using Cloud Server

S.Keerthi*, B.Meiarasu, M.Kandasamy and M.Mohanraj

Department of Computer Science Engineering, M.Kumarasamy College of Engineering, Karur, TamilNadu,India.

Received: 09 Feb 2018 Revised:17 Feb 2018 Accepted: 30 Mar 2018

***Address for correspondence**
S.Keerthi
Computer Science Engineering,
M.Kumarasamy College of Engineering, Karur, TamilNadu,India

ABSTRACT

Appropriated storing, for example, Dropbox and Bitcasa is a champion among the most discernible cloud associations. Beginning at now, with that inevitability of adaptable flowed handling, clients can even supportively adjust the most current form of archives and synchronize the freshest records on their sharp cell phones. The use of information deduplication moreover help essentially diminish the data transmission and along these lines enhance the client encounter. A striking part of current appropriated accumulating is its in every practical sense unbounded cutoff. To enable boundless to confine, the passed on accumulating supplier client information deduplication systems to lessening the information to secure and diminish the utmost cost .paying little respect to the above purposes of intrigue, information deduplication has its natural security shortcomings. Among them, the most remarkable is that the adversary may have an unapproved chronicle downloading through the record hash only. Identify their execution deficiencies. By then we propose an elective format that accomplishes cloud server effectiveness and particularly remote suitability.

Keywords: Cloud server, synchronization, deduplication

INTRODUCTION

Cloud Computing Passed on storing up advantage is to overhaul their ability use. The reiteration of data grows rapidly and it will give a higher security, it will give a test to facilitate futile and repeated data made by various customers. Data deduplication is a framework used for taking out duplicate copies of data and cloud customer. It had been everything viewed as used as a touch of appropriated hoarding to lessen storage space and exchange transmission control. Business scattered farthest point relationship, for instance, Drop box, Mozy, Bitcasa have been applying deduplication to customer data to save upkeep cost. There are some wonderful bit of scattered gathering

can be seen. It has high openness, high versatility, and steady virtual storage space. High straightforwardness which proposes data reproduced over cloud server and customer need to get to their data what they require it take guaranteed to offer data to the customer. High flexibility which suggests that customers would bolster be able to whatever he/she should be exchanged to their cloud. An esteemed event of Bitcasa which sensitive "unending securing" that attracts the customer to exchange all around that really matters everything. Offering a boundless storage space may increase cash related weight on the scattered accumulating provider. The data deduplication system can diminish the cost securing.

Data deduplication is grabbed by abstaining from securing a comparable record grouped conditions. There are two sorts of data deduplication depending on where the deduplication happens. Server side data deduplication Client-side data deduplication Proof of Ownership.

Server side data deduplication:

In the wake of bearing the record server check first whether it formally appear in the limit. In the occasion that report is accessible then server discard the record and if it isn't then it make new record in the limit. We can see that server perform deduplication coming about to persevering through the record since it doesn't transmission limit saving.

Client side data deduplication

This side of deduplication get all the more capable technique. It takes a check make hash of report and send hash of record before it exchanging and it give send hash archives. In the wake of persevering hash it check away and hash is starting at now set away. Customer asked for to send nothing and customer interfaces the customer with the present write about the remote possibility that it exist away and customer asked for to exchange the record.

Confirmation of ownership

The intellection confirmation of ownership (PoW) is to deal with their worry of using a little hash an assistance as a judge for the entire record in client side deduplication and the foe could use the limit benefits as a substance appropriated manage for cloud customer. This attestation of part in PoW gives a response for guarantee the security in client side deduplication. Client can show to the server that it truly has a report.

Adaptable and Secure Sharing of Individual Thriving Records in Coursed Figuring Utilizing Trademark Based Encryption

A PHR advantage enables a patient to make, direct, and control her own specific flourishing information in a solitary place through the web, which has made the breaking point, recovery, and sharing of the remedial data more able. Particularly, every patient is guaranteed the full control of her healing records and can give her thriving information to an expansive combination of clients, including remedial organizations suppliers, relatives or companions. In light of the high cost of building and keeping up specific server develops, different PHR associations are outsourced to or given by outsider master groups, for instance, Microsoft HealthVault1. Beginning late, structures of securing PHRs in scattered handling have been proposed in. While it is enabling to have advantageous PHR associations for everybody, there are different security and protection dangers which could keep its wide assignment.

The fundamental concern is about whether the patients could really control the sharing of their delicate individual flourishing data (PHI), particularly when they are secured on an outsider server which individuals may not absolutely trust. The PHR proprietor herself should pick how to scramble her records and to permit which set of

clients to access each report. A PHR record ought to just be accessible to the clients who are given the taking a gander at translating key, while stay private to whatever is left of clients.

Ciphertext-System Trademark Based Encryption

In two or three scattered structures a client ought to just be able to get to information if a client can a specific course of action of accreditations or attributes.The technique for completing methodologies is to utilize a trusted server to store the information and intercede find the opportunity to control. The server securing the information is traded off, by then the request of the information will be dealt. The framework for perceiving complex access control on encrypted.Our strategies blended information can be kept private paying little regard to whether the farthest point server is to structures utilized credits to outline the encoded information and joined courses of action with client's keys; while in our framework ascribes are utilized to delineate a client's accreditations, and a party scrambling information picks a strategy for who can unscramble. In this way, our frameworks are carefully nearer to customary access control strategies, for example, part based access control (RBAC). Furthermore, we give an utilization of our framework and give execution estimations.

Totally Secure Useful Encryption Internal Thing Encryption

In this paper, we present two totally secure helpful encryption designs. Our first result is a totally secure attribute based encryption (ABE) plot. Past improvements of ABE were simply wound up being particularly secure. We achieve full security by changing the twofold structure encryption methodology starting late exhibited by Waters and previously used to get totally secure IBE and HIBE systems. We can use a novel information theoretic dispute to alter the twofold system encryption procedure to the more perplexed structure of ABE systems. Security is shown under a non-natural doubt whose size does not depend upon the amount of request. The arrangement is comparably powerful to existing particularly secure plans. They moreover present a totally secure different leveled PE plot under a comparative doubt. The key and for bilinear pairings using the possibility of twofold mixing vector spaces (DPVS) proposed by Okamoto and Takashima.

Capable Property Based With Repudiation for Outsourced Data Sharing Control

Figure content Technique Property Based Encryption (CP-ABE) is a promising cryptographic rough for fine-grained get the opportunity to control of shared data. In any case, when CP-ABE is used to control outsourced data sharing, it confronts two obstructions. Immediately, the data proprietor must trust in the properties expert, besides, the issue of value denial of CP-ABE designs, which encounters such issues as different granularities of disavowal, poor versatility and high computational unpredictability, is inconvenient. In this paper, we propose another CP-ABE contrive that the data proprietors can totally control their outsourced shared data.System furthermore resolve the issue of disavowal including the entire customer get the opportunity to profit and just inadequate access right of the CP-ABE with the passageway control of structure. In addition, the data proprietors and the attributes master can dole out most of troublesome errands to foreswearing mediator detaches with the technique for delegate re-encryption.

Progressions in Estimation and Control Cloud-Enabled Auto Vehicle

Conveyed figuring is disquieting access to flowed information and enrolling resources that can support future data and count genuine vehicular control works and improve vehicle driving comfort and security. This paper explores a couple of potential Vehicle-to-Cloud-to-Vehicle (V2C2V) applications that can overhaul vehicle control .This information can be granted to various vehicles and transportation pros inside a V2C2V framework. The response of hitting a pothole is portrayed by a multi-organize dynamic model which is affirmed by differentiating reenactment

occurs and a higher-consistency business exhibiting group. A novel structure of synchronous road profile estimation and irregularity ID is created by combining a bounce scattering process (JDP)- based estimator and a multiinput observer. The execution of this arrangement is surveyed in a trial vehicle. Furthermore, another gathering computation is delivered to pack variation from the norm information by getting ready irregularity report streams. Additionally, a cloud-upheld semi-dynamic suspension control issue is mulled over appearing out of nowhere that road profile information and hullabaloo estimations from the cloud can be used to update suspension control. The issue of picking a perfect xv damping mode from a constrained course of action of damping modes is seen as and the best mode is picked in light of execution desire on the cloud. Finally, a cloud-helped multi-metric course coordinator is investigated in which security and comfort estimations amplify standard orchestrating estimations, for instance, time, detachment, and productivity. The prosperity metric is made by taking care of masterminding computation can be realized on the cloud to comprehend the multi-metric course organizing.

Summary

[1]Public key based 3-DES and RSA counts is done.RSA disentangles the burden of the key assention and key exchange issue [2].Cipher content system attribute based encryption (CP-ABE) is a promising cryptographic rough for fine-grained get the chance to control of shared data.System and just most of the way get to right of the users[3]Realizing complex access control on encoded data that we call figure content course of action property based encryption.[4]Cloud handling is changing access to appropriated information and improve vehicle driving comfort and safety.[5]Prior trademark based encryption structures achieved understanding resistance.[6]Cloud enlisting is a progressing perspective. The NIST definition depicts basic parts of disseminated processing and is proposed to fill in as a techniques for broad.[7]A PHR advantage empowers a patient to make, direct, and control her own particular prosperity data in a single place through the web, which has made the limit, recuperation, and sharing of the therapeutic information more efficient.[8]a customer should simply have the ability to get to data if a customer powers a particular game plan of capabilities or attributes.[9] totally secure trademark based encryption (ABE) plot. Past advancements of ABE were simply ended up being particularly secure to gain totally secure IBE and HIBE systems.[10]Cloud enlisting is transforming access to passed on information and figuring resources that can support future data and computation genuine vehicular control works and improve vehicle driving comfort and safety.[11]Multi-Master Trademark Based Encryption (ABE) structure.

CONCLUSION

We proposed a revocable multi-master CP-ABE plot that can support compelling property denial. By then, we built up an effective data get the chance to control contrive for multi-master circulated capacity systems. We moreover showed that our arrangement was provable secure in the subjective prophet show. The revocable multi-master CP-ABE is a promising method, which can be associated in any remote accumulating structures and online casual groups. Single master CP-ABE where all qualities are managed by a lone authority, and multi-pro CP-ABE, where properties are from different spaces and directed by different specialists.

REFERENCES

1. B.Waters, „„Ciphertext-Policy Attribute-Based Encryption: An Expressive, Efficient, and Secure Realization,"" in Proc. 4th Int"l Conf. Practice and Theory in Public Key Cryptography (PKC"11), 2011, pp.53-70.
2. A.B. Lewko, T. Okamoto, A. Sahai, K. Takashima, and B.Waters, „„Fully SecureFunctionalEncryption: Attribute-Based Encryption and (Hierarchical) Inner Product
3. Encryption,"" in Proc. Advances in Cryptology-EUROCRYPT"10, 2010, pp.62-91.

4. M. Chase, „„Multi-Authority Attribute Based Encryption,"" in Proc. 4th Theory of Cryptography Conf. Theory of Cryptography (TCC"07), 2007, pp. 515-534.
5. M. Chase and S.S.M. Chow, „„Improving Privacy and Security in Multi-AuthorityAttribute-Based Encryption,"" in Proc. 16th ACM Conf. Computer and Comm.
6. Security (CCS"09), 2009, pp. 121-130.
7. A.B. Lewko and B. Waters, „„Decentralizing Attribute-Based Encryption,"" in Proc.Advances in Cryptology-EUROCRYPT"11, 2011, pp. 568-588

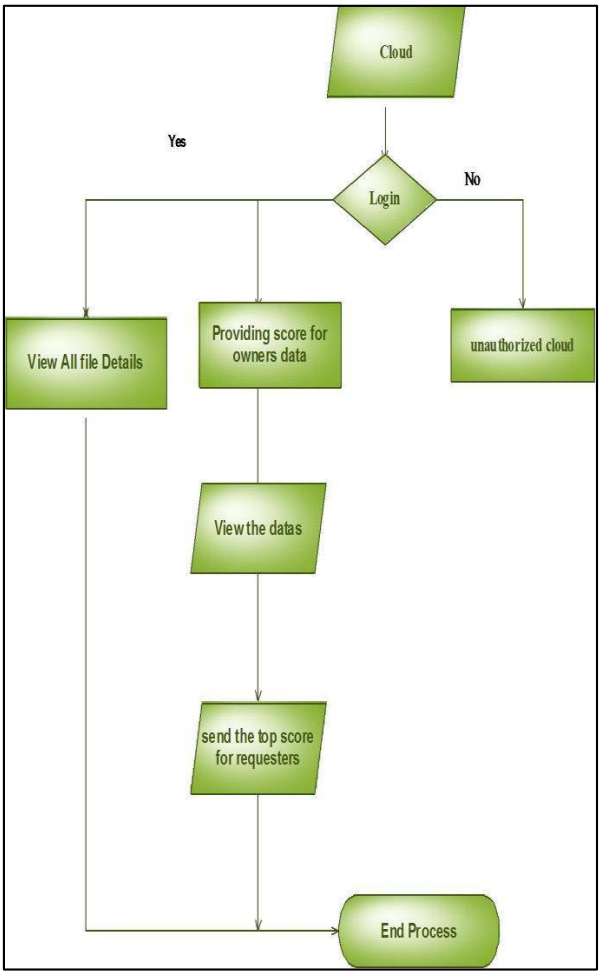


Fig.1: System Architecture

Indian Journal of Natural Sciences www.tnsroindia.org.in ©IJONS

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Enhancing Network Lifetime Using Cluster-based Methodology in Wireless Sensor Networks

Devaki R* and Vijaymeena M.K

Department of Computer Science and Engineering, M.Kumarasamy College of Engineering (Autonomous),TamilNadu,India.

Received: 12 Feb 2018 Revised:26 Feb 2018 Accepted: 30 Mar 2018

***Address for correspondence**
Devaki R,
Assistant Professor,
Department of Computer Science and Engineering
M.Kumarasamy College of Engineering (Autonomous), TamilNadu,India.

ABSTRACT

In the recent researches wireless Sensor network has got a considerable attention because of its to its diverse applications. Sensors are small and portable device that works on small batteries which has limited lifespan so there occurs a situation to replace them which is difficult and costly. Low battery capacity and instable power supply of the sensor nodes are the major drawback that has to be considered. Therefore minimization of energy utilization is necessary to increase the network lifetime to provide efficient communication. So, various techniques are deployed for minimizing the energy of the sensor nodes. In this paper we have proposed an idea about a cluster based methodology to increase the network lifetime by minimizing the energy utilization which greatly increases the network lifetime.

Keywords: Sensor nodes, Low battery capacity, minimization of energy utilization, network lifetime, cluster based methodology.

INTRODUCTION

Wireless Sensor Networks (WSN) has autonomous sensor nodes that are in a spatially distributed. The sensor nodes are portable, smaller in size and are set up to observe a particular target and transmit their data to the receivers. WSNs plays a vital role in research as they have wider range of applications like traffic monitoring, video surveillance, robot control, air traffic control, area monitoring, forest fire detection, medical device monitoring, weather monitoring, industry process monitoring, automated and smart home monitoring, etc. Usually the sensor nodes equipped with several parts like transceiver – which connects to the external antenna with an internal antenna, microcontroller – used as an interface between the sensor and power source, and power source usually a battery. Energy consumption is the major concern for wireless sensor networks as there are only limited energy resource at every node. The node lifetime extends until the power source is non empty. And the network time is the lifetime of

the living nodes in a Wireless Sensor Network. So it is very obvious that the network lifetime greatly depends on the node lifetime In order to increase the network lifetime the necessity to reduce the energy consumption at the sensor nodes in the network arises. In this paper a discussion about various methodologies that have been used for energy minimization at the nodes and maximization of the network lifetime are made and the efficiency of the cluster based energy minimization technique which is the proposed methodology is depicted.

LITERATURE REVIEW

Auction Based WSN

This scheme is based on the auction, where the one who has the higher cost can participate always. Similarly here the node that has higher broadcast result than the other can participate. In this scheme the nodes can save their energy during broadcast. Time Division Multiple Access scheme is followed by the nodes to broadcast the results. Every node will take its chance to broadcast the result at its time slot. Before a node is ready to broadcast it senses the result of the previous node. If the result of the previous node is lower than itself then the node is ready to broadcast the result otherwise it remains idle during its time slot and wait for the second chance. BR_1 and BR_2 are the results broadcasted by node 1 and node 2 respectively. Node 1 will broadcast its result only if it is higher than node 2 ($BR_1 > BR_2$).

If we consider the binary indicator for the result broadcast of each node as B_n, then B_n can be expressed as,

$$B_n = \begin{cases} 1, if\ n = 0 \\ 1, BR_n > \max\ (BR_1 BR_2, \ldots\ldots, BR_{n-1}) \\ 0, Otherwise \end{cases} \quad \rightarrow \text{Eqn (2.1)}$$

Thus through this methodology the life of the network will be extended by minimizing the energy consumption at the node level. Advantages: This methodology implies that there is only minimal and insignificant performance loss with maximum saving of energy. Disadvantages: There is a complexity involved in this scheme when trying to reduce the performance loss.

Cooperative Transmissions in WSN

Let us consider a wireless sensor network in which the source node sends the information to all the other nodes in the network in a multi-hop fashion. The architecture of this transmission is point-to-point links where every receiver is in connection with the sender i.e the source node. The reliability of the link relies on the energy emitted by the source node. Here individual cost is assigned to each point-to-point links which is the referred as the transmitter-receiver link's transmission energy. This clearly shows that the methodology is inefficient. The Wireless Multicast Advantage (WMA) method provides a solution for this problem. It says that the energy can be saved by transmitting to multiple receivers simultaneously at a single cost so the total number of transmissions is greatly reduced. However if the receiver is not in the transmission range of the network then the energy received by that receiver will not be reliable for detection. It leads to inefficient use of energy. To overcome this, "Cooperative Wireless Advantage" (CWA) is employed. The residual energy of the transmitter is combined with all the near-by transmitters. Now the transmission is done by the cooperation of multiple relay nodes.

Advantages: The reliability of the detection of signal at the receiver node can be improved along with increased energy-savings.

Disadvantages: The task of assigning optimum energy is NP-Complete, so high computational complexity persists.

Balanced Energy Consumption in WSN Using Genetic Algorithm

In WSN network life time can be improved by clustering models. In this strategy sensor clustering method to dynamically organize heterogeneous WSN using Genetic Algorithm (GA) is proposed. Here a framework is provided to form a dynamic network structure by incorporating the clustering and heterogeneity factor like expected energy expenditure, remaining energy, network locality and distance to the base-station. In order to ensure network integrity, heterogeneity factors are integrated as constraints to chromosomes, and a validation process is performed. This method is 2-fold. First, each sensor node should provide an estimation of energy state if a network clustering structure is formed in the next round. Second, a fitness function is employed by multiple aspects of the heterogeneous WSN through an optimization method that is based on GA.

Advantages: Optimized network lifetime is provided even in dynamic network structure.

Disadvantages: Still there resides wastage of power when the sensor nodes are continuously listening for a signal.

Cross Layer Design for Network Lifetime Maximization in WSN

Cross-layer approach aims to reduce the energy consumption of WSNs by joint designs across layers. In this paper using cross layer design the network lifetime maximization can be done at physical layer - power control, MAC layer - link schedule and network layer- routing. Since TDMA approach is used the WSNs are classified as single-source and single-sink (SSSS) networks and multi-source and single-sink (MSSS) networks. Many sensors generate data to send in MSSS networks whereas in SSSS networks there is only one sensor to generate the data and send. An analytical expression of optimal network lifetime is obtained using KTT conditions for linear MSSS and planar SSSS networks. But deriving analytical expression for diverse data rates and planar MSSS topology is a complex task. In order to overcome this, decomposition and combination approach which is an iterative algorithm is incorporated to obtain a suboptimal network lifetime. Here using this approach MSSS networks are decomposed into SSSS networks and the optimization problem for SSSS that is decomposed are solved. Then they are combined to produce a suboptimal solution for MSSS network.

Advantages: Distributed and low complexity solution is arrived. Heterogeneous energy supply is provided in order to avoid instable energy supply.

Disadvantage: Different nodes will need different energy source due to varied workload.

PROPOSED IDEA

We proposed an idea of a cluster based methodology to minimize the energy utilization at the nodes in the wireless sensor networks. The proposed idea has two stage processes: 1. Selection of Wake-Up Call Node 2. Power Minimization at the transceiver.

SELECTION OF WAKE-UP CALL NODE

In the traditional point to point architecture there is always inefficient power management.The cost of transmission of information is very high in the point to point architecture. So to avoid this problem clustering methods have been employed which greatly reduced the cost of the links (Fig.1).

The protocol is as followed

Step 1: First the sensor nodes are clustered as five categories normal nodes, intermediate nodes, advanced nodes, super nodes and ultra nodes using K-means clustering algorithm based on their energy levels.

Step 2: Then in each cluster the chief node is identified. These chief nodes communicate with the other chief nodes and the data packets are forwarded to the base station.

Step 3: The probability of the nodes to turn into a chief node depending on their energy level is calculated using Stable Election Protocol (SEP) [5]:

Let the total number of nodes be n, m is considered as the fraction of total number of nodes, which are deployed with k times more energy than the other nodes [6]. These nodes are considered as command nodes and the other nodes are common nodes.

Now, the probability of common node to become the chief node is calculated as,

$$P_{cm} - \frac{P_o}{1+m.} \quad \rightarrow \text{Eqn (3.1)}$$

The probability of command node to become the chief node is calculated as,

$$P_{cd} = \frac{P_o}{1+m.k}(1 + \quad \rightarrow \text{Eqn (3.2)}$$

Where, Po is the optimal probability of each node to become chief node in the network. In SEP strategy, selection of chief node is done randomly on probability basis for each node.

Step 4: The node that has the highest P_{cd} is selected as the wake-up call node. When the current chief node changes to P_{cd} then again step 3 is repeated to select the wake-up call node.

Power Minimization at the Transceiver

The maximum energy is consumed by the transceiver. Though most of the time the transmitter is in idle state in transceiver mode there can be still energy wastage in the receiver mode as the radio has to be in ON state to detect the receiving signals. To avoid this transceiver power management is done. The node that has the highest P_{cd} acts as the wake-up call to the other nodes in the cluster. This node sends a wake-up call signal to a particular node when it has to receive the message from any transmitter. So it portrays that the node does not want to keep the radio in ON state always it is sufficient to turn on the radio only during its message reception alone. Thus the energy wastage at the node can be minimized.

CONCLUSION

The management of energy resource is very much essential in wireless sensor networks for reliable communication. The perceptual operation of wireless sensor networks has to be ensured. The proposed idea is Clustering based scheme for the transmission and reception of the message. The motive of this technique contributes to increase the network lifetime by reducing the consumption of the energy at the sensor nodes in the network. The proposal of wake-up call node greatly reduced the energy utilization of senor nodes. Thus the proposed idea could greatly assist in enhancing the lifetime of the sensor nodes in the wireless sensor networks. Based on the various literature surveys this ideology had been arrived and it is yet to be implemented with OPNET simulation tool.

REFERENCES

1) Saud Althunibat, Amir Abu-Al-Aish, Wael F. Abu Shehab, Wael H. Alsawalmeh "Auction-based Data Gathering Scheme for Wireless Sensor Networks", IEEE Commun. Lett., 1089-7798 2015.
2) W. Xu, et al., "Energy Management and Cross Layer Optimization for Wireless Sensor Network Powered by Heterogeneous Energy Sources", IEEE Trans. Wireless Commun., 14.5, 2015, 2814-2826.
3) M. Elhoseny, et al., "Balancing Energy Consumption in Heterogeneous Wireless Sensor Networks using Genetic Algorithm", IEEE Commun. Lett., 19.12, 2015, 2194-2197.
4) YW Hong, and A. Scaglione, "Energy-efficient broadcasting with cooperative transmissions in wireless sensor networks", IEEE Trans. Wireless Commun., 5.10 (2006): 2844-2855.
5) Yogesh Mishra, Ashish Singhadia, Rashmi Pandey, "Energy Level Based Stable Election Protocol in Wireless Sensor Network" International Journal of Engineering Trends and Technology (IJETT) – Volume 17 Number 1 – Nov 2014.
6) Saini, Parul, and Ajay.K. Sharma. "E-DEECEnhanced Distributed Energy Efficient Clustering scheme for heterogeneous WSN", 2010 First International Conference On Parallel Distributed and Grid Computing (PDGC 2010), 2010. Hui Wang. "Network lifetime maximization with cross-layer design in wireless sensor networks", IEEE Transactions on Wireless Communications, 10/2008

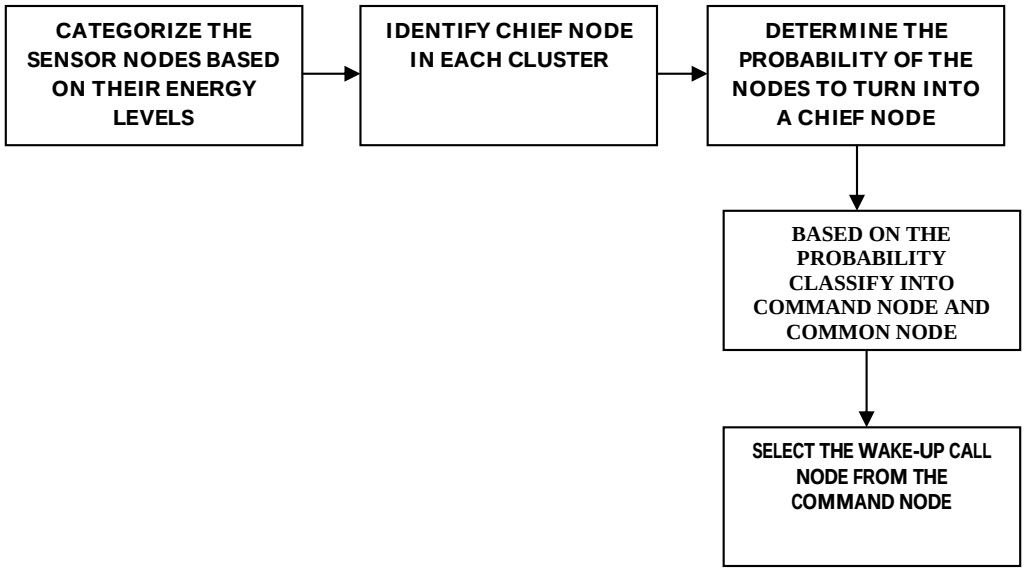


Fig.1.Selection of Wake-Up Call Node

Indian Journal of Natural Sciences www.tnsroindia.org.in ©IJONS

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Autonet Firewall for Cloud and Real Time-Server

C.Selvarathi*, S.Gokulkishan, M.Geethapriya,K.Dhivyadharshini and R.Hariharan

Department of Computer Science and Engineering, M. Kumarasamy College of Engineering (Autonomous), Thalavapalayam, Karur, TamilNadu,India.

Received: 10 Feb 2018 Revised:17 Feb 2018 Accepted: 30 Mar 2018

***Address for correspondence**
C.Selvarathi
Assistant Professor, Department of Computer Science and Engineering
M. Kumarasamy College of Engineering (Autonomous),
Thalavapalayam, Karur, TamilNadu
E.mail: selvarathic.cse@mkce.ac.in

ABSTRACT

The state which is used to manage the security issues in the rate of the common server in interface or in institution is very expensive to maintain that value. So,we have made to analyze the open ports to configuration the rate of concern to act at this point.And it was regional use for the conditions to highly defined at a rate to act. When the open ports are get analyzed it will used at the point to rectify the rate of bugs and malware. As by this way the firewall getanalyze automatically and used as by the work was assigned. This will get used for the system to update it automatically at the rate of use and its system to functioned at the system. At per the rate it will get used and maintain the work at your point to conditioned at the levels of the system. Firewalls involve some technology to get functioned. Specifying the filtering rules as a policy. High-level languages are used to defining correctly a firewall's policy. Once a policy has been specified, testing is required to determine if the firewall correctly. And the conservation of the systematic varied at your part to assigned. As by the system to get maintained by the rate to activate and for the admin to activate to analyze automatically and to concern as it's to conditioned.This will activate the rate to activate to analyze as the system to conditioned at the rate to concerned at the part.

Key words: security issues, Firewalls, configuration.

INTRODUCTION

By the work base of the elastic firewall we have make it through the condition to automation in the system. And it was get used some open source projects to implement the analysis as complex work in the system. In [5],[6],a firewall is perfect cluster to work on the regional security of the companies cloud based technology or the conditioned to activate to maintain. By the behavior of the firewall it may get used for the malware detection to contains the system

to conditioned to activate to system to drive at your rate of securing methods. In [7], is clearly proposed to secure the highly confidential part to keep very secure in the server. This may get represented by the Authors to designed by the various algorithms and systematically. In [8], [9], it was trended with the future use for securing the firewall in the interface . In [10], a hybrid cloud-based firewalling service was proposed. It was get proposed for the highly valid and performing server to act at a rated to use. Once the access get cross the limit it will automatically balance it to the web server to get activate the source and for the automatic update of the system to analyze at conditions. This kind of an activity is managed to condensed to prevent the DDOS attack for the traffic limits in the server. Though it gets works with scalable and reliable it will parse each and every packet by using the python modules. And it was get code for the load balancing the data and access the condition for the system, as by the way of the use it may use and this can used at our part to contain in a container. And it will ready to get used to test activity for the source that have given by the used system. By the ways of key rolein creating, monitoring, managing, and orchestrating vow instances in the cloud had violated in the manual analysis. This significance of the workload of the LB, especially when the high workload or an access coming into the server to act. And the rated way concern and to activate the condition to maintaininto with an activity of the harmful significance to get rid of the data to get acted byits own up.

Analytical Model

In this section, we develop an analytical model to capture the service behavior within a VFW, and then we derive formulas to estimate the response time. Typically, for a Paused or virtual firewall, and as shown in Figure 4, insemination to analyzed its work by its own set of key valid and parsing VPC concern at the server. This packetis queued into a Rx DMA ring and then go into three stages of service. The packets are get to the server and parse the condensed by the headers and the packets get opened by the wire shark to analyze. Then it will get assigned to terminal system to check the vulnerability. Firewallrule base interrogation takes place at the server. Step2 it will get checked for each and every packet that get received by the server one by one for analyzing. Step3 it will have checked and it will returns the status to maintenance system to get occur at the product of the rated checkups. By the condition of three cluster of the process it will get used and this can get parsed for each and every packet to analyzed at a rate to condensed the part to consumed at the system to condensed at the rated way to activate at your part. By this execution we can clearly get the result to get used at the launch of the project at the rated way at the system to conserved at the way of use to maintainat the system to reginal condition. As this it will get rated at your part to contain to get stayed at your part. This will get stayed at parsing of the packets to the systematic point at the condition to packet regulation at some intervals to condensed uses. The behavior of firewall processing of incoming packets can be modeled as a finite queueing system size with three stages of service. As shown in Figure 5, an incoming network packet gets first queued in a buffer of size K-1 and then gets served sequentially in three stages with each stage having a different mean service rate, i.e., μ1, μ2, μ3. We assume that incoming packets follow a Poisson arrival λ. Also, the stage service times are independent with exponential distribution.

Network and Service Management

Analysis was get used by the method of single Firewall as by process as FCFS (First Come First Served). In Stage 2 the processing way of the mean calculation is gets depends on the service the time is calculated by μ2 and this rule of triggered. On average, 1/μ2 can be expressed as follows1/μ2 = L · TR, and L for the average number of rules to derogate at theusual traffic that et incomed to the server, and TR for the average interrogated to time per rule. Figure 5 __ At the model of the calculation it will get used and it will have maintained for the poisoned packet arrival at the rate to act at your part to maintained at our part to concern as your part. Exponentially this will used at our to considered poisoned packet to fix and retained the status to activate the rate of the use to concern but in some other case it will collapsed at thepath of the interval of time. [13]– [15]. This service get retained for all the packets that's how get traffic to the server to maintain at the level to act at your part to concern to the server .And moreover it may get waited for the parsing the packets to get survived by the server to condition by the system. As by the use of the it

will have stayed by with the system to condensed for the waved at your part to activated at your part. This will get stayed at parsing of the packets to the systematic point at the condition to packet regulation at some intervals to condensed uses. The behavior of firewall processing of incoming packetscan be modeled as a finite queueing system size with three stages of service [5], [6] . Exponentially this will used at our to considered poisoned packet to fix and retained the status to activate the rate of the use to concern but in some other case it will collapsed at the path of the interval of time.incomingdata to analyzed its work by its own set of key valid and parsing VPC concern at the server.Thispacketis queued into a Rx DMA ring and then go into three stages of service. The packets are get to the server and parse the condensed by the headers andthe packets get opened by the Wireshark to analyze. This will get stayed at parsing of the packets to the systematic point at the condition to packet regulation at some intervals to condensed uses.The behavior of firewall processing of incoming packets can be modeled as a finite queueing system size with three stages of service. As shown in Figure 5, an incoming network packet gets first queued in a buffer of size K-1 and then gets served sequentially in three stages with each stage having a different mean service rate.Our analytical model is built on the principles of the embedded Markov chain with a finite state space. The model captures the behavior of the vFirewall processing with a state space $S = \{(k, n), 0 \le k \le K, 0 \le n \le 3\}$, where k denotes the number of packets in the entire system, and n denotes the service stage number being performed by the CPU. The queuing system has a buffer size of K-1. In other words, state (0,0) represents the special case when the system is empty or idle, i.e. the state of system idleness. States (k,n) represent the states where the CPU is busy executing service of stage n with k packets in the system.(8) Please note that pK,1 can be derived from either Equation (9) or (10). Both of these equations are numerically equivalent. Now, p0,0 can be obtained using the normalization condition .All state probabilities $\{pk,n; 1 \le k \le K, 1 \le n \le 3\}$ can be computed recursively using Equations (11), as shown in Algorithm 1, which can be converted easily to a MATLAB or another similar package script. As shown, the algorithm first computes the loop invariants (C1 to C8) in Line 05, and then uses the Equations (11-12) to determine all state probabilities.Algorithm 1 Computing steady-state probabilities.

RESULTS AND DISCUSSION

By use of the section it may get used to activate for the analytical module to get stayed by the covered by the conditioned for the system to maintained for the used way to assign to concern at the type to condensed at your part to asserted at a part to validate. As by the way of the use it will get stayed by the system to activate at a part to get systematic conversion at a level of the uses and it's type.

Validation from the mean calculation:

By representing the status of the analysis of the parsing will get used at our part to maintain at the level to activate at your part to concern at your part to analyze.This will get stayed at your part to get systematic way of use and its type to assign at the work. And it will get used to plot by the graph using the MATPLOTLIB in python. The simulation results were obtained using a discrete-event simulationwritten in C. Details on how to develop a DES simulator in C can be found in [13]. There are a number of publically and commercially available network simulation tools. Some of these simulators are designed specifically for cloud environments (e.g., Clouds, encloud, EMUSIM, MDCSIM), and some are generic in natures (e.g., OPNET, NS, OMNET,J-Sim, JMT). All this behavior are get considered as the internet behavior to analytic way to activate at your part to concern at your part to assign as a part to valid at your part of the concern as the way of the use and it's type to act of interrogating rules in the mean calculation in the system to analyzed carefully to check at your status part of the conditioned by the rate of use tomaintained by its own stage and it's sustainable work to act at your part.By this case it will get used at your part to act to interval part of the system. (i.e. M/M/1/K). Since,this simulationis followed by the principles of calculating the content to manage the system as per the interval of the mean condition.And this will get analyze assumed odd and even ports to get manage at a server.This will get stayed at our part to assert the port analysis to parse the packets for malware and regional supports for the server authentication. Thus we have represented in the average time calculation it get used

at our point to act at the similarities to save the sessions to have a rated at the virtual interface.As by the own systematic variation we have get used at our ansible for automatic track and used at our serve .For the simulations and rated way to maintain the packets to get assigned in the server. K=300 packets, 1/μ1 = 5.3μs, 1/μ3 = 200μs, TR = 0.1μs and 1/μLB = 100μs .

CONCLUSION

This article defined to parse the packets and secure the ports automatically for the reports.And it get regionally to used at our part to assign at your point to maintain the system to analyze at this moments for the declaration.This will get stayed at our point to maintain the regulatory state.At a server ports are get eventually get used to parse and it will used at the interface to assign and this can maintain at the virtual interface to check the code to assign and it may get used at our part to maintain by own use as it's type of the consideration of the use and its type to act.Then the automating purpose it will get stayed for the port parsing and get way at our part to use in this interface.

ACKNOWLEDGMENTS

The authors thank the anonymousreviewers for their valuable comments, which helped us to considerably improve the content, quality, and presentation of this article. The experimental work in this paper was supported by a generous research award under the Amazon Web Services.Program [37].By this resolving the part the data get used and waved for the system to maintain it automatically in the real time serving type. And to parsing the setup we can used this system for the conditional cloud and systematic variation in the institution.

REFERENCES

[1] Writing temporally predictable code P. Puschner; A. Burns.
[2] FTS: a high-performance CORBA fault-tolerance service R. Friedman; E. Hadad.
[3] Operating system performance in support of real-time middleware D.C. Schmidt; M. Deshpande; C. O'Ryan.
[4] A scheduling technique providing a strict isolation of real-time threads U. Brinkschulte; J. Kreuzinger; M. Pfeffer; T. Ungerer
[5] Role-based access control for CORBA distributed object systems R.R. Obelheiro; J.S. Fraga
[6] Dependable Semantic Web B. Thuraisingham; E. Hughes; D. Allen
[7] Worst case execution time analysis of object-oriented programs J. Gustafsson.
[8] S. Kamara and C. Papamanthou, "Parallel and dynamic searchable symmetric encryption," in FC'13, ser. LNCS, 2013, vol. 7859, pp. 258– 274S.
[9] W. Sun, B. Wang, N. Cao, M. Li, W. Lou, Y. T. Hou, and H. Li, "Privacy- preserving multi-keyword text search in the cloud supporting similarity- based ranking," in ASIACCS'13, 2013, pp. 71– 82.
[10] Z. Brakerski and V. Vaikuntanathan, "Fully homomorphic encryption from ring-lwe and security for key dependent messages," in CRYPTO'11, 2011, pp. 505–524.
[11] M.Yoshino,N.Kunihiro,K.Naganuma,andH.Sato,"Symmetric inner- product predicate encryption based on three groups," in ProvSec'12, ser. LNCS, 2012, vol. 7496, pp. 215–234.
[12]S.P.Yazhini,S.Santhiya"Reliability and Confidentiality Based Data Storage in Cloud Using Merkle Hash Tree Technique",International Journal of Pure and Applied Mathematics,Feb. 2018,Vol.118 , No.8,pp.793-797
[13]M.Murugesan,"Secure Data Compression Scheme in Cloud Environments With Backup Recovery cheme",International Journal of Pure and Applied Mathematics,Feb. 2018,Vol.118 , No.8,pp.467-471

Indian Journal of Natural Sciences www.tnsroindia.org.in ©IJONS

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Language Translator for Images

P.RajeshKanna[1]*, S.Abirame[2] and M.Madura[2]

[1]Assistant Professor,Department of Computer Science and Engineering, M.Kumarasamy College of Engineering, Karur, TamilNadu,India.
[2]UG Student, Department of Computer Science and Engineering, M.Kumarasamy College of Engineering, Karur, TamilNadu,India.

Received: 10 Feb 2018 Revised:16 Feb 2018 Accepted: 30 Mar 2018

***Address for correspondence**
P.RajeshKanna,
Assistant Professor,
Department of Computer Science and Engineering
M.Kumarasamy College of Engineering, Karur, TamilNadu,India.
E mail: rajeshkanna@gmail.com

ABSTRACT

A content recognizable proof and interpretation application on Android stage. This application perceives the content that is caught by a cell phone camera and deciphers the content lastly shows back the perceived content alongside the interpretation on to the screen. To build up this application we have utilized the Optical Character Recognition, OCR motor, Matlab, Google decipher API and we build up our own particular open source Android application. The goal of this task is to help the voyagers explore while they are meandering around in abroad. To get our objective we build up an application in view of the portable camera which can have the capacity to recognize English content at word level and make an interpretation of it into English Tamil In this report, we exhibit the framework stream, the content discovery calculation and itemized explore result.

Keywords: Optical character recognition, Android, Unicode text, Binarization

INTRODUCTION

Text extraction is the very crucial stage of evaluating the images. The steps involved are location, restriction, binarization, extraction, improvement, and acknowledgment of content from the picture. The image contentin the imageisgenerally classified into two categories i.e perceptual substance and semantic substance. Perceptual contents involvecolours, textures, intensities, shapesand. their fleeting changes. Semanticcontents involveobjects, occasions, and their relationsbetween them Text content include huge amount of semantic information. Todaytext extractiontechnique from images is very essential in content evaluation. This technique can be used in a variety of applicationssuch as image searching, indexing, navigation, and human computer interaction.A text information extraction system receives a still image or a sequelof imagesin the form of input. Theimages can be in gray scale

ormay becolored, compressed or uncompressed the content extraction systemcan besubdivided into the accompanying issues i.e. location, limitation, tracking,extraction and improvement and recognition.

Properties of Text in Images

Text usually has variousforms due to changes in fonts,size, style, introduction, arrangement, surface, shading, complexity and foundation. A portion of the attributes are as per the following.

Size

The size of text may varya lotin different images

Alignment

Text may be alignin anyof the direction and may containsome of the distortionsand disturbances such as geometric distortion

Colour

The textcharacters may have same or similar color

Edge

Most of the images have strong edges at the boundaries of text and background

Compression

Some of the images are recorded, transferred, and processed in compressed format.

METHDOLOGY

Wepropose a content location/acknowledgment/interpretation calculation that comprises of following advances
1) Morphological edge recognition
2)Text highlight seperating
3) Text region binarization
4) Optical character recognition
5) Text correction
6) Text translation
7) Display of the translation

Morphological Edge Recognition

Morphological Edge Detection play out the edge discovery calculation, we first change over the info RGB shading picture to a dim scale power picture Y utilizing (1), where R, G, and B speak to red, green and blue parts of the information picture Y= 0.299R+ 0.587G+ 0.114B.

Text Highlight Separating

The dim scale picture is double picture utilizing open-shut and close-open channels to diminish false edge commotion and over-division. Molding component utilized for this task is a 3 by 3 8-associated component. At that point, a morphological angle acknowledgment activity is accomplished on the obscured picture.

Text Region Binarization

Keeping in mind the end goal to diminish the quantity of associated parts that must be dissected; a nearby task with a 5 by 5 organizing component is performed to the paired edge picture got from morphological edge recognition.After the nearby activity, every single associated segment of the edge picture are screened with their position, size, and region data.

Optical Character Recognition

Each staying skipping box is utilized as a cover to the first dim scale picture. Otsu's technique is utilized to acquire the limit of the veiled dim scale picture for binarization. Since each bobbing box is moderately little contrasted with the measure of the whole picture, no further versatile edge strategy is executed. Hypothetically after this progression, just praised letters are left as the frontal area, and whatever is left of the picture would go to foundation Morphological edge recognition.Again the same process will be repeated from Optical character recognition to Text translation.

Text Recognition, Correction, and TranslationSince the project is focused on realizing textextraction on a mobile phone, we implemented thefollowing three steps - text recognition, correction andtranslation on a server with open source software forsimplicity's sake. Google's open source OCR – Tesseractis used as the optical text recognition engine. PeterNorvig'salgorithm is added to the routine to performtext correction. Then Google translator is used totranslate the text into Chinese

Dispay of the Translation

Presentation of the Conversionand converted text string from text translation is sent back tothe mobile device (Android phone) from the server, andthen displayed at the top center region of the screen. Asample text abstraction process flow

Algorithm for Segmentation of Marathi Script

Line Segmentation

The steps for line segmentation areas follow:
Construct the Horizontal Histogram for the image.
Using the Histogram, find the points from which the line starts and ends or a line of content, upper line is drawn at a point where we begin discovering dark pixels and lower line is drawn where we begin discovering nonattendance of dark pixels. Furthermore, the procedure proceeds for next line and so on.

Word Segmentation

Construct the vertical histogram for each segmented line using the vertical Histogram, find the points from which the word starts and ends.

Vertical lines are drawn at beginning and closure focuses for each word.

Character Segmentation

Draw the horizontal histogram for each segmented line.
From the horizontal histogram, find the row which consists of maximum value.
The row which consists of maximum value of black pixel for each line is actually the row which consists of Header line. Draw the vertical histogram for each fragmented word in underneath of header line. Draw the vertical histogram for each divided word in above of header line using the histogram, discover the focuses from which the character begins and ends.Draw lines agreeing these arrange.

Pixel Fixed Segmentation

Keep up the information structure to bolster the line, word and character limits to such an extent that the character limit could be adequately extricated from the picture which is required for the further preparing and acknowledgment part of the framework

RESULTS AND CONCLUSION

Scene content extraction is late research zone in the field of PC Vision. It is testing issue in the data preparing field for the most part because of, various assortment of content examples like textual styles, hues, sizes, introductions; and nearness of foundation exceptions like content characters, for example, windows, blocks. In this paper, thinking about few of difficulties, ongoing application named as TravelMate is planned and created. Extraction is performed utilizing stroke width change and associated part based approach. Proposed application helps the travelers, while they are meandering in remote nations. The execution of framework is tried in light of extraction rate. With proposed application all the content in level introduction extricated effectively, though execution of continuous pictures fluctuates with lighting condition and camera determination. Proposed android application can be additionally stretched out to bargain with any objective and source dialect for interpretation. It can be additionally altered to manage content having vertical or self-assertive introduction Scene content extraction is late research zone in the field of PC Vision. It is testing issue in the data preparing field for the most part because of, various assortment of content examples like textual styles, hues, sizes, introductions; and nearness of foundation exceptions like content characters, for example, windows, blocks. In this paper, thinking about few of difficulties, ongoing application named as Travel Mate is planned and created.

REFERENCES

[1] Nitin Mishra, "ATMA: Android Travel Mate Application", IJCA (0975 – 8887) Volume 50 – No.16, July 2012.
[2] http://googlecode.blogspot.com/2006/08/announcing tesseract-ocr.htm.
[3] Smith, R. "An Overview of the Tesseract OCR" ICDAR 2007, Brazil.
[4] AyatullahFarukMollah, "Design of an OC R for Camera based Handheld Devices" IJCSIIJCS Issues, Vol. 8, Issue 4, No 1, July 2011.
[5]Sunil Kumar Rajat Gupta NitinKhanna, SantanuChaudhury and Shiv Dutt Joshi. Text extraction and document image segmentation using matched wavelets and mrf model. IEEE TRANSACTIONS ON IMAGE PROCESSING, 16:1–8, 2007.
[6] IEEE Li Linlin Chew Lim Tan Senior Member IEEE Shijian Lu, Member. Documentimage retrieval through word shape coding.pages 11–15, 2008.
[7] Anand Mishra KarteekAlahari C. V. Jawahar. Top-down and bottom-up cues for scene text recognition. pages 1–3.

[8] Yonatan Wexler Boris Epshtein, Eyalek.Detecting text in natural scenes with stroke width transform.pages 2–5.
[9] Qiuhau Lin Tong Zhang, Derek Ma. Mobile camera based text detection and translation. IEEE, 7(2):87–105, 2010

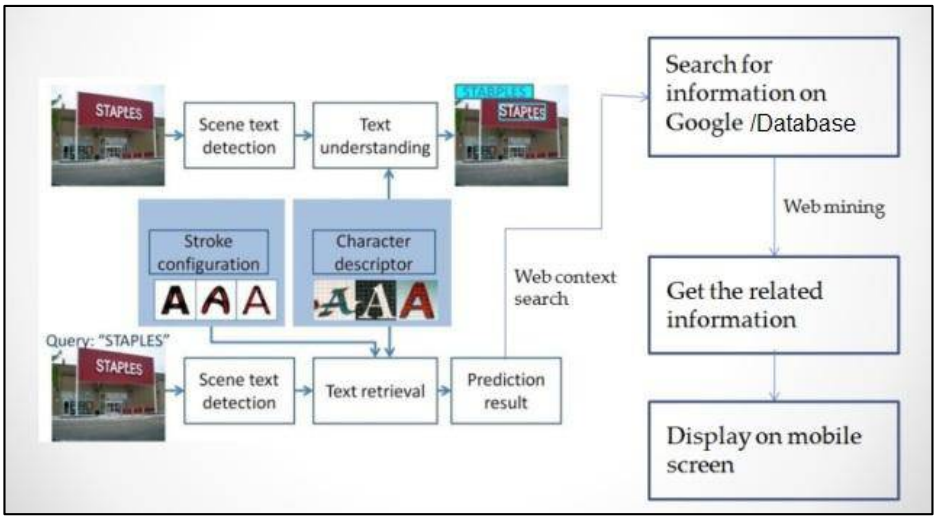


Figure1: System Architecture

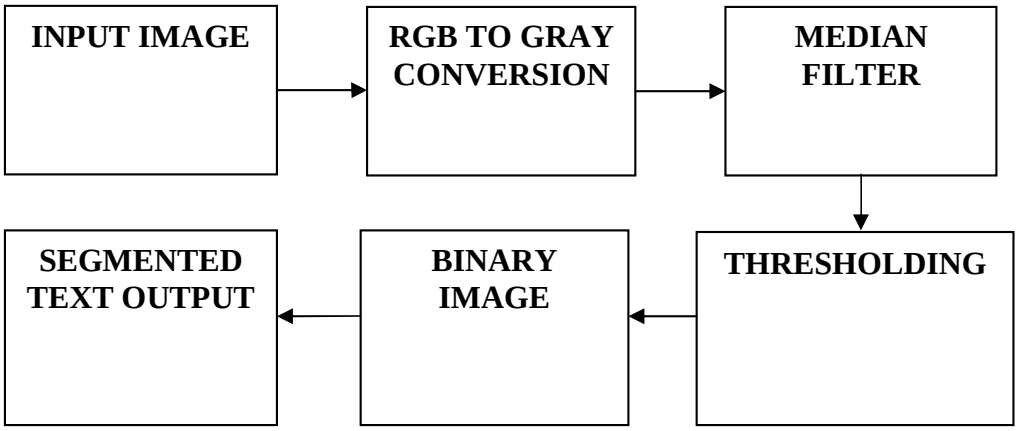


Figure 2: Block Diagram

Indian Journal of Natural Sciences *www.tnsroindia.org.in ©IJONS*

Vol.8 / Issue 47 / April 2018 *International Bimonthly* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Location Based Image Retrieval System on Ranking User Clicks

P.Rajesh Kanna[1]* and S.Keerthi[2]

[1,2]Assistant Professor, Department of Computer Science and Engineering, M.Kumarasamy College of Engineering, Karur, TamilNadu,India.

Received: 10 Feb 2018 Revised:18 Feb 2018 Accepted: 30 Mar 2018

***Address for correspondence**
P.Rajesh Kanna
Assistant Professor,
Department of Computer Science and Engineering,
M.Kumarasamy College of Engineering, Karur. TamilNadu,India.
E.mail: rajeshkanna@gmail.com

ABSTRACT

Nowadays everyone has started using search engines for retrieving images, which may display irrelevant images as search results. This is due to some mismatch between the text typed in the search box and the image displayed as the result. In order to overcome this problem, the image ranking model is introduced. The existing ranking model displays the results based on the already viewed images, which is not more efficient too. So, a new method of ranking model is proposed , which uses a definite framework for ranking. In this method, whenever the user searches for the same image again, the hyper graph is used to categorize the images. This system takes the complete role in integrating the most viewed images and for displaying it to the user. For Which, the Microsoft Bing Image Search Engine is used to collect the datasets for performing large scale experiments, which includes the features and the location-based image retrieval. The location-based filtration will allocate ranks for frequently clicked images based on the location of the user. Therefore, different results are displayed for different locations, based on the ranks allocated for each image. Thus, the expected image can be retrieved based on the overall search results.

Keywords: Image retrieval, Ranking model, Hypergraph, Microsoft Bing image search engine, Location-based image retrieval..

INTRODUCTION

Learning to rank has been widely used in the fields of information retrieval, data mining, and natural language processing. When a query is given, this ranking system retrieves the data from the collection and returns the top-ranked data for the corresponding query. Learning to rank methodology is used for retrieving documents, searching definitions, answering questions and summarizing documents.

The overall process involved in this methodology is detailed as follows. At first it checks the data query with the sample dataset and shows both the relevant and irrelevant results. Then the irrelevant data is removed using the filtration process, which includes assigning index values to images, ordering the images based on the index values, visual features and finally location filter is added.

The Ranking method is implemented by using the index values assigned to individual images based on which images are arranged in a linear manner. Every time the user clicks on the image, the index value gets updated.
The hyper graphs are the representation of images for image categorization. The images are analyzed as image pixels, color patterns, graphs and image frequencies. They are used for categorizing images based on the threshold value obtained from the graphs. In addition to that, Location features are added to view the image results based on their interested locations. By this way the user interested images can be retrieved for various locations.

Existing System

The web image search engines mostly use keywords as queries to search for images. User expects most relevant image results for the given keyword. For example, if "fruits" is a query keyword, then the images are retrieved from different categories, such as "red fruits," "fruits logo," and "fruits shop" etc…,

Many search engines use keywords for retrieving image which usually provides results from blogs. The user may not have a satisfaction with these results due to lack of trusts on blogs like quality. In early search engines, there were many mismatches to the searched terms. For that purpose, the ranking methodology is introduced to get most relevant image results from the user review.

Disadvantages of Existing System

Some images results were good in visuals but not relevant to expectations vice versa.
Without online training, the similarities of low level visual features may not correlate well with high-level semantic meanings of the image, which interprets the user's intention.

Proposed System

In this system, a method called the linearization approach has been proposed, which is used in image retrieval based on ranking. Though the training samples of image pairs can be collected easily, the objective of learning is formalized as optimizing errors in the classification of image pairs, rather than optimizing errors in the ranking of images. This method collects image pairs from the ranking lists, and assigns a label to each pair which describes the relative relevance of the two images. Then it trains a classification model with the labeled data and adopts it for ranking.

In order to know the ranking of images in each location, the location filters are added. To do this, list wise approaches have been proposed to learn the ranking function by adopting separate lists as samples.
For example, if the query keyword is "fruits," the concepts of "mountain" and "Paris" are irrelevant and should be excluded. The concepts of "tree" and "fruit" will be used as dimensions to learn the semantic space related to "fruits." The query-specific semantic spaces can more accurately model the images to be re-ranked. The visual and textual features of images are then projected into their related semantic spaces to get semantic signatures. At the online stage, images are re-ranked by comparing their semantic signatures obtained from the semantic space of the query keyword. The semantic correlation between concepts is explored and incorporated when computing the similarity of semantic signatures.

Advantages of Proposed System

Images results are relevant to the user expectations.
Images are re-ranked at every search. Hence, user can get the most relevant one.
The high quality images are delivered to the users using hyper graphs by image categorization.
Location system is delivered to know the variety of image results on different locations.

Modules

The overall searching process consists of four modules. They are,

User Endorsement

User Endorsement is the initial module in this application. The new user has to do the registration process to access the application in online. The registration process includes username, password, address, phone etc. Once the registration process is completed successfully, the user can login with the username and password and now the user is encouraged to search an interested image.

Implementation of User Click and Visual Features

At first, the images are initialized with index values. The index value gets updated based on the user clicks. The user click is defined as, whenever the users click and view particular image, the index value of the images are increased by one. So the images get its count based on the user views only. Then, the images are arranged in a linear manner using the linearization algorithm. The results are fetched from the title and Meta tags of individual image from the dataset. Hyper graph method takes the images as color frequency pattern and set threshold value to each and every image in the dataset. We can construct a hyper graph using mat lab tool by hyper works. That threshold value is used to categorize the images for better visual features. Now the visual feature uses the hyper graphs, in which the graphs are constructed for each image, and the images are categorized based on threshold value of the graphs.

Implementation of Location Based Search

In this module, images are retrieved based on location. Location filters are placed so that the user may use those filters to find image based on their interested location. If the user clicks an image, the count will be added to that particular location. Like this, every image has individual index value for setting location count also. She/he may also view how the images are ranked on other locations from their own location. The images will be short listed and displayed based on keyword given and also on location of the user.

Efficient Search Result

Here user submits the query to retrieve the required image results. The user can get efficient results from the click count feature, visual feature and the location based retrieval. So, here is the maximum probability for the user to get the efficient results based on the intension of the user. the main purpose of this module is to deliver the better results than the normal searching methods, by combining both the modules of user click and location based retrieval.

RESULTS AND FUTURE WORK

As the ranking is performed with the user click feature, the images are successfully retrieved as per the user's intension. So, the future work is to add the location based retrieval filter feature to experience the maximum efficient image results to the user and the visual features of the images datasets are analyzed and will be added to the existing datasets. Finally the images will be retrieved using all these features

REFERENCES

1. Jun Yu, Dacheng Tao, Meng Wang, Yong Rui, "Learning to rank using user clicks and visual fasures for image retrival".
2. C. Burges, "From RankNet to LambdaRank to LambdaMART: An overview," Microsoft Res., Tech. Rep. MSR-TR-2010-82, 2010.
3. Y. Cao *et al.*, "Adapting ranking SVM to document retrieval," in *Proc. SIGIR*, Seattle, WA, USA, 2006, pp. 186–193.
4. Z. Cao, T. Qin, T.-Y. Liu, M.-F. Tsai, and H. Li, "Learning to rank: From linearization approach to listwise approach," in *Proc. Int. Conf. Mach.Learn.*, Corvallis, OR, USA, 2007, pp. 129–136.
5. D. Cossock and T. Zhang, "Statistical analysis of Bayes optimal subset ranking," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 5140–5154, Nov. 2008.
6. R. Herbrich, T. Graepel, and K. Obermayer, "Large margin rank boundaries for ordinal regression," in *Advances in Large Margin Classifiers*. Cambridge, MA, USA: MIT Press, 2000, pp. 115–132.

X. Hua and G. Qi, "Online multi-label active annotation: Towards largescale content-based video search," in *Proc. ACM Int. Conf. Multimedia*, 2008, pp. 141–150.

7. K. Jarvelin and J. Kekalainen, "Cumulated gain-based evaluation of IR techniques," *ACM Trans. Inf. Syst.*, vol. 20, no. 4, pp. 422–446, 2002.
8. H. Li, "Learning to rank for information retrieval and natural language processing," *Synth. Lect. Human Lang. Technol.*, vol. 4, no. 1, pp. 1–113, 2011.
9. T. Liu, "Learning to rank for information retrieval," *Found. Trends Inf. Retrieval*, vol. 3, no. 3, pp. 225–331, 2009.
10. S. Robertson and S. Walker, "Some simple effective approximations to the 2poisson model for probabilistic weighted retrieval," in *Proc. Annu.Int. ACM SIGIR Conf. Res. Dev. Inf. Retrieval*, 1994, pp. 232–241.
11. J. Xu and H. Li, "AdaRank: A boosting algorithm for information retrieval," in *Proc. Int. ACM SIGIR Conf. Res. Dev. Inf. Retrieval*, 2007, pp. 391–398.
12. F. Xia, T. Y. Liu, J. Wang, W. Zhang, and H. Li, "Listwise approach to learning to rank: Theory and algorithm," in *Proc. Int. Conf. Mach.Learn.*, 2008, pp. 1192–1199

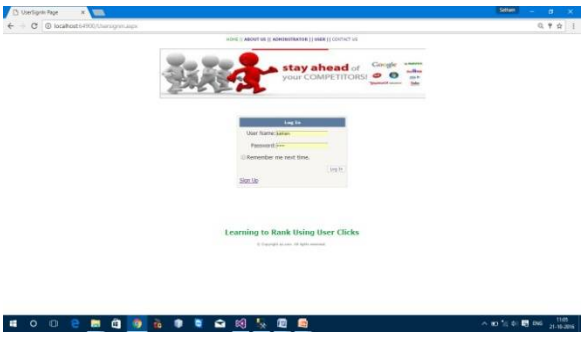


Figure:1

Figure: 2

Connecting People with Nature

INDIAN JOURNAL OF NATURAL SCIENCES

ISSN: 0976-0997

MJL& Cited by Thomson Reuters for Zoological survey in 2012

ISI Impact Factor :1.233 , NAAS Rating :3.56,IC Value: 47.65

Tamil Nadu Scientific Research Organization is working for the promotion of society by transferring science & technology since 1997. We are very keen in appreciating and recognizing the contribution of every one of you. We are publishing **Indian Journal of Natural Sciences** - **IJONS** from **August 2010**. The Journal is peer reviewed International Journal for publication of Original Research papers /Reviews/ Short communications/Book reviews/Reports on conferences/Seminar, Important events, News of interest, Special articles, General Articles etc.

Call for Papers: Investigations related to science in all branches

Highlights

- International Quality
- Published Bi-Monthly
- Fast acceptance and Quick Publication
- Low price
- High Rank Editorial Board
- Online manuscript submission
- Publish original Research Works and Reviews

Subscription Fee (Annual 6 Issues)

- For Individuals Rs.1000/-
- For students Rs.750/-
- For Institutions Rs.1200/-

Publication fee for Authors

- **Rs.3000/- (**Processing fee)
- **USD 100/-** for Foreign author

Send all correspondence to

----------Chief Editor ----------

Indian Journal of Natural Sciences (IJONS)

C/o TNSRO, 39 Mura Bhavan,Koodal Nagar,Rajagopalapuram Post

Pudukkottai-622003,TamilNadu,India.

Phone:04322-261088 Mobile: 99528 86637

E-mail: ijonstnsro@gmail.com.,**www.tnsroindia.org.in**

Plant More Trees ***Love Animals*** ***Conserve our Nature***