Journal of Mathematical Modelling and Applied Computing

Volume No. 13

Issue No. 3

September - December 2018

ENRICHED PUBLICATIONS PVT. LTD

S-9, IInd FLOOR, MLU POCKET,
MANISH ABHINAV PLAZA-II, ABOVE FEDERAL BANK,
PLOT NO-5, SECTOR-5, DWARKA, NEW DELHI, INDIA-110075,

Journal of Mathematical Modelling and Applied Computing

Aims and Scope

The Journal of Mathematical Modelling and Applied Computing is an Indian research journal, which publishes top-level original and review papers, short communications and proceedings on Interdisciplinary Integrative Forum on Modelling, Simulation and Scientific Computing in Engineering, Physical, Chemical Biological, Medical, Environmental, Social, Economic and Other Systems using Applied Mathematics and Computational Sciences and Technology.

Journal of Mathematical Modelling and Applied Computing

Journal of Mathematical Modelling and Applied Computing

(Volume No. 13, Issue No. 3, Sep - Dec 2018)

Contents

Theoritical Aspects of Data Analytics & Its Role in Customer Services and Security

Narendruni Lakshmi Priya[1], Dr. Satheesh Kumar Nagineni [2]

Department of Computer Science and Engineering

[1,2]OPJS University, Churu (Rajasthan), India

ABSTRACT

Every industry draws a stage nearer to understanding the universe of Big Data from how it is being connected in taking care of issues. Most ventures are as yet evaluating whether there is an incentive in actualizing big data, while some different businesses have connected efficiently Big Data analytics. Applications of Big Data were appeared in top ten businesses, for example, banking and, interchanges, media, and excitement, medicinal services suppliers, social insurance suppliers, training, assembling and common assets, government, protection, retail and discount exchange, transportation, vitality, and utilities. Despite the fact that Big Data faces particular difficulties, its usage has been drilled by businesses in these areas. The action of retailing and wholesaling is a piece of our economy and also everyday life.

1. INTRODUCTION

Purchaser and business markets purchase items and services regular as indicated by their requirements and inclination. The retail and discount divisions contribute altogether to the nation's national economy. In the present aggressive and complex business world, the organization needs to depend on the data-organized and new data-unstructured or semi- organized to go down their choices. BDA can bring benefits for e-merchants by enhancing market exchange cost proficiency, administrative exchange cost productivity and time cost productivity. Particularly in the internet business setting, Big Data empowers shippers to track singular client's behaviour and decide the best approaches to change over one-time clients into general clients. The infusion of big data analytics into an organization's esteem tie likens to 5– 6% higher profitability contrasted with their rivals [1].

Late examinations are concentrating on positive components of applying Big Data analytics with little regard for the negativeimpacts of applying Big Data analytics, for example, protection and security [2],

shopping dependence and group impacts. Be that as it may, the positives and negatives of applying big data analytics on clients' reactions have not been accounted for. Before 2008, three models of buyers' behaviour were found; the clients expected to expend more items. In 2008, the global monetary and money related emergency that happened everywhere throughout the world had driven clients to reconsider before purchasing. Subsequently, clients were obtaining less, and their behaviour ended up plainly protective. Today, clients confront monstrous and assorted data.

2. POSITIVE FACTORSFOR APPLYING BIG DATA ANALYTICS

Positive factor of applying Big Data analytics application incorporates offering data seek, proposal system, dynamic valuing and customer administration to associate with the group part. By group diverse data in Big Data period, for example, geographic dissemination, passionate inclinations, customer behaviour on shopping and also social association, side interests, and organizations can accomplish request introduction inclination introduction, relationship introduction, and different approaches to fulfill customers. Internet business sellers utilized data and correspondence advance through utilizing distinctive data mining methods to give customized services to customers, update the site to give better services [3].

E-sellers apply Big Data analytics to make customized offers, set dynamic cost, and offer the correct channel to give consumer value. Applying Big Data analytics by offering virtual shopping knowledge, a more straightforward ordeal of customized items will invigorate consumers want to purchase items. All these four positive applications of positive factor above will help find customers' aim, bring great customers' behaviour lastly lead them to make a move to purchase an item or administration from e-merchants. Sincerely determined consumers are anything but difficult to instigate their buy want and request by arranging data. The speed and comfort of social affair online data are one of the apparent values for customers when they shop on the web. A site utilizing Big Data analytics device can channel and peruse a substantial number of data to customer data [4].

Content digger technology is utilized to fathom inside the web and content pursuit and note the pertinence of history with libraries, indexes, and occurrences. Big Data is about the significance and offering the correct items or services to the opportune individual at the correct cost using the correct channel at the ideal time. For instance, Google customizes its query items in light of clients profile and Amazon offers distinctive landing pages with various items on offer to practically every guest. It returns to knowing your customer by joining diverse data sources to comprehend what they are searching for truly.

A proposal system given the customer's buy behaviour can assess item data, think about the interests of customers, item coordinating and suggest customers substitute or correlative items. Recommender systems help people to distinguish things that may hold any importance with them from a huge accumulation of things by conglomerating contributions from all people. In these systems, proposals are typically made given a blend of past buying or perusing behaviour attributes of the things being considered, and statistic and individual inclination data of customers. Chevalier and Mayzlin [5] showed that other Internet consumers' item suggestions affected consumer buying behaviour at online retailer destinations.

Online business suggestion system can help consumers to pick most loved items that can be actualized in genuine systems, for example, Amazon, Taobao, Google and different sites to advance the deal. Dynamic evaluating is an individual-level value separation technique where costs are charged to the customer, area, item, or time. Dynamic evaluating, regularly alluded to in monetary terms as individual-level cost segregation, has turned out to be considerably more typical with the expanded commonness of Internet promoting.

Dynamic evaluating is characterized as the purchasing and offering of items in business sectors where costs are allowed to modify in reactions to free market activity conditions at the individual exchange level. Along these lines, dynamic evaluating can pull in many retailers with the capacity to utilize the recently accessible data to independently set costs given a given customer's ability to pay. The reason for dynamic evaluating is to augment the dealer's benefit by accusing consumers of the most astounding costs every consumer will pay by controlling the extent and the transient closeness of value contrasts they will utilize. Consumers' responses to this valuing plan technique will significantly affect their fulfilment with buys and their consequent behavioral goals. For instance, Amazon typically changes the cost of things sold on its site on day by day, week by week, or month to month premise by 5%, 10% or 15% [6].

Dynamic valuing hones by merchants in reactions to portion and individual level contrasts have been made more achievable as online clients' conduct increments. Reliable with the suggestions, the present research explores the impacts of different dynamic valuing settings and can be consider as an extra exchange attributes. Financial hypothesis contends that dynamic evaluating (i.e., individual- level price discrimination) is normally useful for the profitability of the firm since it enables the firm to catch a bigger offer of the shopper excess. Nonetheless, confirm from late retail tries different things with Internet based dynamic evaluating proposes that shoppers respond emphatically against this training.

3. BIG DATA IN CUSTOMER SERVICE

Positive factor of applying Big Data analytics application incorporates offering data seek, proposal system, dynamic valuing and customer administration to associate with the group part. By group diverse data in Big Data period, for example, geographic dissemination, passionate inclinations, customer behaviour on shopping and also social association, side interests, and organizations can accomplish request introduction inclination introduction, relationship introduction, and different approaches to fulfil customers. Internet business sellers utilized data and correspondence advance through utilizing unique data mining methods to give customized services to customers, update the site to give better governments [7].

E-sellers apply Big Data analytics to make customized offers, set dynamic cost, and offer the correct channel to give consumer value. Applying Big Data analytics by providing virtual shopping knowledge, a more straightforward ordeal of customized items will invigorate consumers want to purchase items. All these four positive applications of decisive factor above will help find customers' aim bring great customers' behaviour lastly lead them to make a move to purchase an item or administration from e-merchants. Sincerely determined consumers are anything but difficult to instigate their buy want and request by arranging data. The speed and comfort of social affair online data are one of the apparent values for customers when they shop on the web. A site utilizing Big Data analytics device can channel and peruse a substantial number of data to customer data.

Content digger technology is utilized to fathom inside the web and content pursuit and note the pertinence of history with libraries, indexes, and occurrences. Big Data is about the significance and offering the correct items or services to the opportune individual at the right cost using the correct channel at the ideal time. For instance, Google customizes its query items in light of clients profile and Amazon offers distinctive landing pages with various items on offer to practically every guest. It returns to knowing your customer by joining diverse data sources to comprehend what they are searching for indeed. Data look shows that data quality and seeking administration quality [8].

Data quality is a measure of value saw by yield gave a site. Data attributes, for example, refresh, valuable, nitty-gritty, precise, and finished have been seen as critical segments of data quality. Seeking administration quality can be characterized as general customer assessments seeing nature of looking administration as fast responsiveness, reasonable and practical. Given customer's decision and activity, online retailer utilizing Big Data analytics can give continuous policies to customers. This activity may wind up noticeably one of the wellsprings of upper hands to pick up customer's fulfilment. Proposal

systems are worked by popular locales, for example, Amazon, eBay, Netflix, Monster, and other Retail stores where everything is suggested.

This includes a connection between e- sellers and purchasers whereby the purchasers give their data, for example, leisure activities and inclinations, while the e-merchants offer a suggestion fitting their needs, along these lines profiting both. Subtle elements are given on fundamental standards behind proposal systems: client-based community oriented separating which utilized likenesses in client rankings to foresee their interests and thing based cooperative sifting as focuses in the space of things. Synergistic separating systems utilize customer cooperation's and item data with overlooking different elements to make recommendations [9]. Dynamic evaluating is an individual-level value separation technique where costs are charged to the customer, area, item, or time. Dynamic evaluating, regularly alluded to in monetary terms as individual-level cost segregation, has turned out to be considerably more typical with the expanded commonness of Internet promoting. Dynamic evaluating is characterized as the purchasing and offering of items in business sectors where costs are allowed to modify in reactions to free market activity conditions at the individual exchange level. Along these lines, dynamic evaluating can pull in many retailers with the capacity to utilize the recently accessible data to independently set costs given a given customer's ability to pay.

The reason for dynamic evaluating is to augment the dealer's benefit by accusing consumers of the most astounding costs every consumer will pay by controlling the extent and the transient closeness of value contrasts they will utilize. Consumers' responses to this valuing plan technique will significantly affect their fulfilment with buys and their consequent behavioural goals. For instance, Amazon typically changes the cost of things sold on its site on day by day, week by week, or month to month premise by 5%, 10% or 15% [10].

4. BIG DATA AT THE SERVICE OF PRIVATE COMPANIES

At present, Big Data is especially used to the benefit of the private division, very in light of the way that it is the private section which has the information about purchasers. Relational associations, web records, working frameworks, online sellers… these services sit on colossal measures of information about customers and is used to serve the interests of those services, going from opening up arrangements to improving/invigorating their services and perceiving bugs or security defects. In any case, this information could in like manner be used as a piece of the excitement of customers themselves. For instance, customers could choose to what degree occurs they are shown rely upon the examination of their past endeavors and exercises with a particular true objective to get either "changed" proposition or

"fair-minded" rundown things. Decisions to oblige access to their information and how it impacts what they are being shown can in like manner limit dynamic assessing.

By giving the probability to customers to get to their own specific information and scientific counts, the justification of individual risk based assessing discussed in the range entitled "the complete of insurance?" could be exchanged. Through access to their information and computations, customers could guide their probability of having a minor collision in light of their driving conduct, the probability of having a mishap in the midst of trip in light of their activities and the country they go to, the probability of working up an ailment in perspective of their eating routine et cetera. The "last item" could be two-cover: the complete of insurance (since just customers who have a high danger would consider paying for assurance) and in addition an unprecedented homogenization of purchaser conduct or hoodwinking their information to get the best course of action [11].

In the last case, regardless of the possibility that the impact appears to be sure, melding consumer behaviour in such a way may go too far. Returning to the issue of "computerized guidance," such a device could work in light of a legitimate concern for the consumer relying upon how the calculation is arranged. Clients could get an intriguing understanding of their money-related circumstance and open doors for better monetary/budgetary administration given the examination of their data. This raises, once more, the topic of the "lack of bias" of calculations, which may be very reliant on who codes them and whether their interests are "coded" into the calculation. One illustration is the Swedish substantial data accumulation done by open experts in the field of health and a social need which expects needs with inventive arrangements [12].

5. IMPROVEMENT OF THE CUSTOMER EXPERIENCES THROUGH BIG DATA ANALYTICS

With digital data development anticipated that would increment all inclusive by 4,300 percent by 2020 and aggressive tensions building, organizations should now like never before meet the rising requests of their customers. Nonetheless, challenges teem with less assets, restricted spending plans, and more established innovation that don't generally measure up for the undeniably refined assumptions in regards to multi- channel engagement, administration, and esteem. Close by these difficulties, this digital revolution is likewise giving extraordinary chances to enhance the general customer encounter by means of big data analytics. This is the way toward examining and translating these tremendous amounts of data to separate the important, keen, and helpful data that offers some benefit to a customer. Ground breaking organizations are gathering, investigating, and following up on data to enhance various regions

[13]:

- **Market competitiveness**: In manufacturing, incorporating data from R&D, building, and manufacturing units to empower simultaneous designing can fundamentally slice time to showcase and enhance quality.
- **User productivity:**In the public sector, making important data available crosswise over generally isolated divisions can strongly diminish inquiry and preparing time.
- **Sales revenues and profitability:**As per McKinsey, retailers utilizing big data analytics can expand their working edges by more than 60 percent.
- **Customer experience**: Proactive interchanges through favored channels give improved matches amongst

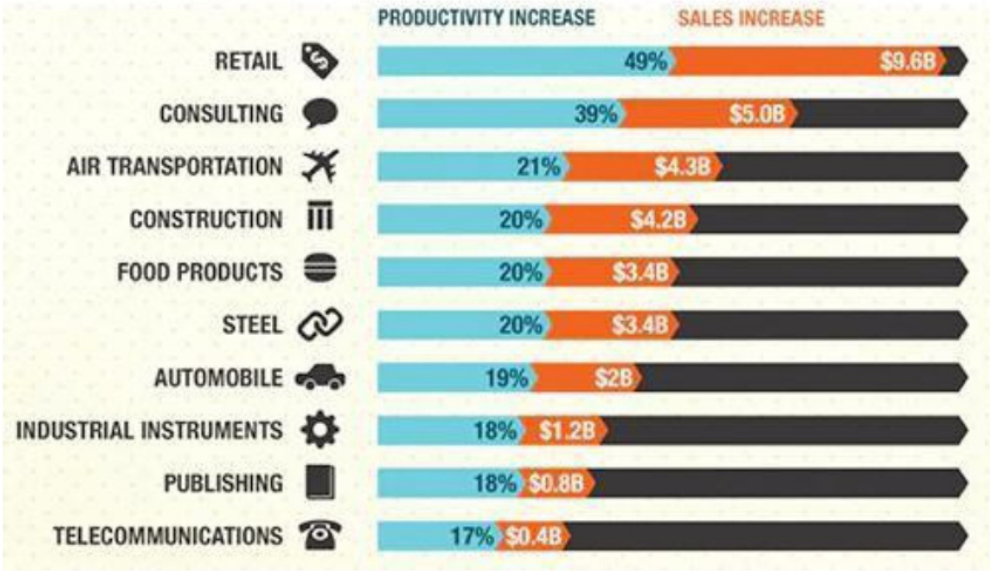


Figure 1: Big Data Analytics Industry Value

Source: Measuring the Business Impacts of Effective Data

The nature of customer experience will dependably be enhanced if a business can rapidly react by means of the right channels to deal with the desires of its customers. Using programming analytics to process the right data sources and measurements, and after that proactively giving pertinent and logical data, is vital. The following are three hints on staying aware of customer desires for customized, on-request record and administration data that will improve the general customer encounter [14].

1. Implement proactive bill shock management: Bill stun is customer "trouble" from unforeseen charges and is typically the aftereffect of broadband clients' powerlessness to measure their gigantic data utilization, particularly while wandering. These disappointed clients can adversely affect the correspondence specialist organization's notoriety and at last prompt stir and income misfortune. Broadband organizations can stay away from this by giving continuous authorization activities and

options. They can send email or content notices as a customer is near achieving the quantity, permit free limited perusing instead of simply crippling web access, execute programmed transmission capacity forming, and divert customers to substitute data intends to stay away from future issues.

2. Create smarter personalized shopping experiences: Opt-in portable advertising correspondences of focused items and administrations would then be able to be offered through customized messages particular to each phase of the purchaser cycle — mindfulness, engagement, thought, change and dedication. Suppose somebody picks in to get showcasing messages from a retailer who has an outlet in the neighborhood shopping center. GPS-coordinated following recognizes that the customer is in closeness to the store and sends the customer an instant message cautioning them to an exceptional one-day offer. The offer is driven by what the retailer definitely thinks about this specific customer, in light of profile, dependability card data, expressed dispositions, and prescient analytics demonstrating. With the customer's advantage aroused, she heads into the store and buys utilizing the coupon code in the instant message.

3. Reduce waiting time in the queue: A service organization, for instance, can address this lasting agony of masterminding a home repair visit by getting the shopper's favored channel of correspondence, affirming the evening before in a computerized way through that favored channel, and advising the customer that the administration tech will call at 8:00 a.m. to tell the customer where he remains in the everyday line. This enjoyments the customer and disposes of the cost of up to three inbound telephone calls: the day preceding ("Is the specialist truly coming tomorrow?"), 8:00 a.m. the day of the visit ("When is he coming?"), and the distraught call on the off chance that he isn't there by 10:30 a.m. ("I need to return to the workplace in time – would he say he is truly preceding twelve?")

6. CONCLUSION

This paper shows a few positive factors and negative elements and their belongings to customer reactions for use of Big Data analytics. Along these lines, organizations need to begin to adjust to the pattern utilizing Big Data analytics with a specific end goal to make due in the dynamic and digitalized markets. This is a procedure that arrangements with data, sources, abilities, and frameworks to make upper hands. The idea of big data has been created, and ought to be connected now to enhance techniques, expectation and basic leadership for better customer relations.

Be that as it may, applying Big Data analytics can likewise have advantages and disadvantages. E-sellers can improve the upsides of applying BDA yet don't slanted to over dependence on BDA keeping in mind

the end goal to maintain a strategic distance from negative viewpoints. Approving checks with their genuine case would make appropriate and successful advertising procedures. What's more, e-sellers that might want to work with Big Data analytics would require having enough data. This multiplies, as a type of run with the online national conduct. Customers contribute their data as well as e-sellers likewise add data to assemble big data.

REFERENCES

1. *Analysys Mason, Big data analytics: how to generate revenue and customer loyalty using real- time network data, January 2013,*
2. *A. Cavoukian, Privacy by Design – The 7 Foundational Principles, available at: https://www.privacybydesign.ca/i ndex.php/about-pbd/7- foundational-principles/*
3. *ISO/IEC, 29100:2011 Information technology – Security techniques – Privacy framework, 2011,*
4. *Goes, P. B. (2014) 'Big Data and IS Research', MIS Quarterly, 38 (3), iii-viii.*
5. *Chevalier, J.A.; Mayzlin, D. The effect of word of mouth on sales: Online book reviews. J. Mark. Res. 2006, 43, 345–354.*
6. *Jifeng, L., Sulin, B. & Han, Z. (2012) 'The Effectiveness Of Online Shopping Characteristics And Well-Designed Websites On Satisfaction', MIS Quarterly, 36 (4).*
7. *Mayring, P. (2003). Qualitative Content Analysis. eighth ed. BeltzVerlag, Weinheim, Germany*
8. *Tan, K.H. et al. (2015). Harvesting big data to enhance supply chain innovation capabilities: An analytic infrastructure based on deduction graph. International Journal of Production Economics, 165, pp.223–233.*
9. *Sanders, N.R. (2016). How to Use Big Data to Drive Your Supply Chain. California Management Review, 58(3), pp.26–48.*
10. *Park, H., Bellamy, M.A. and Basole, R.C. (2016). Visual analytics for supply network management: System design and evaluation. Decision Support Systems, 91, pp.89–102.*
11. *Zhang, Q. et al. (2013). A Case Study of Sensor Data Collection and Analysis in Smart City: Provenance in Smart Food Supply Chain. International Journal of Distributed Sensor Networks, pp.1–12.*
12. *Cao, G., Duan, Y. and Li, G. (2015). Linking Business Analytics to Decision Making Effectiveness: A Path Model Analysis. IEEE Transactions on Engineering Management, 62(3), pp.384–395*
13. *Burrows, R., & Savage, M. (2014). After the crisis? Big data and the methodological challenges of empirical sociology. Big data & society, April–June, 1–6.*
14. *Schouten, E. (2012). Big Data As A Service. Retrieved from: http://edwinschouten.nl/2012/09/19/bigdata-as-a-service/*

A Study On Image Processing for Satellite and Digital Images

Arpita Pandya, Dr. Priya R. Swaminarayan, Chaitali Pandya
Lecturer Institute of Computer Application, P.D.Pandya College, Gujarat University, Ahmedabad, India.
Professor & Head, MCA Department Institute of Science and Technology for Advanced Studies & Research Gujarat Technological University, Vallabh Vidyanagar, Gujarat, India.
Lecturer Institute of Computer Application, P.D.Pandya College, Gujarat University, Ahmedabad, India.
M: 9879370011 pandyacp@yahoo.com

A B S T R A C T

The studies of crop forecasting and prediction of plant leaf disease refer to the studies of visually observable patterns of a particular crop and plant. Image processing techniques are used for crop area estimation procedures and crop yield models, based on the application of satellite remote sensing, statistics and other allied disciplines as well as prediction of plant leaf disease based on the digital images. The research intends to take the help of satellite images, digital image,regional plant information, regional disease forecasting, regional crop information, regional crop forecasting, forecasting methods and image processing techniques. This paper presents an overview of how to use satellite and digital imagery for forecasting crop and leaf disease using image processing techniques .

***Keywords— Crop forecasting, Digital images, Image processing, Plant disease, Satellite* images.**

1. Introduction

Nature of the problem

An image is an array, or a matrix, of square pixels (picture elements) set in columns and rows. **Image processing** is a technique to convert an image into digital form and make some operations on it, in order to get an improved image or to remove some useful information from it. It is a category of signal special consideration in which input is an image, be fond of video frame or photograph and output may

be an image or characteristics attached with that image. generally **Image Processing** system consist of treating images as two dimensional signals while applying already set signal processing methods for them.[9]

Image processing includes the following steps.

- Importing the image with optical scanner or by digital photography.
- Analyzing and manipulating the images which includes data compression and image enhancement and spotting patterns that are not to seen by human eyes like satellite photographs.
- Output is the last stage in which outcome can be altered image or report that is based on image examination.

Types of image processing

The two types of methods used for Image Processing are Analog and Digital Image processing. Analog or visual techniques of image processing can be used for printouts and photographs. Image analysts use a range of basics of explanation while using these visual techniques. The image processing is not just limited to the area that has to be studied, but on facts of the forecaster.

Digital Processing techniques help out in the strategy of the digital images by using computers. As raw data from imaging sensors from a satellite platform contains deficiency. To find over such flaws and to get creativity of information, it has to feel various phases of processing. The three general phases so as to all types of data have to go through while using digital technique are Pre- processing, enhancement and display, information extraction [9]. In this circumstance, image processing has proved itself to be a extremely helpful tool. Images captured throughout satellite can added be analyzed by a variety of methods to gain the needed information.

Previous work

In a research by Domenikiotis, BMVCE – Bhalme and Mooley Vegetation Condition Index is proposed. It is the VCI – Vegetation Condition Index, extracted from NOAA – National Oceanic and Atmospheric Administration. Specific time period of April to August was selected to assess the collective meteorological effects on the cotton by processing VCI with proposed Bhalme and Mooley methodology. The resulting index was observed to be at the same level as Z-Index, which is classified as PDSI – PalmerDrought Severity Index. PDSI is widely used for drought monitoring. BMVCI was

observed to confirm unfavorable conditions for cotton production. The conclusion was derived from a study of 16 years of data. Proposed methodology proved to be a useful tool for early assessment of cotton production in Greece.[1]

Study by A. R. Reddy in 2013 assessed the spatio-temporal variations in cotton yields in India. The study indicated that among cotton growing districts in India, large quantum of spatial as well as temporal variations exist in cotton yields. It was observed that in 33 districts in India, cotton is grown in large area, yet the productivity is low. Study clarified as irrigated area, nitrogen and phosphorus are the main factors causing these variations. Increasing the levels of these three variables would increase the productivity of these districts. Compound growth rates, coefficient of variation, regression and generalized entropy were used to quantify the variability and to find factors responsible for this variation. [2]

A web based tool named as Identification is used to help the people who are not experts in identifying plant diseases in a particular way, which is totally based on the picture selection and/or little text descriptions. It is applicable when no suitable images exist, which represents the symptoms on a specific sample of plant tissues. The user can access this system from anywhere, it can be said as a multi accessed system, because the multi- access key of identification has to be generated, and it from the remote side or desktop computers or smart phone operators can easily use it. In this, the user selects pictures approaching the symptoms and the system give the most probable disease. [3]

The other study for the identification of symptoms of a plant disease, where the images are colored is a machine vision system. The region, where the diseases found, in the digital pictures were improved, separated, and a set of features were removed from each of them. Inputs to a Support Vector Machine (SVM) features were then used as classifier and tests were performed to identify the best organized model [4][5]. Proposed plant classification based on wavelet transforms and support vector machines. The experiment results show that the combine method had higher recognized and faster processing speed.

One study based on leaf image has been done. Some chemicals applied to the plants on the periodic basis. This kind of technique was only applied to the plants where the leaves already have been defected with the disease. Hundreds of chili plants were observed to perform disease forecasting. To detect disease on the chili plant leaf, the image processing technique plays a very important and useful role. This system will help farmers for the future monitoring and plantation .[6]

This paper discussed in the results of a pilot study held by Ukrainian Space Research Institute of NASU-NSAU, in association with the MARS team of the JRC. The purpose of the study was to explore the feasibility, cost-efficiency and specific difficulties of croparea estimation assisted by satellite remote sensing in Ukraine. Several image types like Landsat-5/TM, MODIS, LISS-III, AWiFS and RapidEye etc. were used in the study. Each of this kind of images was compared to the cost efficiency shared with a field assessment of a stratified sample of square segments. Also, image classification algorithms were trained with the help of field data that were collected "along the road". As a finding, it was derived that the TM images from Landsat 5 showed the best results, in spite of the old age of this sensor. Only Landsat TM and MODIS reached cost- efficiency threshold from all the sensors that were tested[7].

It is important to know the crop area for the estimation of total crop production in the country which in turn plays major role in agricultural commodities management. Images received from satellites and derived products can be used effectively for area correction from ground observation. This paper compared results and conclusions of the Ukrainian study in 2010 exploring the possibility and efficiency of crop area estimation in assistance with optical satellite images. Various satellite images with sensors like Landsat-5/TM, MODIS, LISS-III, AWiFS and RapidEye were associated with field survey on a stratified square segment sample of the crop area. The efficiency of various sensors was studied and analyzed under below: (1) Relative efficiency, which indicates how many times the inaccuracy of area estimates can be compact with the help of satellite images. (2) Cost-efficiency that shows how much cost of ground surveys can be reduced for crop area estimation with the help of satellite images. With the help of these criteria, each satellite image type was separately applied to select optimal dataset. It was found that MODIS and Landsat-5/TM only could reach the thresholds set for cost efficiency. Remaining was not cost effective due to their high price. [8]

Purpose

The rationale of image processing is divided into the following groups:

- Visualization - Observe the objects that are not visible.
- Image sharpening and restoration - To create a better image.
- Image retrieval - Seek for the image of interest.
- Measurement of pattern – Measures various objects in an image.
- Image Recognition – Distinguish the objects in an image.

With growing population difficulty all through the world and the need for improved agricultural production there is a specific need for enhanced management of the world's agricultural resources. It is

MATLAB R2010a

It supports the following graphics file formats:

- ✓ BMP (Microsoft Windows Bitmap)
- ✓ HDF (Hierarchical Data Format)
- ✓ JPEG (Joint Photographic Experts Group)
- ✓ PCX (Paintbrush)
- ✓ PNG (Portable Network Graphics)
- ✓ TIFF (Tagged Image File Format) XWD (X Window Dump)

- ➢ MATLAB software provides different image processing techniques.
- ➢ Image Processing Toolbox is used for converting satellite images into data.

III. Proposed System

Satellite Images

Fig. 1 Example of a Satellite image of a crop.

Fig. 2 Example of a Crop area

Digital Images

Fig. 3 Example of Digital images of plant disease.

Above Fig. 1 Shows area to be considered for crop forecasting of satellite image of a crop. Fig. 2: Shows the particular crop area of crop through satellite. Fig. 3 shows the plant affected from the disease.[13]

How to process

The proposed processing steps are below:

- ✓ The obtained satellite images are processed by software.
- ✓ Once we will make processing imagery, we may want to read or write that image so that we can, for example, place the processed image on the web. These instructions need the Image processing tool box.
- ✓ Image processing provides techniques for reading or determining the value of an image.
- ✓ This image processing tool determines the pixel values of an image and converts those pixel values into text data.
- ✓ After getting data from satellite images, compare with actual data using forecasting methods.
- ✓ Fig. 4 shows the overview of steps to devlop the proposed model.

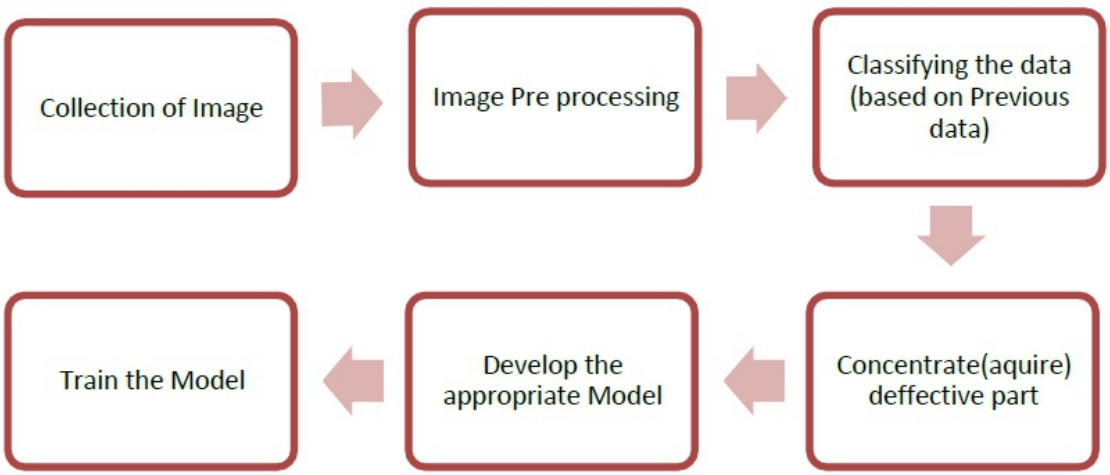


Fig. 4. A process chart

IV. Conclusions

The benefit of this research is to gain reliable, commercial, suitable, and repeatable information on agricultural production, to differentiate, identify, measure and map the area of crop as well as predict the leaf disease using image processing techniques. With the introduction of Remote Sensing Technology

and digital image processing, its great possible inside the field of agriculture have opened new vistas of civilizing the agricultural figures system all over the world.

Although there are some typical methods for image processing, there are no super image processing methods that can be regularly useful to all applications. Thus, it is a difficult task, as well as a hot research topic, to apply or develop image processing method fit for exact applications with high-resolution.

With the help of image processing, crop forecasting examination provides ability to determine prediction of yield production in a farm.

It is now recognized fact that image processing is an significant tool in crop analysis. There are many avenues where cost effective clarification, accurateness, straightforwardness are the major factors. There are sufficient of opportunities to build up a tool which is user forthcoming, easy yet it is cost effective.

Specifically, this paper focuses on the satellite images and digital images, commonly adopted image processing methods and using this method to predict the crop volume of the any crop area and prediction of leaf disease on any plant. This research is targeting on deriving the methodical simple methodology for crop and leaf disease forecasting through image processing.[12]

Applications: Following are the applications of image processing .[10]

- ✓ Intelligent Transportation Systems
- ✓ Remote Sensing
- ✓ Moving object tracking
- ✓ Defense surveillance
- ✓ Biomedical Imaging
- ✓ Automatic Visual Inspection System

REFERENCES

Journal Papers:

[1] Domenikiotis, C. a. (2004). Early cotton production assessment in Greece based on a combination of the drought Vegetation Condition Index (VCI) and the Bhalme and Mooley Drought Index (BMDI). International Journal of Remote Sensing, 25, 5373- 5388. doi:10.1080/01431160410001705033

[2] A. R. REDDY, S. M. (2013, July-Dec). Cotton productivity variations in India: An assessment. Cotton Research Journal, 5.

[3] Elad, I. P. (2012). Identificator: A web-based tool for visual plant disease identification, a proof of concept with a case study on strawberry. 144 - 154.
[4] Smith, A. C. (2009). Image pattern classification for the identification of disease causing agents in plants . Computers and Electronics in Agriculture , 121 - 125.
[5] Liu, j. (2009). A Method of Plant Classification Based on Wavelet Transforms and Support Vector Machine. 5754, 253-260.
[6] Husin, Z. a. (2012). Feasibility Study on Plant Chili Disease Detection Using Image Processing Techniques. (pp. 291-296). IEEE.
[7] Kussul, N. a. (2012). Crop area estimation in Ukraine using satellite data within the MARS project.
[8] Kussul, F. J. (2014). Efficiency assessment of using satellite data for crop area estimation in Ukraine . International Journal of Applied Earth Observation and Geoinformation , 22 - 30.

Web site:

[9] http://www.engineersgarage.com/articles/image-processing-tutorial-applications? page=1
[10] http://www.engineersgarage.com/articles/image-processing-tutorial- applications?page=2
[11] http://www.satimagingcorp.com/applications/natural-resources/agriculture/
[12] http://jpe.oxfordjournals.org/content/1/1/9.full
[13] https://www.plantvillage.com/en/topics/cornmaize/infos/diseases_and_pests_description_uses_propagation
[14] http://www.scilab.org/content/view/full/1092

An Algorith for SSD

Salma Sadiqha
Department of Computer Science and Engineering
Yanbu University College, Royal Commission at Yanbu, Yanbu Al-Sinaiyah, Kingdom of Saudi Arabia.

A B S T R A C T

In this paper, we present "An algorithm for SSD" which emulates the digits in a seven segment display form. The SSD algorithm is a flexible approach uses an object oriented. The SSD algorithm uses stacks and recursive calls used to display digits. The efficiency will be discussed by comparing "The SSD algorithm" to a digital display with stack and without using stacks.

Keywords*: Stacks, application, recursion, object oriented.*

INTRODUCTION

In this paper we will compare the features of the object oriented and procedural based language. In this research we will test only the two applications "The SSD algorithm" to other application displaying digits without using stacks.

The procedure oriented programming focuses on how to accomplish a task, the order of the instructions is important. Disadvantage of the procedural oriented language is that user views problem as series of steps rather than objects. The problems stated in objects are better visualized.

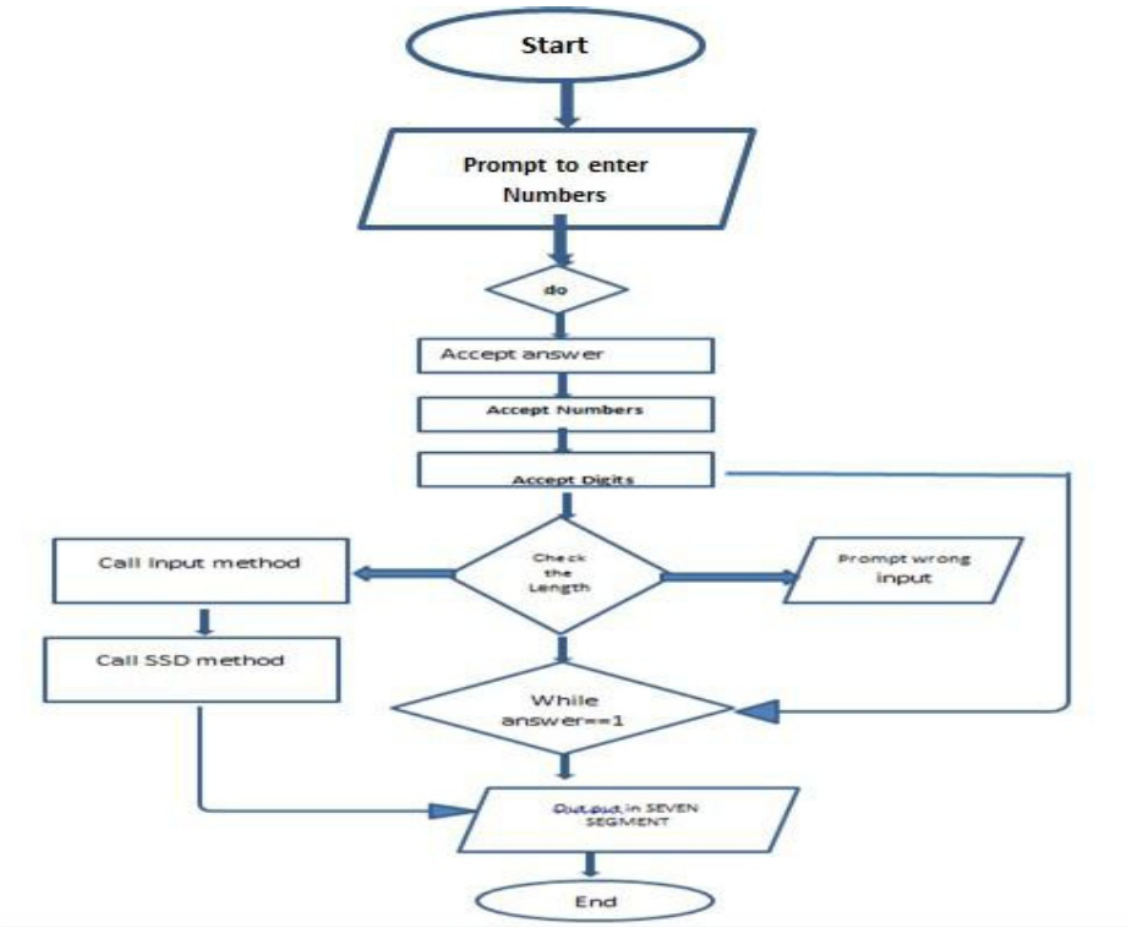
Thus an object oriented programming which breaks program as a concept of objects and methods that interacts to perform a specific task. Using object oriented approach the users can visualize

the solutions to problem more easily. Object oriented programs increases programmer's efficiency where users uses encapsulation and reusability concepts.

This paper also uses one of the data structure concept called as stacks. A stack is a container of objects that are inserted and removed according to the last-in first-out (LIFO) principle. In the pushdown stacks only two operations are allowed: pushing the item into the stack, and popping the item out of the stack. A stack is a limited access data structure that elements can be added and removed from the stack only at one end the top. Push operation adds an item to the top of the stack and pop operation removes the item from the top. If we talk about the daily life example like a stack of books; you can remove only the top book, also you can add a new book on the top.

METHODOLOGY

In this research we used object oriented approach which uses stacks. The stack stores digits and pops in reverse order. Input method is recursively called to make all the accepted digits to in a seven segment display format.



AlgorithmMainMethod:

input(numbersOfDigits); do {
Accepts Digits;

If Check(lengthtoDigitLength) Call Input methods recursively(Call SSD methods
}
else{
print"ERROR Wrong number of digits ...";
}
Print “to continue press 1? “ Input answer;
}
while (answer==1); print "DONE!";}}

Algorithm for Input Method:

Input(parameters)
use logic to get last digit push digit into stack

Algorithm for Convert to Seven Segment Method:

public void SSD(digit){ int ssd = 0;
for (int j=0;j<dig;j++){ if (!stack.isEmpty()) ssd =stack.pop();
check (conv equals ZER0)Call method Zero() Continue calling all digits.
Comparison Analysis:

	Non Procedure	SSD Algorithm
Input	Time taken in millisec	Time Taken in Millisec
10 digits	0	0
20 digits	1	2
30 digits	1	2
30 digits	2	4

Conclusion

The two main objectives of this research are, first to build an object oriented based application for solving real world problems using stacks which provides programmer's efficiency and second is comparing performance of two algorithms “The SSD algorithm” with stacks and without stacks. Finally we can conclude that Object oriented programming is better for improving programmer's efficiency and stacks can be used as a programmer's tool for various real world problems solving with better performance.

REFERENCES

[1] http://iacis.org/iis/2004/Chen.pdf
[2] OBJECT ORINETED PROGRAMMING TEXT BOOK
[3] Data Structures and Algorithms in Java (2nd Edition) 2nd Edition by Robert Lafore.

A Bird's View On Related Literature On Genetic Algorithm

Srujana Rongali[1], Radhika Yalavarthi [2]

[1] *Computer Science & Engineering Department, GITAM University, Visakhapatnam, India.*

[2] *Computer Science & Engineering Department, GITAM University, Visakhapatnam, India.*

A B S T R A C T

As quoted by Goldberg [1], Genetic algorithm (GA) is an adaptive optimization search algorithm aping the evolutionary ideas of natural selection. This Genetic Algorithm method is primarily applied haphazardly on an initial population and later all the individual chromosomes are appraised by a suitability function. This present paper is a review report on Genetic Algorithm and its functionaries on natural genetics and the evolutionary principle which was first proposed by Holland [2]. The GA is a simple but powerful tool for finding the global solution to an optimization problem.

Keywords: *Chromosomes, Evolutionary Principle, Genetic Algorithm, Natural Genetics, Optimization*

GENETIC ALGORITHM

Genetic algorithms (GA), is a generally considered as an adaptive optimization search as like Darwinian natural selection [3] and genetics in biological systems. This methodology is a promising alternative to conventional heuristic methods. The Genetic algorithms functions with a set of candidate solutions named as population. As like the Darwinian principle of 'survival of the fittest', the Genetic algorithm obtains the ideal solution after a cycle of iterative computations. Genetic algorithms transact with large search spaces competently, on a solution to the problem until satisfactory results are acquired by stimulating successive populations of substitute solutions that are represented by a chromosome. Allied with the characteristics of utilization and investigation search, this algorithm has less chance to get local optimal solution than other algorithms. A fitness function considers the quality of a solution in the evaluation step. In this method, crossover and alteration functions are considered as the main operators that haphazardly effect the fitness value. Chromosomes are selected for reproduction by assessing the fitness value. The fitter chromosomes have greater probabilities to be nominated into the recombination pool using the roulette wheel or the tournament selection approaches.

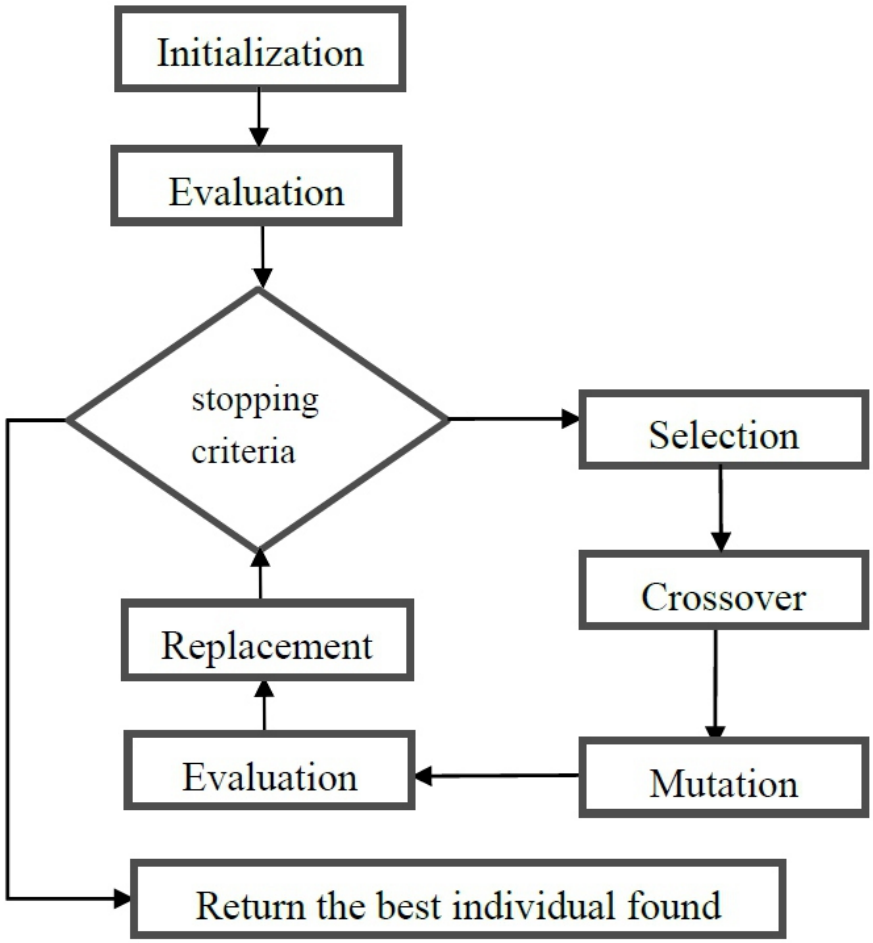


Figure 1: Flow Chart depicting the process of Genetic Algorithm

So the whole figure 1 can be explained as:

Step 1: randomly initialize population

Step 2: ascertain fitness of population

Step 3: repeat

a. choose the parents from population
b. execute crossover on parents creating population
c. execute the mutation of population
d. resolve fitness of population

Step 4: until best individual is good enough

The above figure 1 depicts the Crossover, the critical genetic operator that permits new solution regions in the search space to be surveyed. It is an arbitrary mechanism for exchanging genes between two chromosomes using the different crossover techniques like - one point crossover, two point crossover, or homologue crossover. In mutation the genes may sometimes be altered, that is in binary code, genes change from either 0 to 1 or vice versa. The offspring always replaces the old population using the elitism or miscellany replacement policy and forms a new population in the next generation [3].

Genetic Algorithms Overview

Genetic Algorithms [4] are borne on an analogy with the genetic structure and behavior of chromosomes. They are specified within a population of individuals by means of the following foundations:

- Individuals in a population contend for resources and mates.
- Those individuals who are most successful in each 'competition' will produce more descendants than those individuals that accomplish poorly.
- Genes from `good' individuals propagate throughout the population so that two good parents will sometimes produce offspring who are much better than either parent.
- Thus each successive generation will become more appropriate to their environment.

LITERATURE REVIEWS ON GENETIC ALGORITHM

Zhanchao Li et al., [5], conducted an experiment on protein science. A prior information about function of G protein-coupled receptors (GPCRs) can offer beneficial information to pharmaceutical research. The purpose of their function is quite significant theme in protein science. However, with the speedy increase of GPCRs structures entering into databanks, the breach between the amount of known sequence and the amount of known function is widening quickly. Both are time taking and costly to define their function based only on investigational practices. Hence, it is vitally significant to develop a computational process for swift and precise classification of GPCRs.This study brought in a novel three-layer predictor based on support vector machine (SVM) with feature selection, which is established for forecasting and categorizing GPCRs directly from amino acid classification data. The maximum relevance minimum redundancy (mRMR) is exerted to pre-assessed features with discriminative data while genetic algorithm (GA) is used to find the optimized feature subcategories. SVM is utilized for the production of classification models. The complete accurateness with three-layer predictor at levels of superfamily, family and subfamily are obtained by cross-validation assessment on two non- redundant

datasets. The results are 0.5% to 16% higher than those of GPCR-CA and GPCRPred. The results with high achievement rates specify that the anticipated predictor is a beneficial automated device in forecasting GPCRs. GPCR-SVMFS, is a corresponding implementable program for GPCRs forecast and sorting that can be acquired freely on appeal from the researchers.

Shuangyin Liu et al., [6] worked on prediction model based on support vector regression (SVR) to solve the aquaculture water quality prediction problem. The SVR parameters must be set carefully to build an effective SVR model. A hybrid approach known as real-value genetic algorithm support vector regression (RGA–SVR) study was presented, which searches using real-value genetic algorithms for the optimal SVR parameters and adopts the optimal parameters to construct the SVR models. To predict the aquaculture water quality data collected from the aquatic factories of YiXing in China, this approach is applied. From actual experiments using monitored aquaculture water quality data, the hybrid approach of SVR with genetic algorithm optimization is able to provide reliable data on the water quality prediction of large-scale intensive aquaculture. Based on the root mean square error (RMSE) and mean absolute percentage error (MAPE), the test results demonstrate that RGA–SVR outperforms the traditional SVR, back-propagation neural network models. To predict aquaculture water quality, the RGA–SVR model is proven to be an effective approach.

SU Bai-hua, and WANG Ying-lin [7] considered that in natural language processing, Semantic textual similarity is a common task. This measures the degree of semantic equivalence of two textual snippets. In recent times, machine learning methods have been put on to this task, including methods based on support vector regression. Still, there happens to be number of features consisting of the learning process. Certain features are noisy and irrelative to the outcome. In addition, the prediction performance of the SVR model will be significantly affected by different parameters. The researchers recommend genetic algorithm to select the effective features and optimize the parameters in the learning process all together. To assess the proposed method, they adopted the STS-2012 dataset in the trial. The anticipated GA-based method has improved regression performance when compared with the grid search.

Jui-Sheng Chou et al., [8] studied the hybrid system as an able device to dole out with construction engineering and management problems. The researchers recommend an optimized hybrid artificial intelligence model to assimilate a fast messy genetic algorithm (fmGA) through a support vector machine (SVM). In the initial phase of public-private partnership projects the fmGA-based SVM also called GASVM is utilized for initial forecast of dispute tendency. Predominantly, the SVM provides learning and curve fitting whereas the fmGA optimizes SVM parameters. To find the performance evaluation of proposed hybrid intelligence classification model area under the curve and synthesis

index, the measures are taken in form of accuracy, precision, sensitivity and specificity. When compared to other baseline models (i.e., CART, CHAID, QUEST and C5.0) and previous works, investigational assessments specify that GASVM attains improved cross-fold forecast accurateness. The forecasting results prove the proactive warning and decision-support information desired to accomplish potential disputes.

To screen the features of a dataset GA was chosen by Jack and Nandi ; Shon et al., [9], [10]. The chosen subset of features is then fed into the SVM for classification testing. A GA-based technique to find a subset of features for SVM in machine monitoring was developed by Zhang et al., [11]. Applying a GA approach to amend the RBF width parameter of SVM with feature selection was anticipated by Samanta et al., [12]. However, since these approaches only ponder the RBF width parameter for the SVM, they may slip the optimal parameter setting. A GA- based feature selection and parameters optimization for SVM was presented by Huang and Wang [3]. This GA-based feature selection and parameter optimization was utilized for credit scoring.

A study was conducted by Ilhan Ilhan and Gulay Tezel [13] on SNPs (Single Nucleotide Polymorphisms) that are promising tools for disease-gene association studies. However, such a study is inhibited by the high outflow of genotyping millions of SNPs. Hence, it is required to obtain a suitable subset of SNPs to accurately represent the rest of SNPs. Many methods have been developed for the purpose of selecting an appropriate subset of tag SNPs but all of them only offer low prediction accuracy. The authors developed, a new method is developed as GA–SVM with parameter optimization method that benefits from support vector machine and genetic algorithm to forecast as well as to select tag SNPs, respectively. Moreover, it also uses particle swarm optimization algorithm to optimize C and γ parameters of support vector machine. Experiment was carried out using various datasets and the attained results established that, this method can provide better prediction accuracy in recognizing tag SNPs when compared to other methods at present.

GTagger, a heuristic method using genetic algorithms suggested by Mahdevar et al., [14] used correlation and Shannon entropy to compute fitness function. This method results in low prediction accuracy.

The work done by Shutao Li et al., [15] for selecting high discriminative genes from gene expression data has become a vital research. When a vast number of noisy, redundant genes are filtered, this cant only improve the performance of cancer classification but also reduce the rate of medical diagnoses. In this research, a hybrid Particle Swarm Optimization and Genetic Algorithm method are used for gene

selection and Support Vector Machine is implemented as the classifier. The suggested method has been tested on three benchmark gene expression datasets like the data of Leukemia, Colon and breast cancer. Experimental results suggest that the proposed strategy can reduce the dimensionality of the dataset and improve classification accuracy.

Mingyuan Zhao et al., [16] performed a study on Generalization performance of support vector machines with Gaussian kernel which is influenced by its model parameters, i.e. error penalty parameter and the Gaussian kernel parameter. A new area distribution model is proposed after researching on the characteristics and properties of the parameter which results simultaneously with Gaussian kernel. The distribution model comprises of optimal straight line, reference point of area boundary, optimal area, transition area, under fitting area and overfitting area. To improve the classification performance of support vector machines, a genetic algorithm based on change area search a new area search is adopted. Comparison of experiments reveal that test accuracy of the genetic algorithm based on change area search a different is better than the two-linear search method.

A study by Y. Rong et al., [17], study recognized results of surface electromyography (sEMG) recorded under circumstances of a maximum voluntary contraction (MVC) in exhausted states using wavelet packet transform and energy analysis. The sEMG signals were listed in 10 young men from the right upper limb with a handgrip. The corresponding energies of certain occurrences were regulated as feature vectors by sEMG signals that were decayed by wavelet packet transform. To discriminate muscle states, a back-propagation neural network, a support vector machine (SVM) and a genetic algorithm-based SVM (GA-SVM) worked as classifiers. The outcomes display that muscle fatigue, MVC could be recognized by level-4 wavelet packet transform and GA-SVM more precisely than when using other methods. The classification accurate rate reached to 97.3% with seven fold cross-validation. The authors claim that proposed method can be used adequately to reflect the muscle activity.

Youngjoo Lee and Jeongjin Lee [18] performed a study on Support vector machine (SVM) with a binary tree architecture that is widespread as it requires the least number of binary SVM to be trained and tested. To design the tree architecture with optimal binary various efforts have been designed. On the other hand, these approaches generally construct a binary tree by a greedy quest. They consecutively divide the classes into two groups so that they consider only local optimum at each node. Genetic algorithm (GA) has been freshly presented in multiclass SVM for the local partitioning of the binary tree structure. The authors proposes a global optimization method of a binary tree structure using GA to develop the classification precision of multiclass problem for SVM. Unlike earlier researches on multiclass SVM using binary tree structures, this present approach universally finds the optimal binary tree structure. The

study proposes an enhanced crossover plan to include the determination method of crossover points and generation method of offspring to preserve the maximum information of apparent tree structure for the efficient utilization of GA. Investigational consequences revealed that the planned method provides higher accuracy than any other competing methods used as benchmark within the stipulating time. The presentation of the method for small size problems is comparable with other challenging methods while more practical developments of the classification accuracy are acquired for medium and large size problems.

Jian Zhao et al., [19] studied on the various treating parameters as model variables. They stressed on quality of plastic part, its warp, volumetric shrinkage and sink marks are to be minimized. The study is based on on two-stage optimization system where the first stage is an improved efficient global optimization (IEGO) algorithm which is adopted to estimate the nonlinear association between processing parameters and the measures of the part quality. In the second stage, to achieve improved spread of design solutions and better convergence near the true Pareto optimal front, non-dominated sorting-based genetic algorithm II (NSGA-II) was used. A cover of liquid crystal display part is optimized to show results that the Pareto fronts obtained by NSGA-II are distributed uniformly. Hence this algorithm has better convergence and vigor. The pair-wise Pareto frontiers illustrate that there is a significant trade-off amid warpage and volumetric shrinkage, and that there is no noteworthy trade-off between sink marks and volumetric shrinkage and between sink marks and warpage.

Chih-Hung Wu et al., [20] developed a genetic-based SVM (GA-SVM) model that can automatically determine the optimal parameters, C and σ, of SVM with the peak predictive accuracy and generalization ability simultaneously. The authors initiated to optimize the parameters of SVM for predicting bankruptcy by employing a real-valued genetic algorithm (GA). The mentioned GA-SVM model was tested on the prediction of financial crisis in Taiwan. The accuracy was compared with other models in multivariate statistics and artificial intelligence. Experimental results of proposed model indicate a very promising hybrid SVM model for predicting bankruptcy in terms of both predictive accuracy and generalization ability. The results of the GA-SVM can stand as a guide of investment for investors and government.

CONCLUSION

In the operations of selection, crossover, and mutation, the GA unites over successive generations towards the global optimum. This simple operation produces rapid, suitable and robust techniques is largely for the reason that GAs combine direction of an effective and efficient procedure. Since

population indirectly contain much more evidence than simply the individual fitness scores, GAs combine the good information concealed in a solution with good information from another result to produce many more outcomes with worthy information inherited from both parents, absolutely preceding towards optimality.

The expertise in the application of Genetic algorithm helps to discover and develop simultaneously a growing amount of theoretical validation, and successful application to the real- world problems. It also reinforces the conclusion that Genetic Algorithm is a powerful, effective optimization technique.

REFERENCES

[1] *D. E. Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning. Addison- Wesley Longman Publishing Co., Inc, 1989.*

[2] *J. H. Holland, Adaptation in Natural and Artificial Systems - An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence. Ann Arbor: The University of Michigan Press, 1975.*

[3] *C. L. Huang and C. J. Wang, "A GA-based feature selection and parameters optimizationfor support vector machines," Expert Syst. Appl., vol. 31, no. 2, pp. 231–240, 2006.*

[4] *"Introduction to Genetic Algorithms." [Online]. Available: https://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol1/hmw/article1.html.*

[5] *Z. Li, X. Zhou, Z. Dai, and X. Zou, "Classification of G-protein coupled receptors based on support vector machine with maximum relevance minimum redundancy and genetic algorithm," BMC Bioinformatics, vol. 11, no. 325, pp. 1471–2105, 2010.*

[6] *S. Liu, H. Tai, Q. Ding, D. Li, L. Xu, and Y. Wei, "A hybrid approach of support vector regression with genetic algorithm optimization for aquaculture water quality prediction," Math. Comput. Model., vol. 58, no. 3–4, pp. 458–465, 2013.*

[7] *S. U. Bai-hua and W. Ying-lin, "Genetic Algorithm Based Feature Selection and Parameter Optimization for Support Vector," J. Shanghai Jiaotong Univ., vol. 20, no. 2, pp. 143–148, 2015.*

[8] *J.-S. Chou, M.-Y. Cheng, Y.-W. Wu, and A.-D. Pham, "Optimizing parameters of support vector machine using fast messy genetic algorithm for dispute classification," Expert Syst. Appl., vol. 41, no. 8, pp. 3955–3964, 2014.*

[9] *L. B. JACK and A. K. NANDI, "FAULT DETECTION USING SUPPORT VECTOR MACHINES AND ARTIFICIAL NEURAL NETWORKS, AUGMENTED BY GENETIC ALGORITHMS," Mech. Syst. Signal Process., vol. 16, no. 2–3, pp. 373–390, 2002.*

[10] *T. Shon, Y. Kim, C. Lee, and J. Moon, "A machine learning framework for network anomaly detection using SVM and GA A Machine Learning Framework for Network Anomaly Detection using SVM and GA," in Proceedings of the 2005 IEEE Workshop on Information Assurance and Security United States Military Academy, 2017, no. February.*

[11] *L. Zhang, L. B. Jack, and A. K. Nandi, "Fault detection using genetic programming," Mech. Syst. Signal Process., vol. 19, no. 2, pp. 271–289, 2005.*

[12] *B. Samanta, K. R. Al-Balushi, and S. A. Al-Araimi, "Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection," Eng. Appl. Artif. Intell., vol. 16, no. 7–8, pp. 657–665, 2003.*

[13] *I. Ilhan and G. Tezel, "A genetic algorithm – support vector machine method with parameter optimization for selecting the tag SNPs," J. Biomed. Inform., vol. 46, pp. 328–340, 2013.*

[14] *G. Mahdevar, J. Zahiri, M. Sadeghi, A. Nowzari-dalini, and H. Ahrabian, "Tag SNP selection via a genetic algorithm," J. Biomed. Inform., vol. 43, no. 5, pp. 800–804, 2010.*

[15] *S. Li, X. Wu, and M. Tan, "Gene selection using hybrid particle swarm optimization and genetic algorithm," Soft Comput., vol. 12, no. 11, pp. 1039–1048, 2008.*

[16] *M. Zhao, R. Jian, J. Luping, F. Chong, L. Jianping, and Z. Mingtian, "Parameter selection of support vector machines and genetic algorithm based on change area search," Neural Comput. Appl., vol. 21, no. 1, pp. 1–8, 2012.*

[17] *Y. Rong, D. Hao, X. Han, Y. Zhang, J. Zhang, and Y. Zeng, "Classification of Surface EMGs Using Wavelet Packet Energy Analysis and a Genetic Algorithm-Based Support Vector Machine," Neurophysiology, vol. 45, no. 1, 2013.*

[18] *Y. Lee and J. Lee, "Expert Systems with Applications Binary tree optimization using genetic algorithm for multiclass support vector machine," Expert Syst. Appl., vol. 42, pp. 3843–3845, 2015.*

[19] *J. Zhao, G. Cheng, S. Ruan, and Z. Li, "Multi-objective optimization design of injection molding process parameters based on the improved efficient global optimization algorithm and non-dominated sorting-based genetic algorithm," Int. J. Adv. Manuf. Technol., vol. 78, no. 9, pp. 1813–1826, 2015.*

[20] *C. H. Wu, G. H. Tzeng, Y. J. Goo, and W. C. Fang, "A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy," Expert Syst. Appl., vol. 32, no. 2, pp. 397–408, 2007.*

Sensor Network Security Algorithm Based On Evidence Theory

Sachika , Pramod Kumar, Pooja Khokher

Student ,MDU/ MERI College Sampla,Jhajjar , Haryana , India

Asstt. Professor Computer Science, MDU/Nehru College, Jhajjar, Haryana, India

Computer Science, Govt. College ,Dujana ,Jhajjar, Haryana, India

A B S T R A C T

WSN i.e. Wireless sensor network are having low cost & multi functional resources and nodes with sensors to receive & transmit the data throughout the network .These nodes are very critical to the environmental situations .Incase of very sensitive applications the matter of securitis at the peak as these type of network deals with a maximum factor of interaction through network .Hence, very much sensitive to both type of attacks: external & internal .Internal attacks are more crucial in this case as we have to make interaction with the network information again & again in which that malicious node is participating every time , that is the reason ,it needs hard efforts to identify the suspicious node. One more reason why it is crucial is that that suspicious node is having all the rights for accessing the information as any other network node .In this paper we will make use of a three stage model to identify the attacker node that is present in the network .Firstly, all the nodes checked for 3 properties throughput , energy ,delay to guess their behavior for abnormality. Secondly, neighbor nodes are made to check their properties that there are interacting normally. Thirdly , these are again checked with the use of evidences previously obtained that they are strong enough to prove the node as attacker one. If yes the node is removed from the network .The experiments gives appropriate results with the last stage.

Index Terms*— Attacker , Dampster Shafer Theory ,3-level check algo , Energy, Nodes*

I. INTRODUCTION

As a new type of network technology rising from the military field, WSNs can be widely applied to civil, commercial, medical field and other critical ones. Covering which the security problems are more important and noticeable. In WSNs, the attacks are complicated and various, the threats faced by WSNs are not only from external attackers, but also from internal nodes which are a part of the communication due to the rights allotted to them for network use. In such a systems it becomes difficult to mark which

one is attacker or suspicious .If a right one is marked as attacker it may cause remarkable harm to the network information as that node become aware of checking, & behave intelligently in inspection of network. Another reason is that comparing with the external attacks, the internal attacks are more difficult to defense because that the key mechanisms are ineffective for internal malicious nodes, thus the internal attacks can make worse threats to the network. So it needs to be solved urgently for the legitimate nodes to detect and further eliminate the malicious nodes.

WSN is a network with functionality as:

Sensing + Processing + communication

Sensing-

It is storing of information i.e. collected data about any change in the defined property likewise temperature, pressure , moisture etc. It is the recording of physical data about any parameter .After that, it transmits it to the controller for further processing.

Processing-

It takes the data from sensors & processes it and supervises the other components of the node .The processor may be micro controller, desktop microprocessor.

Communication-

Sensor nodes often make use of ISM band, which includes free radio, spectrum allocation and global availability. The possible choices of wireless transmission media are radio frequency (RF), optical communication (laser) and infrared. Lasers communication requires less energy, but needline-of-sight for successful communication and is sensitive to atmospheric conditions. Infrared, like lasers, needs no antenna but it is limited in its capacity of broadcasting.

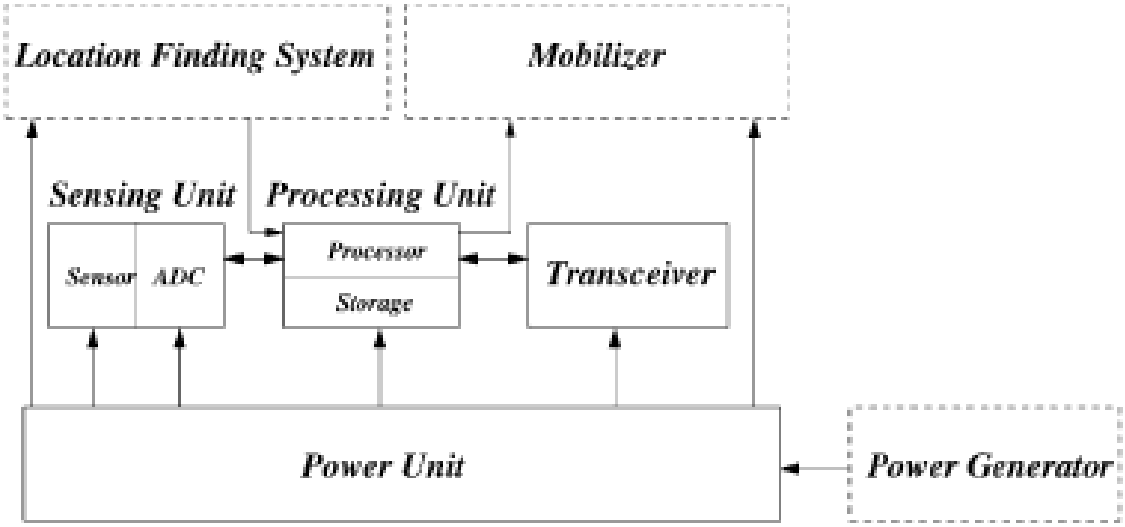


Figure 1. Components of WSN sensor node

Related Work-

Presently, the manners of identifying the malicious nodes of WSNs as follows:

A statistics-based malicious node detection scheme is

proposed by Ana Paula. A series of regulations are predefined to define the normal behaviors of nodes and further judge the anomaly or any deviation in behaviors of nodes, there is no interaction among nodes, so the rate of false alarming is quite high.

A rule-based malicious node detection scheme in Ad Hoc is proposed by

Chin-Yang].This scheme uses the monitoring points distributing in the network to monitor nodes whether operate in accordance with the routing norms in the process of AODV route query phase, then a finite state machine formed by the norms is used to identify nodes as normal state, suspected state, and intrusion state.

A Markov Chain based anomaly detection algorithm proposed by Bo Sun. This algorithm only aims at identifying the malicious nodes which launch route spoofing attacks, and needs to preserve a large amount of status information about changing of routing table. So it is not suitable for the resource-constrained sensor networks.

ACK based anomaly detection algorithm is proposed by

D. Tian and N. D. Georganas. In this algorithm, a next-hop

ACK feedback technology is used to identify the unreliable communication links.

A checkpoint-based multi-hop acknowledgement scheme

for detecting selective forwarding attacks is proposed by Bo

Yu . In this scheme, part of intermediate nodes along a forwarding path can be randomly selected as checkpoint nodes which are responsible for generating acknowledgement for each packet received. On receiving an incident packet, a checkpoint will create an ACK for this incident packet and upstream transmit the ACK along the routing path. If an intermediate node does not receive a sufficient number of ACK packets, an alarm packet which designates its downstream neighbor node as the suspicious node will be created and transmitted to the source node by such an intermediate node.

Proposed Work-

In this work we give a algorithm for detecting the suspicious node with a 3 degree confirmation model. In this algorithm WSN is created with predefined no. of sensor nodes which are checked for communication under specified parameters i.e. energy , throughput , Delay time .

In our work , these parameters play a important role for the evaluation of a node to be suspicious or normal .Its behavior is checked either it is normal or abnormal .That means more energy , more delay , and high throughput or u pretended throughput .

Secondly , It is checked by a set of neighbor nodes .In other words a set of monitoring nodes are assigned for the purpose of rechecking the behavior of the above saved suspicious nodes so that non confusion or mistake could be there in decision .

Thirdly, Saving all the evidences of above checking, Evidence –belief theory is used for checking the plausibility & belief of these evidences .This theory implemented to find out the correctness of the doubt on the said node.

Evidence theory -a degree of belief also called mass is represented as a belief function. Probability values are assigned to *sets* of possibilities rather than single events: their appeal rests on the fact they naturally encode evidence in favor of propositions.

Shafer's framework allows for belief about such propositions to be represented as intervals, bounded by two values, *belief* (or *support*) and *plausibility*:

belief ≤ *plausibility*.

Here, subjective probabilities (*masses*) are assigned to all subsets of the frame; usually, only a restricted number of sets will have non-zero mass (*focal elements*). *Belief* in a hypothesis is constituted by the sum of the masses of all sets enclosed by it. It is the amount of belief that directly supports a given hypothesis or a more specific one, forming a lower bound. *Belief* (usually denoted *Bel*) measures the strength of the evidence in favor of a proposition *p*. It ranges from 0 (indicating no evidence) to 1 (denoting certainty). *Plausibility* is 1 minus the sum of the masses of all sets whose intersection with the hypothesis is empty. Or, it can be obtained as the sum of the masses of all sets whose intersection with the hypothesis is not empty. It is an upper bound on the possibility that the hypothesis could be true, *i.e.* it "could possibly be the true state of the system" up to that value, because there is only so much evidence that contradicts that hypothesis.

The algorithm works as follows:

This algorithm works in 3 stages Initial Check, Monitor Check, Belief Check

Initial Check:

In this a network is designed of WSN nodes including sensor nodes , One Base Station , some of the suspicious nodes.

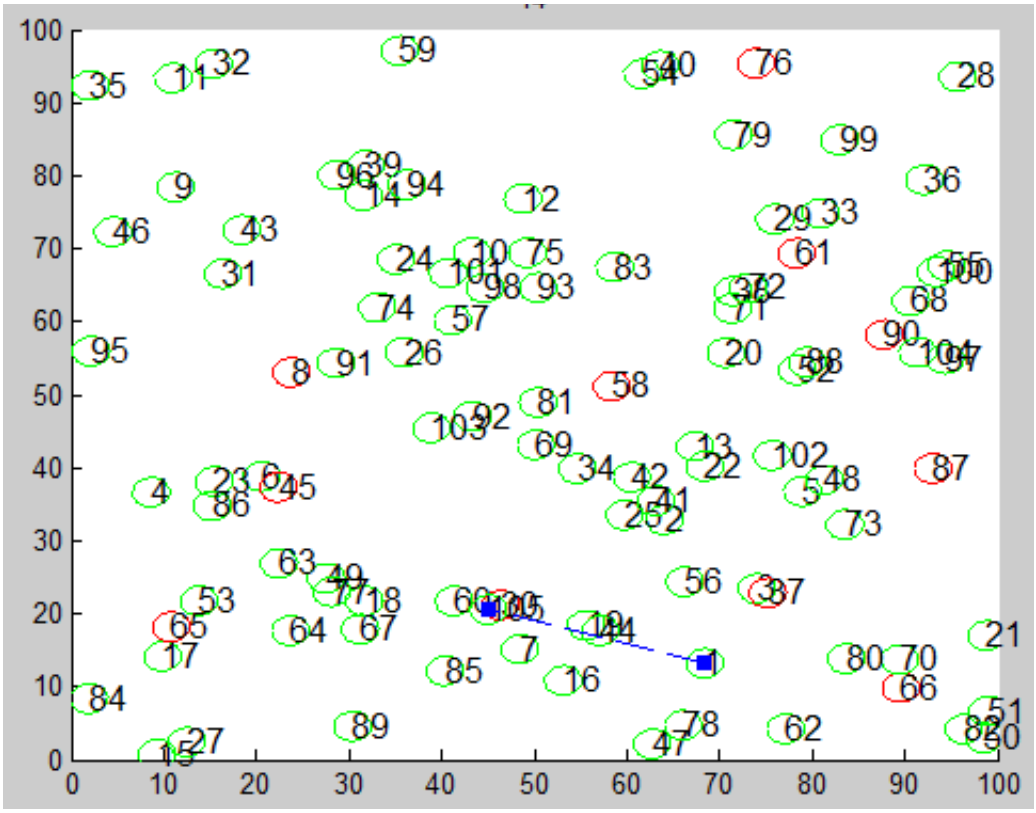


Figure 2: Designing a Sensor network with BS and attacker node

Initially all the parameters of network is set to the initial values i.e. Energy, delay , receiving ,transmitting signals etc. These parameters are updated timely according to the communication between the nodes and processing of information

Monitor Check:

In this check all the suspicious nodes are made under checking with the assignment of some of observer nodes / monitor nodes that keeps a supervision on the communication with these nodes. If found abnormally behaving ,saved for next stage check .

These checking is done by making 100s of looping or iterations so that no mistake could be left in the decision in checking the attacker one .Using of iteration has its purpose that in single transmission node may behave abnormal due any positive reason & in other iteration it could behave normal. No risk could be taken in this guessing as that attacker node has authorization rights & that node could harm in a second .

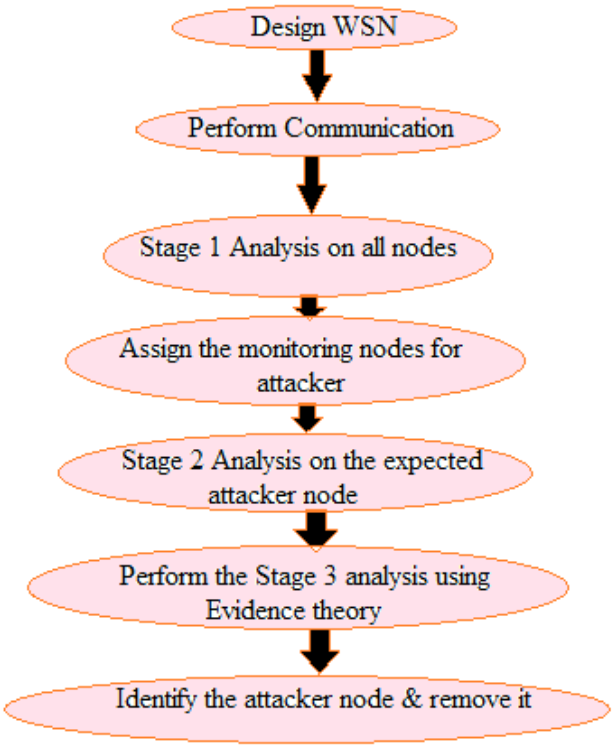


Figure 3: Working of complete algorithm

Belief Check:

The last check is Belief Check or the Evidence theory check .In this , we check the beliefs and plausibly on the certainity or probability of the attacker node present in the network . Bounded by two values, *belief* (or *support*) and *plausibility*:

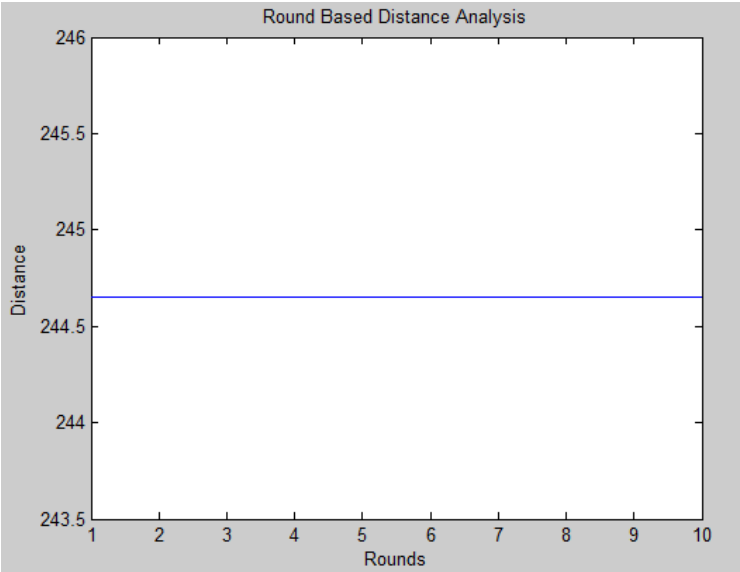


Figure 4: showing Distance covered in communication belief ≤ plausibility

Evidence Theory Checks the evidences arrived from the previous stage results and compare the results with its own values .In this check we also use the iterations for the purpose of being sure about our decision .Here two types of nodes are made in this theory –Alive nodes , Dead nodes. Alive nodes which covers the belief , Dead nodes which cancels the beliefs or does not stand with belief.

Our focus is on the alive nodes as these nodes could be attacker nodes .Hence checked with 100s of iterations giving accurate results when implemented on Matlab environment .

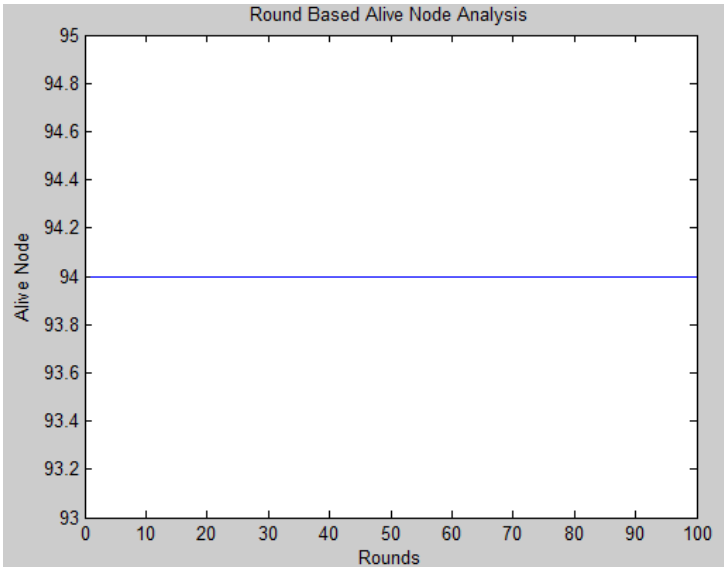


Figure 5: showing the Dead nodes list

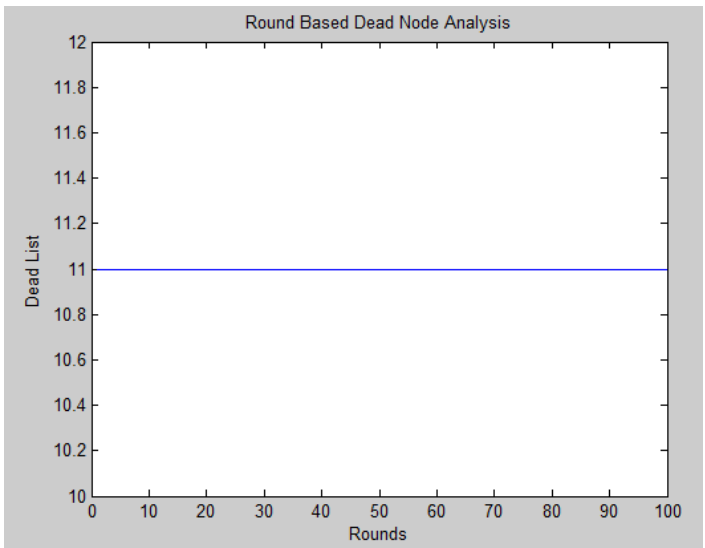


Figure 6: Showing Alive nodes list

After no. of iterations completed, this theory gives the result about the attacker nodes using parameters of energy, delays, throughput.

Conclusion:

WSN security is such a critical issue as these network gives very crucial information & are not supposed to be wrong in information. We make a algorithm for a 3 level check on the identification of the intruder node detection in case of sensor network .Previously no. of protocols are created for thesenetworks for the security but our work is not so much complex as compared to those protocols .We give a algorithm simple to implement on network. With use of evidence theory , there is scope of use of any other theory in this work .

Future Scope:

With the vast area of applications i.e. military ,medical , air services etc. these networks require more secured data service. Hence in our work other theories could be added for more safer & quick results. Another concept that could be added is the manipulation of network or acquisition of change dynamically keeping the working of security algorithm same .This flexibility is needed to be there for network to accommodate more changes with changing environment .

References:

[1] Muhammad R Ahmed,"A Novel Two-Stage Multi-Crieteria Evaluation for Internal Attack in WSN,"2013 13th International Symposium on Communications and Information Technologies (ISCIT) 978-1-4673-5580-3/13 © 2013 IEEE

[2] Guisheng Yin,"A Novel Reputation Model for Malicious Node Detection in Wireless Sensor Network,"978-1-4244-2108-4/08 © 2008 IEEE

[3] Jerzy Konorski,"Data-Centric Dempster-Shafer Theory-Based Selfishness Thwarting via Trust Evaluation in MANETs and WSNs,"978-1-4244-6273-5/09 ©2009 IEEE

[4] Wassim Znaidi,"Hierarchical Node Replication Attacks Detection in Wireless Sensors Networks," 978-1-4244-5213-4/09 ©2009 IEEE

[5] Ochirkhand Erdene-Ochir,"Resiliency of Wireless Sensor Networks: Definitions and Analyses,"2010 17th International Conference on Telecommunications 978-1-4244-5247-7/09 ©2009 IEEE

[6] Wu Yang,"Research on Reputation Evaluation Model for WSN Nodes Based on Vertical and Horizontal Analysis,"2010 Second International Conference on Networks Security, Wireless Communications and Trusted Computing 978-0-7695-4011-5/10 © 2010 IEEE

[7] Shio Kumar Singh, M. P. Singh D. K. Singh: " Applications, Classifications, and Selections of Energy-Efficient Routing Protocols for Wireless Sensor Networks", (IJAEST) INTERNATIONAL JOURNAL OF ADVANCED ENGINEERING SCIENCES AND TECHNOLOGIES Vol No. 1, Issue No. 2, 085 – 095

[8] KiranMaraiya, Kamal Kant, Nitin :Wireless Sensor Network:" A Review on Data Aggregation", International Journal of Scientific & Engineering Research Volume 2, Issue 4, April -2011 1 ISSN 2229-5518 ,IJSER © 2011

Instructions for Authors

Essentials for Publishing in this Journal

1 Submitted articles should not have been previously published or be currently under consideration for publication elsewhere.

2 Conference papers may only be submitted if the paper has been completely re-written (taken to mean more than 50%) and the author has cleared any necessary permission with the copyright owner if it has been previously copyrighted.

3 All our articles are refereed through a double-blind process.

4 All authors must declare they have read and agreed to the content of the submitted article and must sign a declaration correspond to the originality of the article.

Submission Process

All articles for this journal must be submitted using our online submissions system. http://enrichedpub.com/ . Please use the Submit Your Article link in the Author Service area.

-

Manuscript Guidelines

The instructions to authors about the article preparation for publication in the Manuscripts are submitted online, through the e-Ur (Electronic editing) system, developed by **Enriched Publications Pvt. Ltd**. The article should contain the abstract with keywords, introduction, body, conclusion, references and the summary in English language (without heading and subheading enumeration). The article length should not exceed 16 pages of A4 paper format.

Title

The title should be informative. It is in both Journal's and author's best interest to use terms suitable. For indexing and word search. If there are no such terms in the title, the author is strongly advised to add a subtitle. The title should be given in English as well. The titles precede the abstract and the summary in an appropriate language.

Letterhead Title

The letterhead title is given at a top of each page for easier identification of article copies in an Electronic form in particular. It contains the author's surname and first name initial .article title, journal title and collation (year, volume, and issue, first and last page). The journal and article titles can be given in a shortened form.

Author's Name

Full name(s) of author(s) should be used. It is advisable to give the middle initial. Names are given in their original form.

Contact Details

The postal address or the e-mail address of the author (usually of the first one if there are more Authors) is given in the footnote at the bottom of the first page.

Type of Articles

Classification of articles is a duty of the editorial staff and is of special importance. Referees and the members of the editorial staff, or section editors, can propose a category, but the editor-in-chief has the sole responsibility for their classification. Journal articles are classified as follows:

Scientific articles:

1. Original scientific paper (giving the previously unpublished results of the author's own research based on management methods).

2. Survey paper (giving an original, detailed and critical view of a research problem or an area to which the author has made a contribution visible through his self-citation);

3. Short or preliminary communication (original management paper of full format but of a smaller extent or of a preliminary character);

4. Scientific critique or forum (discussion on a particular scientific topic, based exclusively on management argumentation) and commentaries. Exceptionally, in particular areas, a scientific paper in the Journal can be in a form of a monograph or a critical edition of scientific data (historical, archival, lexicographic, bibliographic, data survey, etc.) which were unknown or hardly accessible for scientific research.

Professional articles:

1. Professional paper (contribution offering experience useful for improvement of professional practice but not necessarily based on scientific methods);
2. Informative contribution (editorial, commentary, etc.);
3. Review (of a book, software, case study, scientific event, etc.)

Language

The article should be in English. The grammar and style of the article should be of good quality. The systematized text should be without abbreviations (except standard ones). All measurements must be in SI units. The sequence of formulae is denoted in Arabic numerals in parentheses on the right-hand side.

Abstract and Summary

An abstract is a concise informative presentation of the article content for fast and accurate Evaluation of its relevance. It is both in the Editorial Office's and the author's best interest for an abstract to contain terms often used for indexing and article search. The abstract describes the purpose of the study and the methods, outlines the findings and state the conclusions. A 100- to 250-Word abstract should be placed between the title and the keywords with the body text to follow. Besides an abstract are advised to have a summary in English, at the end of the article, after the Reference list. The summary should be structured and long up to 1/10 of the article length (it is more extensive than the abstract).

Keywords

Keywords are terms or phrases showing adequately the article content for indexing and search purposes. They should be allocated heaving in mind widely accepted international sources (index, dictionary or thesaurus), such as the Web of Science keyword list for science in general. The higher their usage frequency is the better. Up to 10 keywords immediately follow the abstract and the summary, in respective languages.

Acknowledgements

The name and the number of the project or programmed within which the article was realized is given in a separate note at the bottom of the first page together with the name of the institution which financially supported the project or programmed.

Tables and Illustrations

All the captions should be in the original language as well as in English, together with the texts in illustrations if possible. Tables are typed in the same style as the text and are denoted by numerals at the top. Photographs and drawings, placed appropriately in the text, should be clear, precise and suitable for reproduction. Drawings should be created in Word or Corel.

Citation in the Text

Citation in the text must be uniform. When citing references in the text, use the reference number set in square brackets from the Reference list at the end of the article.

Footnotes

Footnotes are given at the bottom of the page with the text they refer to. They can contain less relevant details, additional explanations or used sources (e.g. scientific material, manuals). They cannot replace the cited literature.
The article should be accompanied with a cover letter with the information about the author(s): surname, middle initial, first name, and citizen personal number, rank, title, e-mail address, and affiliation address, home address including municipality, phone number in the office and at home (or a mobile phone number). The cover letter should state the type of the article and tell which illustrations are original and which are not.

Notes: