

Annexure – IX
Faculty Publications,
Patents and Awards

Bio-inspired routing protocol for wireless sensor network to minimise the energy consumption

ISSN 2043-6386
Received on 29th October 2019
Revised 29th February 2020
Accepted on 1st June 2020
doi: 10.1049/iet-wss.2019.0198
www.ietdl.org

Padmalaya Nayak¹ ✉, Ch Praneeth Reddy¹

¹Department of Computer Science and Engineering, Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad-500 090, India
✉ E-mail: padma_nayak@yahoo.com

Abstract: The minimisation of energy consumption has become an emerging topic in wireless sensor networks (WSNs) as these networks enable a wealth of new applications. The internet of things (IoT) application is one of them and the current hype around the IoT is huge. Therefore, the development of efficient communication protocols for WSNs is a major concern. In this context, various research communities have triggered several optimisation techniques to provide energy-efficient solutions to WSNs. This study aims to apply the genetic algorithm (GA) in WSNs clustering and to evaluate its performance over another optimisation technique. The proposed protocol is analytically analysed and compared with a fuzzy logic (FL)-based routing protocol and traditional routing protocol like LEACH and K-means using a Java-based custom simulator. Simulation results show that there is a trade-off between GA-clustering and FL-clustering, but the overall performance of GA-clustering is very promising for obtaining optimal energy consumption.

1 Introduction

Rapid growth in the field of internet of things (IoT) along with wireless communication has led to the development of tiny, smart sensor nodes those play a major role in implementing IoT. It facilitates billions of devices to share the data by their sensing and communicating capability. It seems like every day a new company announces some IoT-enabled product. Wireless sensor network (WSN) consists of hundreds or thousands of sensor nodes organised in an *ad hoc* pattern to observe and interact with the physical world. Each sensor node consists of four elements; sensing unit, processing and storage unit, power supply and transceivers. The sensing unit is responsible for measuring the physical parameters in the real world such as temperature, pressure, humidity, acoustic signal, vibrations, vehicular movements etc. [1]. These sensed values are handled by the processing unit and forwarded to the base station (BS) through intermediate nodes either by single-hop or multi-hop fashion. Energy consumption, limited bandwidth and limited memory is the main challenging issue in designing a protocol in WSN. Most of the time the sensor networks are deployed in unattended terrain, where recharging and replacement of battery is quite impossible. Despite of enormous

constraints, the applications of WSNs are huge in range that vary from military surveillance to health-care monitoring, agriculture, inventory control, industrial automation etc. The basic architectural model of WSN is shown in Fig. 1.

Many researchers have put their effort into designing routing protocols for WSN since last decade and proved their energy efficiency through simulation results [1–14]. Clustering-based routing technique is one of these efficient techniques, where the whole sensor network is partitioned into small size networks (clusters) to resolve the scalability issue of the network. In these networks, each cluster is controlled by an efficient node known as cluster head (CH). Whenever any event occurs, all the sensor nodes inside a cluster send the sensed data to the respective CHs. These CHs aggregate the sensed data and send it to the BS either directly or through the other CHs till it reaches the BS. The objective of the clustered-based routing protocol (equal or unequal) is purely application specific. However, the overall target is to minimise energy conservation and to overcome hot spot problems. Some of the added features which cannot be compromised are discussed below.

- **Scalability:** sensor nodes are deployed in large numbers that range from hundreds to thousands of nodes depending on the requirement of the applications in real time scenario. When the sensor nodes are large, clustered-based routing can only provide scalability in large scale by dividing the sensor field into number of levels and dividing each level into number of clusters.
- **Data aggregation/fusion:** data fusion occurs at BS level, but in clustering-based algorithm local data fusion occurs at CH level and global data fusion occurs at BS level.
- **Load balancing:** rotation of CH among all the sensor nodes is required within a cluster in order to balance the network load. Proper selection of CH must be taken care of to accomplish the objectives.
- **Fault tolerant:** as sensor networks are deployed in an open environment without human monitoring system, it must be robust enough to handle the changes in physical properties of ever-changing environment.
- **Stable network topology:** stable network topology is an essential feature to minimise the energy consumption.
- **Increased lifetime:** energy conservation is the major concern in WSN that directly lengthens the network lifetime. So, proper

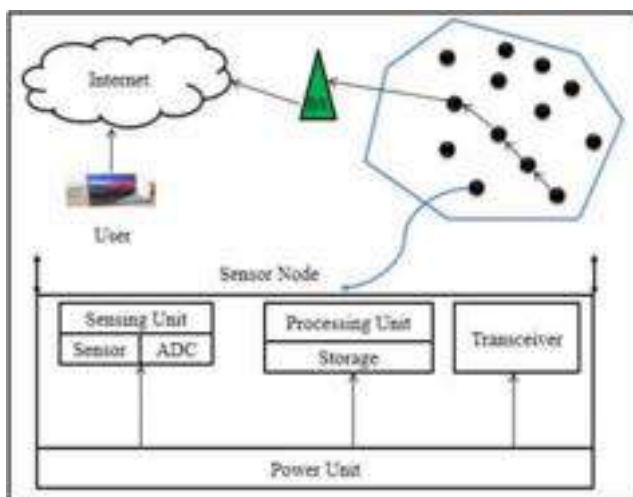


Fig. 1 Basic structure of WSN



Routing in wireless sensor networks using machine learning techniques: Challenges and opportunities

Padmalaya Nayak^{*}, G.K. Swetha, Surbhi Gupta, K. Madhavi

Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, India

ARTICLE INFO

Keywords:

WSNs
Artificial Intelligence
Machine Learning Techniques
Routing

ABSTRACT

Energy conservation is the primary task in Wireless Sensor Networks (WSNs) as these tiny sensor nodes are the backbone of today's Internet of Things (IoT) applications. These nodes rely exclusively on battery power to maneuver in hazardous environments. So, there is a requirement to study and design efficient, robust communication protocols to handle the challenges of the WSNs to make the network operational for a long time. Although traditional technologies solve many issues in WSNs, it may not derive an accurate mathematical model for predicting system behavior. So, some challenging tasks like routing, data fusion, localization, and object tracking are handled by low complexity mathematical models to define system behavior. In this paper, an effort has been made to provide a big outlook to the current "researchers" on machine learning techniques that have been employed to handle various issues in WSNs, and special attention has been given to routing problems.

1. Introduction

A WSN is a collection of a large number of sensor nodes, usually deployed in remote areas to monitor environmental parameters like temperature, humidity, moisture, etc. The sensor nodes are equipped with various types of sensors like acoustic, pressure, motion, image, chemical, weather, pressure, temperature, optical sensors, etc. Due to this diversity of sensor nodes, the applications of WSNs are huge in a range that starts with healthcare, military, defense, agriculture to our day to day life. Despite huge applications, WSN faces many typical challenges like limited energy sources, computational speed, memory, and limited communication bandwidth, making the sensor network degrade in performance and decreasing the network lifetime [1]. Developing different algorithms for different applications is quite a challenging task. In particular, the designer of WSNs must emphasize on various issues like data aggregation, clustering, routing, localization, fault detection, task scheduling, event tracking, etc. The various challenges and issues in WSNs are illustrated in Fig. 1. The complete description is given in section III. Among all the tasks, routing is one of the important tasks as major percentage of the energy consumption takes place while routing the data packet from the source node to the destination either through a single hop or multi-hop fashion. While routing the data, the sensor network designer must focus on all the sensor node's energy consumption issues to keep the network operating

for a long time. Every routing protocol has its own characteristics and specifications based on network applications and structure.

Machine Learning (ML) is a part of Artificial Intelligence introduced in the late 1950s. Over the period, it evolved and moved towards algorithms that could computationally feasible and robust enough to handle different problems like classifications, clustering, regression, and optimization in the field of medical, engineering, and computing. ML is one of the most exciting and influential technologies in today's world. ML provides computer systems with the ability to learn automatically without human involvement and take action accordingly. It creates a model by analyzing complex data automatically, quickly, and accurately. ML has the ability to learn from the generalized structure to provide a general solution to improve system performance. Because of the diversified applications, it is applied in various scientific fields of medical, engineering, and computing like manual data entry, automatic detection of spam, medical diagnosis, image recognition, data cleansing, noise reduction [144,145], etc. Recent studies prove that ML has been applied to solve many issues in WSNs. Applying ML in WSNs not only improves the system performance but also reduces the complex tasks like reprogramming, accessing the large amount of data manually, and extracting useful information from the data. So, ML techniques are extremely helpful for fetching large amounts of data and extract useful information [2–4]. For more clarity, the requirements of Machine Learning Techniques in WSNs are briefly explained in the below

^{*} Corresponding author.

E-mail addresses: padmalaya@griet.ac.in (P. Nayak), royal_surbhi@yahoo.com (S. Gupta).



A machine learning based IoT for providing an intrusion detection system for security

Dhanke Jyoti Atul^a, R. Kamalraj^b, G. Ramesh^c, K. Sakthidasan Sankaran^d, Sudhir Sharma^e, Syed Khasim^{f,*}

^a Engineering Science (Mathematics), Bharati Vidyapeeth's College of Engineering, Lavale, Pune 412115, India

^b MCA Department, School of CS & IT, Jain University, Bangalore, Karnataka, India

^c Department of Computer Science and Engineering, Gokaraju Rangaraju Institute of Engineering & Technology, Hyderabad-500090, Telangana State, India

^d Department of Electronics and Communication Engineering, Hindustan Institute of Technology and Science, Chennai, India

^e Computer Science & Engineering, Birla Institute of Applied Sciences, Bimtal-Narinal (Uttarakhand) - 263136, India

^f Computer Science & Engineering, Dr. Samuel Grady Institute of Engineering & Technology, Markapur, Andhra Pradesh, India

ARTICLE INFO

Keywords

Cyber-physical system (CPS)
Energy-aware smart home
Machine Learning
IoT
Security

ABSTRACT

Digital communication is provided with an effective communication platform to share and transfer information. The emergence of the Cyber-Physical System (CPS) is a platform incorporated with electronic devices that enables the services through a digital platform. The considerable challenges of this system are security issues, abnormality, and service failure. Hence, the requirement of providing an effective system, which should be overcome these issues. This paper analyzes these problems and providing the paradigm in terms of enhanced communication paradigms, specifically propose Energy Aware Smart Home (EASH) framework. With this work, the problem in communication failures and types of network attacks are analyzed in EASH. With the utilization of the machine learning technique, the abnormality sources of the communication paradigms are differentiated. To evaluate the performance, we analyze the proposed work based on its accuracy, performance, and efficiency. Hence, we obtain better results especially the result shows an 85% accuracy rate. In the future, we try to enhance a high accuracy rate for further development.

Introduction

The advancement in modern technology enables effective communication to every field, in specific the Cyber-Physical System is a novel platform to provide a better platform to share and transfer information from one end to another via the different communicational channel. This technology enables advancement in communication transferring platforms; therefore, the development of the economy is reached a high destination. However, security and resiliency is still challenging and should be considered for security enhancement processes. The two considerable factors [1] of communication failure are security and component failures. The hackers or intruders enable malicious activity towards the CPS system; meanwhile, the development of it is tremendous and omnipresent in modern societies. Hence, security failure is a serious topic due to the abnormal behaviors of the system; however, the implications of these are not the same as other systems. Therefore, the development of new technology is to tackle these issues in terms of

minimizing the attacks, providing security service, and protecting existing services by controlling abnormal behaviors. In this process, the parameter specification is the complex task of differentiation which needs the research on the CPS system based on its specific components [2-3], before incorporating with a holistic strategy. The modern world incorporated with tremendous development due to the emergence of technology. The technology growth provides huge benefits to the consumers; in specific, IoT is the platform for an effective communication process [4-5] that enables various devices for simplifying the difficulties in transferring information from source to source. However, the IoT has nowadays faced huge difficulties with the security issues are caused by component failures and malicious attack lead by attackers from outside. Lack of sensor nodes resources on the IoT, network complexity and communication available to the open wireless transmission is vulnerable to attack. The Intrusion Detection System (IDS) helps to identify the network of anomalies and take necessary measures to guarantee the security and dependability of the activity of IoT applications [6-7]. The

* Corresponding author.

E-mail address: khasim716@gmail.com (S. Khasim).



Support Vector Based Regression Model to Detect Sybil Attacks in WSN

R Vijaya Saraswathi¹, L Padma Sree², K Anuradha³

¹Assistant Professor, VNRVJiet, Hyderabad, India. E Mail: vijayasaraswathi_r@vnrvjiet.in

²Professor, VNRVJiet, Hyderabad, India. E Mail: padmasree_l@vnrvjiet.in

³Professor, GRIET, Hyderabad, India. E Mail: kodali.anuradha@yahoo.com

ABSTRACT

A Wireless Sensor Network (WSN) is a wireless network that includes minute sensor nodes. Sensors monitor physical and environmental conditions. WSNs are well used in military and civilian applications. Also, WSNs being deployed in an unattended area, may get inclined to different types of attacks, leading to harmful effects for the nodes. Sybil attack is one such type of a node that claims illegitimately multiple identities. Legitimate node shares data to the malicious node leading to loss of data. Hence, the proposed work attempts to secure the network with the use of Machine Learning Techniques. The proposed methodology results in a quantitative based machine learning Sybil attack detection methodology to analyze the performance of nodes.

Key words : WSN, Sybil Attack, Machine Learning, SVM, Regression.

1. INTRODUCTION

Wireless Sensor Networks (WSNs) comprise of an enormous number of sensor nodes[15], indistinctly sent over an area, cooperating to monitor and gain information about an environment. Sensor nodes are capable of collaborating with one another with the use of surrounding environment factors such as light, temperature, sound, vibration. The sensed measurements are then converted into digital signals and processed to reveal some properties of the phenomena around sensors. WSN is a wireless network that consists of base stations and many sensor nodes. Following Figure 1. shows the architecture of WSN. Sensor nodes have limited resources regarding energy, transmission range, computation, available bandwidth and memory.



Figure 1: Wireless Sensor Network

They are typically installed in a distant or hostile location and are left unattended to monitor and report.

Therefore, limited resources of nodes need to be used efficiently in order to extend network lifetime and acquire better throughput. These networks will set up in a wide range of applications like health care, environmental monitoring and military surveillance are few to mention. Hence the usage of wireless sensor networks demand the prominence on ensuring security.

The deployment and use of sensor nodes in an unattended area, with the limited resources[17] has let the nodes prone to many malicious attacks[16]. Also, in WSNs, data dissemination has to be done often, and sensor nodes need to be deployed in an arbitrarily surroundings, There is a scope that an attacker opponent can be easily attacked to a WSN. There is a possibility that an attacker may eavesdrop messages, compromise a sensor node, alter the integrity of the data, inject fake messages, and misutilize network resources.

1.1 Sybil Attack

Sybil Attack [14] is one of the most vulnerable attack among many attacks of WSN[13], in which malicious node creates a huge number of fake identities in order to gain an excessively high advantage through a byzantine method. A Sybil node using only one physical device that generates a random number of additional node, that will identify the disruption of normal functioning of the WSN with the use of multipath routing of multiple disjoint paths between source and

Optimized deep belief network and entropy-based hybrid bounding model for incremental text categorization

Optimized
deep belief
network

V. Srilakshmi
CSE, JNTUA, Anantapur, India

K. Anuradha
CSE, GRIET, Hyderabad, India, and

C. Shoba Bindu
CSE, JNTUA, Anantapur, India

Received 18 March 2020
Revised 30 May 2020
Accepted 2 June 2020

Abstract

Purpose – This paper aims to model a technique that categorizes the texts from huge documents. The progression in internet technologies has raised the count of document accessibility, and thus the documents available online become countless. The text documents comprise of research article, journal papers, newspaper, technical reports and blogs. These large documents are useful and valuable for processing real-time applications. Also, these massive documents are used in several retrieval methods. Text classification plays a vital role in information retrieval technologies and is considered as an active field for processing massive applications. The aim of text classification is to categorize the large-sized documents into different categories on the basis of its contents. There exist numerous methods for performing text-related tasks such as profiling users, sentiment analysis and identification of spams, which is considered as a supervised learning issue and is addressed with text classifier.

Design/methodology/approach – At first, the input documents are pre-processed using the stop word removal and stemming technique such that the input is made effective and capable for feature extraction. In the feature extraction process, the features are extracted using the vector space model (VSM) and then, the feature selection is done for selecting the highly relevant features to perform text categorization. Once the features are selected, the text categorization is progressed using the deep belief network (DBN). The training of the DBN is performed using the proposed grasshopper crow optimization algorithm (GCOA) that is the integration of the grasshopper optimization algorithm (GOA) and Crow search algorithm (CSA). Moreover, the hybrid weight bounding model is devised using the proposed GCOA and range degree. Thus, the proposed GCOA + DBN is used for classifying the text documents.

Findings – The performance of the proposed technique is evaluated using accuracy, precision and recall is compared with existing techniques such as naive bayes, k-nearest neighbors, support vector machine and deep convolutional neural network (DCNN) and Stochastic Gradient-CAViaR + DCNN. Here, the proposed GCOA + DBN has improved performance with the values of 0.959, 0.959 and 0.96 for precision, recall and accuracy, respectively.

Originality/value – This paper proposes a technique that categorizes the texts from massive sized documents. From the findings, it can be shown that the proposed GCOA-based DBN effectively classifies the text documents.

Keywords Incremental learning, Bounding model

Paper type Research paper



Predictive Method for Diabetic Medical Records Data Analysis Using Machine Learning and Hadoop

¹T. Ajay kumar, ²Dr. B. Sankara Babu

¹PG Scholar, ²Professor

^{1,2}Department of Computer Science and Engineering

^{1,2}Gokaraju Rangaraju Institute of Engineering and Technology

¹ajaykumarthallapalli55@gmail.com, ²sankarababu.b@griet.ac.in

Abstract

Presently days from social insurance businesses huge volume of information is creating. It is important to gather, store and procedure this information to find information from it and use it to take huge choices. Diabetic Mellitus (DM) is from the Non Communicable Diseases (NCD), and loads of individuals are experiencing it. Presently days, for creating nations, for example, India, DM has become a major medical problem. The DM is one of the basic diseases which has long haul difficulties related with it and furthermore pursues with different medical issues. With the assistance of innovation, it is important to fabricate a framework that store and break down the diabetic information and predict potential dangers likewise. Predictive investigation is a strategy that incorporates different information mining systems, ML algorithms and measurements those utilization present and past informational collections to pick up understanding and predict future dangers. In this work “machine learning calculation in Hadoop MapReduce environment are executed for Pima Indian diabetes informational index to discover missing qualities in it and to find designs from it. This work will have the option to predict kinds of diabetes are far reaching, related future dangers and as per the hazard level of patient the sort of treatment can be given”.

Keywords: Healthcare industry, Hadoop, MapReduce, ML, Predictive Analysis

Introduction

Predictive examination which is help to human services associations to assess information on the past conduct and predict probability of future conduct to empower better choices and results of their patient[1]. Predictive models can settle on human choices increasingly viable and profoundly computerize a whole basic leadership process. It progressively, predictive examination utilizes information from the IOT to improve wellbeing and execution of patient results. Medicinal services industry faces numerous provokes that make us to realize the significance to build up the information investigation of the diabetes mellitus.

BigData is developing as an answer for the issues related with enormous measure of information. The huge measure of information produced would now be able to be utilized so as to give an internal perspective on what is truly occurring and recognize the developing patterns. Large Data can likewise be utilized in the field of medicinal services so as to make the framework increasingly powerful. Huge Data alludes to the enormous measure of information which might be organized or unstructured and can't be handled utilizing a social database model. Unstructured



Stochastic gradient-CAViaR-based deep belief network for text categorization

V. Srilakshmi¹ · K. Anuradha² · C. Shoba Bindu¹

Received: 25 November 2019 / Revised: 24 March 2020 / Accepted: 2 July 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract

Text categorization is defined as the process of assigning tags to text according to its content. Some of the text classification approaches are document organization, spam email filtering, and news groupings. This paper introduces stochastic gradient-CAViaR-based deep belief networks for text categorization. The overall procedure of the proposed approach involves four steps, such as pre-processing, feature extraction, feature selection, and text categorization. At first, the pre-processing is carried out from the input data based on stemming, stop-word removal, and then, the feature extraction is performed using a vector space model. Once the extraction is done, the feature selection is carried out based on entropy. Subsequently, the selected features are given to the text categorization step. Here, the text categorization is done using the proposed SG-CAV-based deep belief networks (SG-CAV-based DBN). The proposed SG-CAV is used to train the DBN, which is designed by combining conditional autoregressive value at risk and stochastic gradient descent. The performance of the proposed SGCAV + DBN is evaluated based on the metrics, such as recall, precision, F-measure and accuracy. Also, the performance of the proposed method is compared with the existing methods, such as Naive Bayes, K-nearest neighbours, support vector machine, and deep belief network (DBN). From the analysis, it is depicted that the proposed SGCAV + DBN method achieves the maximal precision of 0.78, the maximal recall of 0.78, maximal F-measure of 0.78, and the maximal accuracy of 0.95. Among the existing methods, DBN achieves the maximum precision, recall, F-measure and accuracy, for 20 Newsgroup database and Reuter database. The performance of the proposed system is 10.98%, 11.54%, 11.538%, and 18.33% higher than the precision, recall, F-measure, and accuracy of the DBN for 20 Newsgroup database, and 2.38%, 2.38%, 2.37%, and 0.21% higher than the precision, recall, F-measure and accuracy of the DBN for Reuter database.

Keywords Text categorization · Deep belief network · Stochastic gradient descent · CAViaR · Vector space model

1 Introduction

One of the most important processes in the knowledge discovery process is text mining [1]. As most of the information on the internet is unstructured, hence it is appropriate to regulate the data process text mining-based algorithms. Several intelligent algorithms, like neural networks, case-based reasoning and probability reasoning and combination of text processing technology, have been proven effective in the text mining process. The text mining algorithms regulate the unstructured document and help in extracting the key

concept and the relationship among the characters. The text mining algorithms help in text classification by accessing useful knowledge and information from the database [2]. The documents present online have improved, due to the growth of the internet. Regularly text documents contain research articles, blogs, journal papers, and newspapers, and so on. This vast number of documents may be valuable and useful [3, 4]. Documents are generally denoted by “bag-of-words”, such as every phrase or word present in documents once or several times took as the feature. For the given data set, a collection of entire phrases or words forms the “dictionary” with several features. To ameliorate the “curse of dimensionality” problem and to boost up the learning process, it is necessary for performing feature reduction to mitigate the feature size [5].

TC is utilized for finding relevant categories from the given text. It is the core method for several applications,

✉ V. Srilakshmi
srilakshmiv064@gmail.com

¹ CSE, JNTUA, Anantapur, India

² CSE, GRIET, Hyderabad, India



RESEARCH ARTICLE



Intelligent signalling system to control traffic in vehicular ad hoc networks

R V S Lalitha^{1*}, R Srinivas², Pullela SVVSR Kumar³, K Kavitha⁴,
P V S N S Sameera⁵

¹ Professor, Department of C.S.E, Aditya College of Engineering & Technology, Surampalem, 533 437, India

² Professor, Department of C.S.E, Aditya Engineering College, Surampalem, 533 437, India

³ Professor, Department of C.S.E, Aditya College of Engineering, Surampalem, 533 437, India

⁴ Associate Professor, Department of C.S.E, Gokaraju Rangaraju Institute of Engineering & Technology, Hyderabad, 500 090, India

⁵ Software Engineer, TCS, Bangalore, India

OPEN ACCESS

Received: 16-07-2020

Accepted: 25-07-2020

Published: 07-08-2020

Editor: Dr. Natarajan Gajendran

Citation: Lalitha RVS, Srinivas R, Kumar PSVSR, Kavitha K, Sameera PVSNS (2020) Intelligent signalling system to control traffic in vehicular ad hoc networks. Indian Journal of Science and Technology 13(28): 2876-2882. <https://doi.org/10.17485/IJST/v13i28.1043>

* **Corresponding author.**

rslalitha@gmail.com

Funding: None

Competing Interests: None

Copyright: © 2020 Lalitha et al.. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Published By Indian Society for Education and Environment (iSee)

ISSN

Print: 0974-6846

Electronic: 0974-5645

Abstract

Traffic congestion problem persists more at junctions and causes inconvenience to public. Owing to this, people may not reach their destinations in time. Although there are alerts in mobile phones regarding arrival times of flight services and other journey related information due to sudden traffic jams at junctions these advanced technology innovations are being becoming fragile. Since long, many researchers have been putting their efforts to find remedies for reducing traffic congestion. **Objectives:** The key focus is on balancing forwarding time and waiting time at junctions based on the number of vehicles arrived at that junction. **Methods:** The proposed system uses Internet of Things (IoT) based monitoring to control signaling system. IR sensors are used to count number of vehicles passing over the lane by triggering clock for object detection. The signaling time can be changed dynamically based on the vehicle count, so that more time is allocated to the lanes which have more traffic and the remaining time is adjusted among other lanes. This reduces congestion at dense traffic lanes. **Findings:** There will be time variant setting of signal lights based on the density of the traffic. The dense traffic lanes will be allotted more time and will be cleared first. Expanding on this point, the time adjustment is done based on the vehicle count not on periodical fixation of intervals. **Novelty:** The ecosystem developed provides an ultimate solution to vehicle users for comfortable movement on the roads without being delayed.

Keywords: Traffic density; IR Sensor; VANET; ITS; IoT; clock

1 Introduction

The exponential increase in the number of vehicles leads to transportation issues. Due to flooding of vehicles and signaling intervals vehicles are forced to wait at junctions. The performance of the road clearing mechanisms is to be improved to avoid roundabout travels. Research was being worked on finding optimal path. The optimal

A Framework for Subset Pruning using REP Tree

¹Dr.R.V.S.Lalitha, ²Chaallapalli Sujana, ³Dr.K.Kavitha, ⁴Mylavarapu Kalyan Ram

¹Professor, Department of C.S.E, Aditya College of Engineering & Technology,
Surampalem,India,

²Assistant Professor, Department of C.S.E, Aditya College of Engineering & Technology,
Surampalem,India,

³Associate Professor, Gokaraju Rangaraju Institute of Engineering & Technology,
Hyderabad,India,

⁴Sr. Assistant Professor, Department of C.S.E, Aditya Engineering College (A),
Surampalem,India.

Corresponding author:rvslalitha@gmail.com

Abstract: Basically the count of confirmed cases depends upon number of tests so far carried out. The rate of confirmed cases relies on number of test carried out so far. Based on the figures arrived, the Government organization will impose precautionary measures. In case of patient data analysis for any pandemic disease decision making plays crucial role. If the test reports are inaccurate, underreporting takes place. The reports should be provided on time so as to safeguard the lives of patients. Identification of misclassified examples is a well known problem and is drawing significant attention in health monitoring units. By and by exactness is the primary concern for evaluation and treatment of the individuals. In order to categorize correct classification labels, two typical algorithms are considered, ZeroR and RIPTree. The ZeroR classifies all the instances with majority of the labels without including any predictors. Since the RIPTree is less prone to error, it provides correct information about misclassified instances. The evaluation metrics can be streamlined with the adaptation of these two algorithms. The two algorithmic outcomes can be cross verified with Repeated Incremental Pruning to produce error reduction (RIPPER) algorithm which classifies true positives exactly. The application of the above algorithms assists in confirming the cases, with varied conditions and datasets.

Keywords: ZeroR, RIPTree, RIPPER, Classification, Pruning, Rules.

I. Introduction

Corona virus is being spread globally since December. No proper diagnostics is available either for curing or for confirming the cases. As viral tests evaluate current infection of the individuals, the analysis of report for confirmed/unconfirmed is playing crucial role. Finding outstanding classifier is a typical task, because the algorithm must support large datasets at par with accuracy. Researchers need to drill down their data exploration mechanisms by applying

Applications of Raspberry Pi and Arduino to Monitor Water Quality Using Fuzzy Logic



Padmalaya Nayak, Chintakindi Praneeth Reddy and Devakishan Adla

Abstract Water is one of the most important natural resources that must be monitored, analyzed continuously for safety and survival of human life. The traditional method relies on collecting water samples, testing, and analyzing the water samples in specific laboratories which is not only cost effective but also causes access latency, and delay in disseminating the information among the end users. The huge growth of wireless technology and VLSI design has brought a tremendous change in developing small microsensors that are being utilized for various monitoring applications since the last decade. In this paper, an effort has been made to measure the drinking water quality with less cost in a hardware platform with the help of some water-related Sensors, Raspberry Pi, and Arduino Microcontroller. The proposed method utilizes the Fuzzy Logic algorithm and the experimental result shows that the proposed method has many more advantages over traditional systems. It is also observed that the proposed system works effectively in a real-time environment with immediate response and less cost.

Keywords WSN · Raspberry Pi · Arduino · Fuzzy logic

1 Introduction

The WSNs applications provide many challenges even though these sensor nodes are very tiny, battery operated and can be deployed randomly or deterministically to monitor the environmental parameters. The applications are huge that ranges from military, civil, health care, agriculture, disaster hit areas, water quality, and many

P. Nayak (✉) · C. P. Reddy
CSE Department, Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad
500009, Telangana, India
e-mail: padma_nayak@yahoo.com

C. P. Reddy
e-mail: chpraneeth1539@gmail.com

D. Adla
CSE Department, Vidya Jyoti Institute of Technology, Hyderabad 500009, Telangana, India

© Springer Nature Singapore Pte Ltd. 2020
V. K. Solanki et al. (eds.), *Intelligent Computing in Engineering*,
Advances in Intelligent Systems and Computing 1125,
https://doi.org/10.1007/978-981-15-2780-7_17

Constrained Clustering Approach in Wireless Sensor Networks to Minimize the Energy Consumption

Ramamurthy Garimella, *MEC, Hyderabad, India, E-Mail: rama.murthy@mechyd.ac.in*
Padmalaya Nayak, *GRIET, Hyderabad, India, E-Mail: drpadmalaya2010@gmail.com*
Mohammed Nazeer, *MCET, Hyderabad, India, E-Mail: nazeer8584@gmail.com*
K. Ramesh babu, *VJIT, Hyderabad, India, E-Mail: krameshbabucse@vjit.ac.in*

Abstract—The applications of Wireless Sensor Networks (WSNs) are massive in the range that starts from the defense, civil, military to our day to day routine life. Energy conservation in WSNs is a crucial issue to make the network operational for a long time. It is entirely worth noting that most of the energy consumption occurs during the data transmission. Routing protocol plays a significant role in carrying and forwarding the data to the destination. Clustering is one such branch of routing technique that increases the Network Lifetime unexpectedly through minimal energy consumption. In this regard, this research paper aims to formulate an exciting clustering problem in WSNs where the coordinates of patterns are random variables. The clustering problem is solved using the CENTROID approach, such as Cluster Head (CH) selection is done with an optimization interpretation. The proposed protocol is verified through the ns-2 simulator, and the results obtained from the simulations show that the proposed constrained clustering LEACH (CC-LEACH) offers its superiority over the existing protocols such as LEACH and LEACH-C in terms of various performance parameters.

Index Terms—WSN, Mobility, Clustering, Hessian Matrix

I. INTRODUCTION

In the research area of pattern recognition, data mining, and many other human endeavor fields, pattern vectors are grouped into classes based on the proximity of metrics. Various clustering algorithms, such as the fuzzy clustering algorithm, grid-based algorithm, k-means algorithm, and many more, have been proposed in the current literature. These algorithms are found to be of practical utility in clustering under the unsupervised learning algorithm (commonly employed for classification problems). There has been extensive research effort on developing the communication protocols in WSNs in the last few years. As WSNs operate in hazardous environments, the replacement/recharging of battery in the sensor field is not possible in many practical applications. Thus, the strategy of the protocols/algorithms in such networks must be energy-aware. Clustering is the most popular technique in a hierarchical system consisting of hundreds or thousands of sensor nodes. Instead of each node sending the data to the base station, only cluster head (CH) can transmit and receive

the data to the base station, which reduces energy utilization and increases the network lifetime. LEACH and other energy-efficient protocols select the CH by exchanging the hello messages among each other within a cluster ensuring that cluster head is placed at the centroid (that happens by minimizing the squared distances from centroid/CH to sensor nodes in a cluster). The research work discussed in [1] uses inter clustering to transmit the data from one CH to another CH until it reaches the base station. The research in [2] [3] discusses Cluster Based Topology (CBT), where clustering is widely used in a static and mobile environment to perform various tasks in different applications. The author in [4] discusses the Low Energy Adaptive Clustering Hierarchy (LEACH) protocol as one of the oldest hierarchical routing protocols in which the sensor nodes communicate within its cluster. The work in [5] proposes a modified LEACH protocol called Low Energy Adaptive Clustering Hierarchy Centralized (LEACH-C), where the BS executes the clustering process in a centralized manner to minimize energy consumption. Hybrid Energy-Efficient Distributed Clustering (HEED) protocol [6] presents a CH selection procedure by considering node degree, residual energy, and relative distance. The elected CH is found to be more realistic and more energy-aware. The research works in [7-20] discuss many potential clustering techniques to make the sensor networks energy efficient.

Despite intensive research efforts on clustered-based routing in WSNs, a question remained unanswered is the interpretation of CENTROID computation. In this research paper, we address and solve the problem of the optimization interpretation of CENTROID. We also address and solve an interesting, constrained clustering problem. Further, we consider the case where the components of pattern vectors contain random variables. With such random pattern vectors, we formulate and solve the clustering problem of potential practical interest. This research paper focuses on the quadratic optimization problem to select the CH by considering the distance between the sensors and many other factors. The cluster head is determined to be CENTROID of sensor position coordinates. As far as our knowledge is concerned, the derivation of CENTROID patterns is an exciting optimization problem and is not addressed in the literature.

Analysis of Big Data in Healthcare and Life Sciences Using Hive and Spark



A. Sai Hanuman, R. Soujanya and P. M. Madhuri

Abstract Big data is a declaration used to recognize the database whose area is afar the potential of typical database software tools to store, organize and examine. Big data has shown a new path toward the mankind. With several theoretical and technological obstacles in health huge processing, it is onerous to transfer knowledge into fortunate and valuable applications. Meeting the challenge of handling big data in healthcare information construction procedure, this paper proposes a referential architecture on the Hive and Spark platform to overcome the problems in healthcare big data process. Hive is a noteworthy project as a result of it permits exposing the simplest components of Hadoop, specifically map reduce and knowledge storage. Spark may be a memory-based computing framework that features a higher ability of computing and fault tolerance, supports batch, interactive, iterative and flow calculations. Experiment results of data upload, data query and data analysis show that the performance of the proposed framework is greatly improved, and a brief summary of the performance and the differences between two methods of Hive and Spark is also discussed.

Keywords Big data · Hadoop · Healthcare · Hive · Spark and Scala

A. Sai Hanuman · R. Soujanya · P. M. Madhuri (✉)
Department of Computer Science and Engineering, GRIET, Bachupally, Hyderabad,
Telangana, India
e-mail: pmmadhuri18@gmail.com

A. Sai Hanuman
e-mail: a_saihanuman@hotmail.com

R. Soujanya
e-mail: soujanya96@gmail.com

Smart Privacy Protection for Location-Based Services using Queueing Modelling

Chamana H Ram Sai Jeetesh^{1,*}, Yerubandi Sai Sriram², Kavitha Kayiram³, V.N. Rama Devi⁴

^{1,2} Student, Department of CSE, GRIET, Hyderabad, India

³ Associate Professor, Department of CSE, GRIET, Hyderabad, India

⁴ Associate Professor, Department of Humanities and Basic Sciences, GRIET, Hyderabad, India

Abstract. In the digital era, we are greatly dependent on the popular applications of the Location Based Services (LBS) in our day-to-day activities. The smart phone comes with a variety of applications which acquire the user location and build up user profile like the user activities, hobbies, places of visit, food orders etc. Such sensitive information in the LBS server can pose privacy risk for the user. To safe guard the user from such threat we propose a smart privacy protection technique in this paper that can conceal the user location when using the location based services. We adopt the generation of dummy locations to obfuscate the user original location from the LBS server. The server generates the result set for the dummy user locations. In this work we try to optimize the things at server as well as user ends with two objectives. The first goal is to work towards identifying the overlap in result sets and generate unique and reduced result set with which the communication load on the network can be reduced. The second goal is to prioritize the result set by Queueing model for the result set through which waiting time of the customer can be minimized. We have also illustrated that this model show good performance in terms of the reduced communication load through experimental results.

1 INTRODUCTION

With ever growing digital population, internet connectivity in every device and uninterrupted service the Location Based Services (LBS) [1] are in high use. The advent of smart technology in digital lives have made drastic change in the marketing, services and communication industry. With LBS, we are able to navigate the geography and also avail many other services at the tip of the finger. The smart phones are resource rich and able to offer many of its features based on user location information. The location based services are able to cater to the user requirements like near-by hospital, restaurant, fuel filling station etc. The service providers are continuously monitoring the user requests and their device location for their business as well many third party servers are also gathering user data and them to advertisers. This poses a threat for the user in sharing the location. Often based on the user location and services availed information the user movement profile can be built which poses a privacy risk. Also many untrustworthy service providers will be tracing the profiles for some statistical analysis. Of course location based services are worthy, but at the same time user privacy is also at stake. Without having to compromise on location privacy, still the LBS can be used to the full extent by giving user privacy in sharing the location.

Hence, many approaches were proposed in the literature for privacy preserving in LBS.

The geographical boundaries we use to identify a location are the latitude and the longitude coordinates. Similarly, Global Positioning System (GPS) is used to identify a device location and for tracking the movement of the device with the help of network infrastructure. First, the mobile phone is able to detect the latitude and the longitude coordinates of the current location. This is sent to the LBS requesting for service details. The service provider will send a reply based on the details received in the request. This customized service is obtained from the service provider in reply. The service provider has the information regarding the home, office locations, health conditions, political views, banks, and many other details about the person. This can pose a crucial security threat for the person using location based services. Therefore user privacy needs to be protected.

In this Paper, we propose a smart privacy protection technique that can hide the original geographical location of the user from the LBS server. With this scheme, neither the attacker will be able to guess the true location of the user nor will the attacker ever understand about the dummy locations and their behaviour. As the first step towards obfuscation of the user original location, a set of dummy locations around the user location are generated within a given cloaking radius and the region is termed as a cloaking region. These set of

* Corresponding author: jeetesh.egnr@gmail.com

A Survey on Hybrid Machine Translation

Anusha Anugu¹, Gajula Ramesh²

¹PG Scholar, Department of Computer Science and Engineering, GRIET, Hyderabad, Telangana, India.

²Associate Professor, Department of Computer Science and Engineering, GRIET, Hyderabad, Telangana, India.

Abstract—Machine translation has gradually developed in past 1940’s. It has gained more and more attention because of effective and efficient nature. As it makes the translation automatically without the involvement of human efforts. The distinct models of machine translation along with "Neural Machine Translation (NMT)" is summarized in this paper. Researchers have previously done lots of work on Machine Translation techniques and their evaluation techniques. Thus, we want to demonstrate an analysis of the existing techniques for machine translation including Neural Machine translation, their differences and the translation tools associated with them. Now-a-days the combination of two Machine Translation systems has the full advantage of using features from both the systems which attracts in the domain of natural language processing. So, the paper also includes the literature survey of the Hybrid Machine Translation (HMT).

1. Introduction

The process of retrieving and evaluating the information from the document repositories is known as "Information Retrieval (IR)" [1]. The user who needs information has to send request in the form of a query in natural language. Then the information related output will be retrieved from the IR system. The process of IR system [2] is as shown in figure 1.

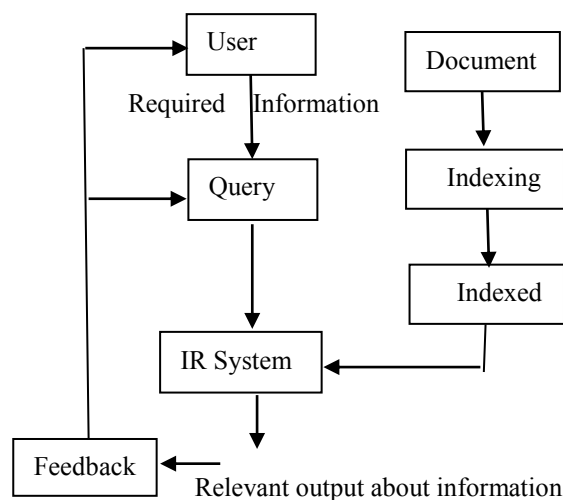


Fig. 1. Information Retrieval Process

Now-a-days, searching information in different languages has been increased rather than original language which creates a problem in IR system. Then the translation has evolved.

The domain of research work of 'Natural Language Processing (NLP)' which fulfils the interaction among the distinct classification of the nation is known as "Machine Translation". As the human made translation is expensive and time taking process, MT system is used which reduces the time and cost. MT is an automated application used by the computer to translate one language into other. In 1940's the research in MT's has been started. It is more advantageous to the industries for consumer maintenance, increasing the capacity for the accomplished translators.

2. Machine translation techniques

There are four main techniques in Machine Translation [2]. They are

2.1 Direct Machine Translation:

In initial days this type of translation is used. It translates word after word with a few word-order adjustments. It depends on dictionary look-up. Without the analysis of internal structure and grammatical correlation, the source sentence is morphologically analyzed to derive target sentence.

This process involves three steps

1. Morphological Analysis —root words are extracted from the words in source language.
2. Dictionary Lookup —searches for the matching words for target language words.

A Survey on NLP based Text Summarization for Summarizing Product Reviews

Ravali Boorugu
Assistant System Engineer
Tata Consultancy Services Limited
Hyderabad.
ravali.boorugu@gmail.com

Dr. G. Ramesh
Associate Professor
Department of CSE
GRIET, Hyderabad.
ramesh680@gmail.com

Abstract—No one can imagine life without a smartphone and internet nowadays. It has become essential for people of all age groups. With an increase in the usage of internet and smartphones, there has been a steady increase in online shopping too. Everyone wishes to get their products delivered at their home without any hassle. How to detect which products are genuine and pick the best among the unlimited options at the same price? Every user looks at the reviews before ordering anything online. Nevertheless reading those long reviews is not easy for everyone. Therefore, there must be something that can reduce the long reviews to short sentences of limited words depicting the same meaning. Text Summarization can come in hand in this aspect. Many NLP researchers are interested in Text Summarization. This paper is a survey on the various types of text summarization techniques starting from the basic to the advanced techniques. According to this survey, seq2seq model along with the LSTM and attention mechanism is used for increased accuracy.

Keywords— Text Summarization, Product Review Summarization, NLP Techniques.

I. INTRODUCTION

There has been a continuous increase in the number of internet users every year. With an increase in Internet users, comes a great deal of information that gets stored online every second. There is a need for summarizing this data without losing the original meaning of the data. Thus the process of Text

Summarization comes into the picture with its benefits spread over different fields such as Machine Learning, Natural Language Processing, Artificial Learning, Semantics etc.,

Online Shopping has become a common thing these days as a wide variety of products are available at a single place. The ease of ordering a product and getting the product delivered directly to home at a convenient date and time has attracted many people. Along with these, the discount offers being offered by online shopping sites are making the people stick to online ordering. Everyone refers to the product reviews before buying a product. Then they can conclude which is the best product to buy among the different products available. Suppose a user needs to buy a laptop. Then he must go through different kinds of laptops available at his budget, make a note of different reviews for each product and choose the best among the available laptops. This is a tedious and time taking task. In addition to this, some users' reviews are so long that the user could get the actual meaning of it only after closely going through the review. Thus there is a need for minimizing the review to a shorter representative sentence which depicts the same meaning as the whole content.

This is the area where Text Summarization comes into picture with a great deal of benefits that could help us in choosing the best product from the whole lot. Text Summarization methods are broadly categorized into different types as shown in the figure below.

A Machine Learning Approach for Air Pollution Analysis

Dr.R.V.S.Lalitha¹, Dr.K.Kavitha², Y.Vijaya Durga³, K.Sowbhagya Naidu⁴, S.Uma Manasa⁵

¹Professor, Dept. of C.S.E, Aditya College of Engineering & Technology, Surampalem,

²Associate Professor, Dept. of C.S.E, Gokaraju Rangaraju Institute of Engineering & Technology, Hyderabad

³Assistant System Engineer TCS, Hyderabad

⁴Alumni, Aditya College of Engineering & Technology, Surampalem,

⁵Associate Consultant, Syntel.

Abstract--- Air pollution occurs due to presence of harmful gases like dust and smoke. Inhaling these gases leads to health problem. The inhaling of dust leads to breathing problems and lung issues which are of major concern in human life. The green house gases like synthetic chemicals present due to emission of human activities. The major green house gases are Carbon dioxide, chlorofluorocarbons, water vapour, ozone, methane and Nitrous Oxide. Greenhouse gases absorb infrared radiation. Air pollution is monitored by Governments and various local agencies. The prime responsibility of the proposed system is to detect the concentrations of major air pollutant gases that are present in the air which cause harm to humans. Air Pollution Detection is developed using IoT for detecting pollutant gases using MQ2 and MQ135 sensors. The gases like Carbon Dioxide (CO₂), Carbon Monoxide (CO), Ethyl Alcohol, Nitric Oxide, Nitrogen Dioxide (NO₂) and Sulfur Dioxide (SO₂) will be detected using these sensors. The detected parameters are analyzed using Machine Learning (ML) algorithms to estimate air quality. The ecosystem developed helps in learning correlation among gases. This helps in estimating impact and level of air pollutants to measure air quality.

Keywords---Air pollution, Internet of Things, MQ2, MQ135 and Correlation.

I. INTRODUCTION

In order to protect our Environment we need to monitor the pollutant gases. Several epidemiological studies have been taken up for formulation of air pollution analysis. Air pollution is one of the factors that cause deaths from lung cancer and respiratory diseases. Air pollution may have direct impact on adult deaths. World Health Organization (WHO) says that air pollution has impact on environment related deaths. Air pollution analysis assists in reducing mortality risk of individual death rate. The prediction analysis is done using ML to take preventive measures to increase life span of individuals. IOT based Air Pollution Detection measures presence of dangerous gases in the air. The ecosystem is

developed using Raspberry PI 3, MQ2 and MQ135 sensors. WiFi module is used to upload detected parameters to ThingSpeak cloud for storing them in database. A web server is used to deploy data from ThingSpeak for analyzing it through ML. The paper is organized into four sections. Section 1 stated the significance of air pollution detection. In Section 2 the previous research works are discussed. The design and implementation details are given in Section 3. And the Machine Learning based analysis is done in Section 4. Conclusion states the role and impact of air pollution analysis for protecting human lives.

II. RELATED WORK

Diffusion tubes are usually used to monitor air pollution. They are made of plastic material with a rubber stopper attached at each end. These are designed for detecting NO₂. They are large in size and are not efficient. The materials need to perform sampling process are diffusion tubes, tube holders, survey sheet, maps, clip board, Re-sealable samples bag1 and a pen. Initially, diffusion tubes are located in a specific area which is divided into grids. Then position tubes are positioned vertically downwards and cable ties are attached if it is fixing to a pipe. Fix sample in a specific location where free circulation of air around the tube. Remove the white cap allow it for exposure. Fill date and times in the record sheet. Note the tube condition, changes in site conditions, or anything that might affect the results. D.Arun Kumar(2018) presented IoT based eco system to estimate quality of the air using various sensors like gas, temperature, humidity, rain and smoke. By

Energy Efficient Data Retrieval in Wireless Sensor Networks for Disaster Monitoring Applications

Kavitha Kayiram

Dept. of CSE,

Gokaraju Rangaraju Institute of Engineering and Technology,

Hyderabad, India

kavitha.bits@gmail.com

Dr. Avinash Sharma

Dept. of CSE,

M.M. Engineering College,

Maharishi Markandeshwar (Deemed to be University)

Mullana, Ambala, India.

asharma@mmumullana.org

Dr. P.Chandra Sekhar Reddy

Dept. of CSE,

Gokaraju Rangaraju Institute of Engineering and Technology,

Hyderabad, India

pchandureddy@yahoo.com

Dr. R.V.S.Lalitha

Dept. of CSE,

Aditya College of Engineering and Technology,

Kakinada, India.

rvslalitha@gmail.com

Abstract—Occurrence of natural disasters is a major concern for people residing in disaster prone areas and this increases the need for Disaster Monitoring. Wireless Sensor Network (WSN) technology offers a better solution in disaster monitoring to facilitate effective rescue operations. WSNs have the capability of quick sensing and transmission of critical data and thus prove effective in reliable monitoring and rescue operations. As the networks are deployed in disaster prone regions, the power, storage resources are limited and data retrieval is the major concern. Many schemes like sleep-scheduling, collaborative sensing have been implemented to solve these issues. But, there are many limitations to enhance the network lifespan, network traffic and effective data retrieval associated with it. In this paper, we propose a novel collaborative sensing scheme along with duty cycling to enhance the lifetime of the network. Our mathematical evaluation prove the energy efficiency of our scheme for disaster applications in WSN.

Keywords—Wireless Sensor Network, Collaborative Sensing, Disaster Monitoring System, Sleep-scheduling, Duty Cycling.

I. INTRODUCTION

Disasters occur in the nature for various reasons but the consequence of a disaster disturb natural vegetation, human life causing heavy loss. Hence, it is desirable to detect the disaster and warn early to plan timely evacuation of the space and save human life. However some disasters give a late warning with high intensity such as Tsunami, floods etc. and there is very less time for precautions and safety measures. These early warning systems generate alarm signals to indicate the dangers associated with disasters. It is highly desirable for rescue system to get the precise location to launch timely rescue operations. Hence, timely reporting and data backup during the event is important for reducing the victims and damages in incidents. Another major issue is that the disaster may sometimes lead to damage in the network both in terms of hardware and communication medium. This makes it difficult to obtain information about the

event and to plan an effective data backup from the network. The Wireless Sensor Network (WSN) [1] offers a better solution to the afore-mentioned scenarios with their cutting edge technology. Typical applications of WSN in disaster monitoring systems include Mine Fire warning system, Earth-Quake Detection system [2], Land-slide Detection system, Volcano monitoring system, Fire detection system etc.

The WSN technology can be used in disaster monitoring to facilitate early warning mechanism. In general the sensor nodes of a WSN are deployed in the region prone to disaster for continuous monitoring and event detection. The sensor nodes monitor the physical phenomenon and generate an early warning in case it senses any unusual behavior in the environment. The sensor nodes are meant to detect certain parameters in the environment like humidity, pressure, temperature, sound intensity, etc. Depending on the application, these sensor nodes are deployed in the disaster prone region to monitor the environment continuously. Each sensor node is equipped with a battery as power source and it consumes this energy to perform sensing, transmission and reception operations in a continuous manner. The lifetime of the battery is limited. As the networks are deployed in remote geographic locations and it is impossible to recharge or replace battery if battery is completely exhausted. We need to conserve the battery power by minimizing the data transmission activity in the network. Hence we propose to use duty cycling in WSN to minimize the data transmissions in the network while also catering to the goals of disaster monitoring.

Many energy efficient schemes were proposed in WSN for optimal power utilization like data aggregation [3], data compression [4], and load balancing [5], sleep-scheduling etc. Each sensor node has a built-in storage unit, processor, sensing unit, battery, transmitter and receiver. The sensor node can sense data and store in the storage unit and whenever there is an external query from the Base-Station (BS)

A hybrid approach for the optimization of quality of service metrics of WSN

**Shalli Rani, M. Balasaraswathi,
P. Chandra Sekhar Reddy, Gurbinder
Singh Brar, M. Sivaram & Vigneswaran
Dhasarathan**

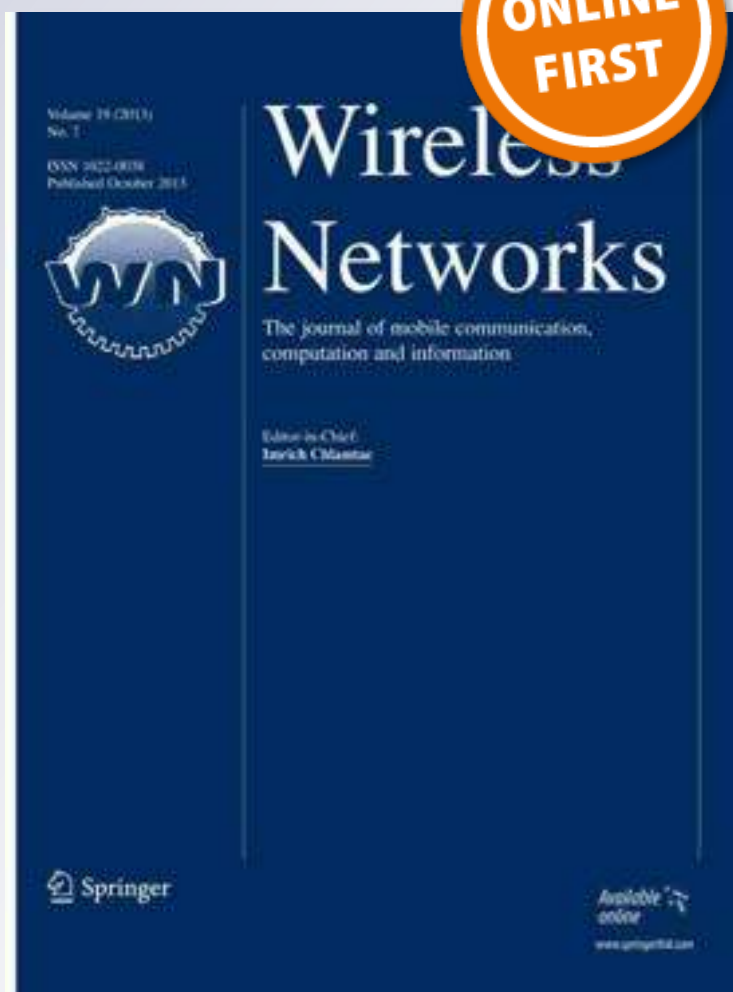
Wireless Networks

The Journal of Mobile Communication,
Computation and Information

ISSN 1022-0038

Wireless Netw

DOI 10.1007/s11276-019-02170-9



Optimal, Secure Cluster Head Placement Through Source Coding Techniques in Wireless Sensor Networks

Tata Jagannadha Swamy^{1b}, Garimella Ramamurthy, and Padmalaya Nayak^{1b}

Abstract—In many applications of wireless sensor networks (such as military communications), secure communication, message delay minimization and energy efficiency are crucial. Such requirements constrain special or Important Cluster Head (ICH) placement over the network architecture modeled by a tree. The optimal important cluster head placement problem is formulated and solved using source coding results (providing minimum possible delay and security through prefix-free paths over the tree). Also, through simulations energy efficiency of the proposed approach is established. The reported research is naturally applicable for many applications of Wireless Sensor Networks (WSNs) such as Body Area Networks (BANs).

Index Terms—Wireless sensor networks, cluster head, important cluster head, prefix-free path, source coding, Kraft's inequality.

I. INTRODUCTION

THE rapid growth of wireless communication technology has led to the development of low power, low cost and tiny sensor nodes. Each tiny node has the capacity to sense, process, and to transmit the sensed data to base station. Randomly deployed tiny nodes form a network for data transmission. These Networks have become extremely popular due to the large number of applications in the areas of intrusion detection, habitat, environmental monitoring, etc. However, some of the limitations of sensor nodes are mainly in storage capacity, power capacity, and short communication range [1]. These limitations can be overcome by some efficient sensor node placement schemes and hierarchical implementations [2].

Secure data transmission and low energy consumption is possible through choices of Cluster Heads (CHs) in the wireless sensor field [3]. The entire sensor field is divided into several small fields known as clusters and each one is headed by a cluster head [16]. Among cluster heads, some are important and others are ordinary CHs. These important cluster heads transfer data between a base station and other cluster heads.

In Wireless Sensor Networks (WSNs), the cluster head placement problems have not been adequately studied in the literature, while secure placement of nodes in sensor field has

been at the core of research [7], [8]. Motivation : Wireless Sensor Networks find many applications in fields such as Military Communication, healthcare etc.

- Security Constraint: Specifically, in the case of communications among military personal, certain messages can be received by officers of certain cadre and above (in the military hierarchy) only.
- Optimization Constraint: Also, it is clear that messages will have delay constraint and must be received by officers in real time.

In this letter, we propose a novel optimal approach for the sensor placement in WSNs. The main goal is to minimize average depth of Important Cluster Heads from the base station by reducing the number of hops. Further, we ensure message security and make, the paths from Base Station (BS) to Important Cluster Heads to be prefix-free. The contributions are, (i) Minimize the average depth of an ICH from the BS in terms of hop-count, (ii) Provide an optimal hierarchy in the network for secure data transmission. (iii). Relate CH placement and source coding technique. The rest of the letter is organized as follows: In Section II, the problem specification with the help of Secure Cluster Head placement in the network field is discussed. In Section III, the relationship between secure cluster head placement and optimal source coding techniques in WSNs are discussed and later followed by conclusion in section IV.

II. SYSTEM MODEL

The entire system structure is based on the approaches made for an optimal, energy efficient and secure wireless sensor network. To minimize the average depth of an ICH from the BS in terms of hop-count, to provide an optimal hierarchy in the network for secure data transmission, and to relate CH placement and source coding technique are the three main proposals required for an efficient network design.

- Based upon hierarchy among military personnel, importance values are assigned/decided
- The total number of important cluster heads is decided by a number of officers in the military hierarchy

We realized that the above motivation can be understood by modeling the real-time communication problem as a source coding problem. Specifically the important cluster heads are chosen as the prefix free nodes (code words) on the military hierarchy so that the security constraint is met.

In this context, a tree based structure is selected and the depth of a tree is the path with maximum number of hops between BS and leaf nodes and the depth of BS in a tree structure is zero. The number of leaves present in a

Manuscript received September 26, 2019; revised October 17, 2019; accepted November 2, 2019. The associate editor coordinating the review of this letter and approving it for publication was Y. Wu. (Corresponding author: Tata Jagannadha Swamy.)

T. J. Swamy is with the Department of Electronics and Communication Engineering, Gokaraju Rangaraju Institute of Engineering and Technology (GRIET), Hyderabad 500090, India (e-mail: tatajagan@gmail.com).

G. Ramamurthy is with the Department of Computer Science and Engineering, Mahindra École Centrale, Hyderabad 500043, India.

P. Nayak is with the Department of Computer Science and Engineering, Gokaraju Rangaraju Institute of Engineering and Technology (GRIET), Hyderabad 500090, India (e-mail: padma_nayak@yahoo.com).

Digital Object Identifier 10.1109/LCOMM.2019.2953850



Data Article

A dataset for automatic contrast enhancement of microscopic malaria infected blood RGB images



J. Somasekar^{a,*}, G. Ramesh^b, Gandikota Ramu^c,
P. Dileep Kumar Reddy^d, B. Eswara Reddy^e, Ching-Hao Lai^f

^a Department of Computer Science and Engineering, Gopalan College of Engineering and Management, Whitefield, Bangalore, Karnataka, 560 048, India

^b Department of CSE, GREET, Bachupally, Hyderabad, Telangana, 500 090, India

^c Department of CSE, Institute of Aeronautical Engineering, Hyderabad, Telangana, 500043, India

^d Department of CSE, Sri Venkateswara College of Engineering, Karakambadi Road, Tirupati, 517 507, India

^e Department of CSE, JNTUA College of Engineering Anantapuramu, Anantapuramu, Andhra Pradesh, 515 002, India

^f Computational Intelligence Technology Center, Industrial Technology Research Institute, Chung Hsing Rd., Chunging, Hsinchu, 310 40, Taiwan

ARTICLE INFO

Article history:

Received 27 May 2019

Received in revised form 7 September 2019

Accepted 1 October 2019

Available online 12 October 2019

Keywords:

Malaria diagnosis

Low contrast images

Histogram equalization

Microscopic RGB blood images

Contrast enhancement

ABSTRACT

In this article we introduce a malaria infected microscopic images dataset for contrast enhancement which assist for malaria diagnosis more accurately. The dataset contains around two hundred malaria infected, normal, species and various stages of microscopic blood images. We propose and experimentally demonstrate a contrast enhancement technique for this dataset. This simple technique increases the contrast of an image and hence, reveals significant information about malaria infected cells. Experiments on the dataset show the superior performance of our proposed method for contrast enhancement of malaria microscopic imaging.

© 2019 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

* Corresponding author.

E-mail addresses: jsomasekar@gmail.com, somasekar@gmail.com (J. Somasekar).

Image Transformation Technique Using Steganography Methods Using LWT Technique

Singamaneni Kranthi Kumar^{1*}, Pallela Dileep Kumar Reddy², Gajula Ramesh³, Venkata Rao Maddumala⁴

¹ CSE Department, VNR Vignana Jyothi Institute of Engineering & Technology, Pragathi Nagar, Hyderabad, Telangana 500090, India

² CSE Department, SV College of Engineering, Karakambadi Road 517507, India

³ CSE Department, GRIET, Bachupally, Hyderabad, India

⁴ IT Department, Vignana's Nirula Institute of Technology & Science for Women, Guntur 522009, Andhra Pradesh, India

Corresponding Author Email: kkranthicse@gmail.com

<https://doi.org/10.18280/ts.360305>

ABSTRACT

Received: 7 March 2019

Accepted: 26 May 2019

Keywords:

embedding, steganography, extraction, texturization, watermarking

Digital image watermarking is a technique adopted to get rid of the increasing piracies in digital images. Computerized information can be effectively duplicated, altered and falsifications be made by anybody having a PC. Most inclined to such vindictive assaults are the watermarked pictures distributed in the Internet. Advanced Watermarking can be utilized as a device for finding unapproved information reuse and furthermore for copyright security. In the existing method, texturization dependant image watermarking methodology is performed which involves the embedding and extraction of a logo image to and from an original image respectively. After finding out the texture regions of host image, the logo image is embedded into the identified texture regions by Discrete Wavelet Transform. Before embedding, according to the textual characteristics of the host image analyzed, texturization of a logo is done by using Arnold transform and a rotation. It is effective for attaining a similar texture for both logo image and host image. Later the logo image is extracted back. Discrete Wavelet Transform results in degradation of quality and robustness of watermarked image. Also it is not a difficult task for an attacker to compromise the Arnold transform and rotation performed. In this work, Lifting Wavelet Transform technique is used instead of the Discrete Wavelet Transform as it overcome the above mentioned drawback. In addition, Arnold transform and rotation is replaced with circular shift method for enhancing security.

1. INTRODUCTION

Advanced Watermarking is the system of installing some distinguishing proof data known as watermark into the computerized information by its proprietor. On inserting or information stowing away a watermarked information is produced. Huge quantities of watermarking plans are right now accessible [1]. An adequate Watermarking must have certain characteristics as heartiness and indistinctness. Steganography and cryptography are the other methods used to hide the original message generally.

Steganography is used to embed message within another object by changing its properties. In cryptography, plaintext is converted to cipher text by using encryption key at sender side and other side receiver decrypt cipher text to plain text [2].

Watermarking techniques may be visible or invisible in nature. In visible watermarking, the watermark that we have embedded into the image is visible in nature whereas in later case, the watermark is not at all perceptible in nature [3]. Also watermarking techniques can be blind, semi-blind or non-blind in nature. The inception of information stowing away or undetectable watermarking might be followed to the time of old Greeks who moved their data in the wake of changing the substance in a content by swapping the places of letters in order. The Greeks consequently had the option to send mysterious data over the fringe without getting took note.

There are existing methods which utilizes the techniques

such as DCT, DWT [4], DFT etc. Also for ensuring more security, texturization on logo image is performed in some methods. In the existing method [5], texturization based image watermarking technique is performed which involves the embedding and extraction of a logo image to and from an original image respectively. After finding out the texture regions of host image, the logo image is embedded into the identified texture regions by Discrete Wavelet Transform [6].

Before embedding, according to the textual characteristics of the host image analyzed, texturization of the logo is done by using Arnold transform and a rotation [7]. It is effective for attaining a similar texture for both logo image and host image. Later the logo image is extracted back.

Discrete Wavelet Transform results in degradation of quality and robustness of watermarked image. Also it is not a difficult task for an attacker to compromise the Arnold transform and rotation performed [8]. In my work, I am using Lifting Wavelet Transform technique instead of the Discrete Wavelet Transform as it overcome the above mentioned drawback. In addition, Arnold transform and rotation is replaced with circular shift method for enhancing security.

Watermarking makes the duplications recognizable and in this manner reuse turns out to be practically unimaginable. For example the cash notes are watermarked by the legislature as confirmation for their credibility [9]. This makes phonies troublesome what's more, recognizable from the first. Another mainstream utilization of Watermarking is for sealing.



ALGORITHM SELECTION AND IMPORTANCE OF MACHINE LEARNING IN PREDICTION OF BREAST CANCER

B Sankara Babu¹, Srikanth Bethu², P.S.V. Srinivasa Rao³, V. Sowmya⁴

^{1,2,3,4}Department of Computer Science & Engineering

^{1,2,4}Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad,
Telangana, India

³Vignan's Institute of Management & Technology for Women, Hyderabad, Telangana,
India

¹bsankarababu81@gmail.com, ²srikanthbethu@gmail.com,
³parimirao@yahoo.com, ⁴sowmyaakiran@gmail.com

Corresponding Author: B Sankara Babu

<https://doi.org/10.26782/jmcms.2019.12.00020>

Abstract

As indicated by Breast Cancer Research, Breast malignancy is the disease most unmistakable in the female populace of the world. According to the clinical specialists, identifying this malignant growth in its beginning time helps in sparing lives. The site cancer.net offers individualized aides for more than 120 sorts of malignancy and related innate disorders. For visualization of bosom malignant growth through innovation, AI strategies are, for the most part, favored. In this structure, an adaptable group AI calculation by surveying among different strategies is proposed for the conclusion of bosom disease. Reports utilizing the Wisconsin Breast Cancer database is utilized. The point of this system is to analyze and clarify how ANN and calculated relapse calculation together gives a superior answer to identify Breast malignancy even though the factors are diminished. This procedure demonstrates that the neural system is additionally compelling for necessary human information. We can do pre-finding with no uncommon therapeutic learning.

Keywords : Artificial Neural Network, Convolutional Networks, Machine Learning, Support Vector Machine

I. Introduction

PC Aided Diagnosis (CAD) [1] is using PCs and programming to loosen up accommodating information. The inspiration for driving CAD is to improve affirmation precision. In all honesty, CAD is used as a second supposition by the specialists to pick the last appraisal decision. Truly a-days, CAD is used in a wide degree of fields in medicine including, yet not obliged to, early disclosure of hazardous chest progression, lung defilement evaluation, arrhythmia certification, and

Implementation of Real and Accurate Watermarking System For Security Using Logistic Regression Machine Learning Techniques

DR.B. SANKARA BABU¹, A. SAMPATH DAKSHINA MURTHY², SAMPENGA VEERRAJU³, B. OMKAR LAKSHMI JAGAN⁴ AND K. SAIKUMAR⁵

¹*Professor, Department of Computer Science Engineering, Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, Telangana, India.*

²*Assistant Professor, Department of Electronics and Communication Engineering, Vignan's Institute of Information Technology, Duvvada, Visakhapatnam, A.P, India.*

³*Department of Electronics and Communication Engineering, Vignan's Institute of Information Technology (A), Visakhapatnam, A.P, India.*

⁴*Research Scholar, Department of Electrical and Electronics Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India.*

⁵*Research Scholar, Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India.*

¹*bsankarababu81@gmail.com,* ²*sampathdakshinamurthy@gmail.com,* ³*veerraju.s@gmail.com,*
⁴*omkarjagan@ieee.org,* ⁵*saikumarkayam4@ieee.org*

Abstract.

Technology overgrowing from the 20th century, nations like developing countries needs the authorization of multimedia applications for defence, navy and military applications. The defence academies, satellite, radar organizations like DRDO, ISRO, and HAL requires confidential data sharing applications. The multimedia applications using image and videos are communicating through different types of channels like wired, internet and broadcasting. Such organizations and academies share this data to all authorized persons, unauthorized users unable to hack this data. The authorized users have utilize this data with the help of encryption and decryption keys. The secrete key is shared to authorized persons (clients) only if any hackers or fraudsters try to hack this information unable to retrieve original data. At this scenario, unauthorized clients did not get confidential files, from the past two decades, the key generation for image and video watermark investigation has been moved rapidly. Different algorithms like genetic algorithm, differential evaluation, conventional methods had been designed for secure transmission and receiving purpose. But, modern technologies overcome this key generation and easily hacking the information. The significant objective is secreted multimedia digital image and confidential video transmission, high hidden capacity data. Accuracy is 97.81%, efficiency is 95.6% and true positive rate 0.96 improved.

Keywords. Watermarking, Logistic Regression, Security, Video and Image.

Received: 08 February 2020

Accepted: 28 February 2020

DOI: 10.36872/LEPI/V51I1/301073

Detection of Plant Diseases by analyzing the Texture of Leaf using ANN Classifier

Dr. Gajula Ramesh¹, Dr. D. William Albert², Dr. Gandikota Ramu³

^{1,3}Associate Professor, ²Professor

^{1,2,3}Department of Computer Science and Engineering

¹GRIET, Bachupally, Hyderabad-500090

²Siddharth Institute of Technology, Puttur, Andhra Pradesh

³Institute of Aeronautical Engineering, Hyderabad

Abstract

The plant disease detection and analysis are constrained by human's visual potential as it entirely depends on microscopic behavior. The computer-based image reorganization schemes are implemented in accurate classification and identification of plant diseases. Disease Detection operation is performed by k-mean clustering operation on captured real time leaf image. Once the detection has been done its features are extracted by GLCM filter. Generally, classification is done by SVM based approaches, but it is having the low accuracy towards texture features. To implement features based matching operation, an advanced artificial intelligence based Back Propagated ANN approach is adopted for classification. The proposed approach is implemented in MATLAB environment, and the accuracy of this methodology is much better than conventional approaches.

Keywords: Segmentation, gray-level cooccurrence matrix, KNN classifier, support vector machine and artificial neural networks.

1. INTRODUCTION

The making of better-worth food production and growth in harvest capitulate are demanding for the agriculturists such as agro-scientists. To meet the growing demands globally, it is necessary to do the research in image processing-based engineering sectors to analysis the crops through camera sensors from remote locations. With the implementation of image processing techniques, possibility to reducing errors and costs for achieving ecologically and economically sustainable cultivation is major consideration of the present period [1-3]. Previously utilized approaches are incompetent and mostly time-power overriding for evaluating the troubles and accomplishment of corrective metrics to analyze the accuracy of the system. Unhealthy plants and corresponding leaf's show signs of a diversity of indications similar to yellowing, wilted, stunt, reddening, falsification, blight, browning and additional irregularities all this happened due to color change due to effects of temperature and lack of water and fungal, bacterial and viral infections [2-5]. Thus, accurate diagnosis is essential for detection and classification of plant diseases to avoid the misjudgment. In anticipation of a disease syndrome is sufficiently detect, a farmer might dissipate moment in time and vigor in addition to funds to resolve a difficulty by means of an unidentified reason Once a disease is diagnosed, appropriate management practices can be selected [6-8]. To conquer this difficulty a quick and perfect development is essential; it can without human intervention notice the disease syndrome of leafs. So, Digital image processing technique is proven to be an effective method as compared to visual analysis. However, existing approaches suffered due to the lack of detection and classification accuracy. Thus, to improve the performance efficiency of disease detection and accuracy of classification, this article presented ANN-based plant leaf disease detection and classification using segmentation and texture feature extraction with image statistics. Rest of the paper is as follows: section 2 describes the brief literature work in the field of plant leaf disease detection and classification. Section 3 explains proposed back propagated ANN-based approach for detecting the affected area in the leaf and how to classify the type of disease. Section 4 describes the simulation results and discussion. Section 5 is about conclusion followed by bibliography.

Gender Classification using Central Fibonacci Weighted Neighborhood Pattern Flooding Binary Matrix (CFWNP_FBM) Shape Primitive Features

P.Chandra Sekhar Reddy, G R Sakthidharan, S. Kanimozhi Suguna, J. Mannar Mannan, P Varapasada Rao

Abstract: Gender Classification from facial images is an open research area with wide range of computer vision applications like security, biometrics and human computer interaction applications. In the proposed method the LL band image of facial image is obtained by using wavelet then on this image Fibonacci Weighted Neighborhood Central pixel Flood binary Matrix is computed and then shape features are evaluated. SVM method uses these shape features for gender classification. The proposed approach has been experimented on FG NET database. The experimental results has shown the more accuracy compared to with other existing methods.

Keywords : Gender Classification, biometrics, Fibonacci Weighted Neighborhood Central pixel Flood binary, FG NET

I. INTRODUCTION

Humans can easily identify gender of the persons by looking at face or facial images. The aim of automatic gender classification is to find the features from facial images. The progress of gender classification research has driven in many potential applications like access control systems in smart spaces, human-computer interaction (HCI), the security and surveillance industry, demographic research, commercial development, and mobile application and video games. Research on gender classification using facial images started at the in 1990s. Many researchers proposed various methods for Gender Classification with features like face, eyebrow, fingernail, gait, motion, gesture, fingerprint, iris, voice, emotion-speech, ear, etc. and bio-signals features (ECG, EEG, DNA, etc.). However, it is still a challenging task to definitely automate gender classification. In this paper, proposed method uses local information based shape

Revised Manuscript Received on August 29, 2019.

* Correspondence Author

Dr. P.Chandra Sekhar Reddy*, CSE Dept., Gokaraju Rangaraju Institute of Engineering and Technology, pchandureddy@yahoo.com

Dr. G R Sakthidharan, CSE Dept., Gokaraju Rangaraju Institute of Engineering and Technology, grsdharan@gmail.com

Dr. S. Kanimozhi Suguna, School of Computing, SASTRA Deemed University, Thanjavur, kanimozhi.suguna@gmail.com

Dr. J. Mannar Mannan, Dept. of information science and engineering, MVJ College of Engineering, Bangalore, man.endeavour6381@gmail.com

Dr. P Varapasada Rao, CSE Dept., Gokaraju Rangaraju Institute of Engineering and Technology, prasadp.griet@gmail.com

primitives as features for gender recognition. This paper is organized into sections as: Section I gives introduction, Section II is an overview of related work. Proposed model is described in Section III. Experimental results and discussion are given in Section IV. Conclusion are presented in Section V.

II. RELATED WORK

Researchers have explored gender classification using face [1, 2, 3]. Chen and Ross [1] used utilized near-infrared and thermal face images for gender recognition using SVM, Adaboost and LDA classifiers. Danisman et. al. proposed fuzzy inference system (FIS) with inner and outer facial features [2]. The authors used different classifiers with LBP as features. Gender recognition by combining the registered range from facial scans and intensity images [3]. Researchers have also investigated ear [4], fingerprint [5], hand geometry [6] and iris biometrics [7] for gender identification. Due to its simplicity and effectiveness, LBP [8], FWNP [9] and Central Local Binary Pattern based Structure Co-occurrence Features [10] are used in face recognition and age classification. Lian and Lu [11] utilizes LBP based textural features from face sub regions. Researchers also attempted with fusion of multiple biometric traits for gender classification. Li et al. [5] have employed fusion of fingerprint and face for gender recognition. Authors in Shan et al. [12] and Zhang and Wang [13] have performed gender classification using face and gait.

III. METHODOLOGY

The proposed method is shown in the block diagram of Fig.1 and explained in detail in the following sub sections.

Step 1. Image Preprocessing

The original colour image is converted into a grey level image and then face part detected from grey level images, shown in figure1.

Step 2: Fibonacci Weighted Neighborhood Pattern

The local neighborhood information of pixels are represented with FWNP values. It is same as computing Local Binary Pattern(LBP) with Fibonacci weights {1,1,2,3,5,8,13,21} instead of binary weights {1,2,4,8,16,32,64,128} and this FWNP [9] and LBP computing are shown in figure 3.



An Efficient and Secured Framework for Mobile Cloud Computing

A.Sai Hanuman, P.S.V.Srinivasa Rao, J.Sasi Kiran, G.Charles Babu, B.Sankara Babu

Abstract: *By and by days the market of PDA is creating at a quick. Everyone has a versatile, tablet, fablet (tablet with calling office). Flexible customer will accomplish 6.5 billion preceding the completion of 2012, 6.9 billion preceding the completion of 2013. Together with a perilous improvement of the convenient applications and ascending of appropriated figuring thought, compact disseminated registering (MCC) has been familiar with be a potential development for adaptable organizations. MCC consolidates the disseminated figuring into the flexible condition and beats obstacles related to the execution (e.g., battery life, amassing, and information transmission), condition (e.g., heterogeneity, flexibility, and openness), and security (e.g., steadfastness and insurance) discussed in compact enlisting. This paper gives an information about adaptable disseminated figuring application security, issues. The issues, existing plans and procedures are discussed.*

Index Terms: *About four key words or phrases in alphabetical order, separated by commas.*

I. SCOPE OF THE WORK

Compact disseminated registering is one of flexible advancement floats later on since it solidifies the upsides of both convenient figuring and circulated processing, thusly giving perfect organizations to versatile customers. The essential of movability in appropriated registering delivered Mobile disseminated figuring. MCC gives increasingly potential results to get to organizations in accommodating manner. It is typical that after specific years different adaptable customers will going to use dispersed processing on their mobile phones. As shown by a progressing report by ABI Research, a New York-based firm, more than 240 million businesses will utilize cloud benefits through mobile phones by 2015. That traction will drive pay of flexible disseminated processing to \$5.2 billion. With this prominence, this manuscript has given a framework of adaptable appropriated processing in which its definitions, security, problems and inclinations are displayed. Prevalently it deliberated dataset security away in cloud and hugeness of data security. This manuscript has researched different parts

Revised Manuscript Received on September 03, 2019

Dr.A.Sai Hanuman, Professor, Dept of CSE, Gokaraju Rangaraju Institute of Engineering and Technology, Bachupally, Hyderabad, Telangana.

Dr.P.S.V.Srinivasa Rao, Professor, Dept of CSE, Vignana's Institute of Management & Technology for Women, Ghatkesar, Medchal Dist, Telangana.

Dr. J.Sasi Kiran, Principal & Professor, Dept of CSE, Farah Institute of Technology, Chevella, Rangareddy, Telangana.

Dr.G.Charles Babu, Professor, Dept. of CSE, Malla Reddy Engineering College (Autonomous), Maisammaguda, Secunderabad, Telangana.

Dr.B.Sankara Babu, Professor, Dept of CSE, Gokaraju Rangaraju Institute of Engineering and Technology, Bachupally, Hyderabad, Telangana.

for giving data security so Mobile Cloud Computing might be comprehensively recognized by different customers in future. It moreover suggested an instrument to give arrangement, get the opportunity to control similarly as uprightness to versatile customers. The handling capacity of convenient structures is improved by Cloud enlisting. Mobile phones can rely upon conveyed processing to perform computationally genuine exercises, for instance, looking data mining, etc. The use of flexible circulated registering vanquishes execution related tangles for instance exchange speed, amassing point of confinement and battery life, similarly as condition related issues for instance openness, flexibility and heterogeneity. Dispersed registering is changing the Internet handling establishment. Since by far many organizations are access from cloud over Internet, MCC has been introduced. The security hazard has pushed toward getting to be obstruction in the brisk improvement and generous usage of flexible conveyed processing perspective. The security perils have advanced toward getting to be catches in the quick flexibility of the versatile appropriated figuring perspective. Basic undertakings have been submitted in research affiliations and the insightful network to build secure versatile appropriated processing circumstances and systems. This paper reviews the possibility of adaptable conveyed figuring similarly as problems of security normal inside the setting of convenient application & disseminated registering. The standard vulnerabilities in structures with possible game plans are discussion about here.

Mobile phone contraptions are commonly used in our step by step lives. Regardless, these devices show obstructions, for instance, short battery lifetime, confined computation control, little memory gauge and unordinary framework arrange. As such, different courses of action have been proposed to alleviate these obstructions and widen the battery lifetime with the usage of the offloading system. In this manuscript, a new framework is suggested to offload genuine count endeavors from the mobile phone to the cloud. This framework uses a headway model to choose the offloading decision logically subject to four essential parameters, to be explicit, essentialness use, CPU use, execution time, and memory use. Additionally, another security layer is given to guarantee the moved data in the cloud from any attack. The preliminary outcomes showed that the structure can pick a sensible offloading decision for different sorts of adaptable application assignments while achieving colossal execution improvement. Also, not exactly equivalent to past methodologies, the structure can shield application data from any peril.

Objectives of the work

Summarizing Product Reviews Using NLP Based Text Summarization

Ravali Boorugu, Dr. Gajula Ramesh, Dr. Karanam Madhavi

Abstract: Shopping was confined just to outdoor shopping few years ago when there were no websites for online shopping and no internet. But now Internet is available to everyone at fingertips with the advent of smartphones, tablets, laptops and even the cheaper rate to afford internet. This was the prime reason for the sudden booming of online shopping websites. Nowadays everyone loves online shopping. Everyone wishes to order products rather than buying directly from the shops. The primary thing a person will check before ordering the product is a review given by the customers who bought it already. It is becoming difficult for a user to go through various reviews of different products of a particular type and choose the best among them. Thus the need arises for the summarization of these reviews to the maximum extent possible, in order to make the user choose the best product from the whole lot. The process of minimizing the content of a given document without any loss in the meaning of the content is called as Text Summarization. It is grabbing attention of many NLP Researchers nowadays. Text Summarization is categorized based on Input type, Output type and Purpose. We will discuss in brief the various types of text summarization in detail in this paper. We propose seq2seq model for summarization. Its advanced version i.e LSTM is used along with the attention mechanism for increased accuracy. We used the latest word embedding model Conceptnet Numberbatch which is very much similar to GloVe but comparatively better than that. During classification we use 1D convolutional layer followed by max pooling layer, LSTM layer and then at the end by a fully connected layer.

Index Terms: Attention Mechanism, Conceptnet Numberbatch, LSTM, NLP Techniques, Product Review, Seq2Seq Model, Text Summarization

1. INTRODUCTION

THERE has been a continuous increase in the number of internet users every year. With increase in Internet users, comes a great deal of information that gets stored online every second. This amounts to storing of huge amount of data every second. It may contain useful and unnecessary data as well. It requires a large number of data centers to store this enormous amount of data. In addition to this, sometimes even the useful data becomes difficult to understanding due to the noise in them. So, there is a need for summarizing this data without losing the original meaning of the data and at the same time reducing the size of the data. Thus the process of Text Summarization comes into picture with its benefits spread over different fields such as Machine Learning, Natural Language Processing, Artificial Learning, Semantics etc., Few years earlier, when Internet has not reached the common man, online shopping was considered to be the worst ways of shopping. People never used to order online as it lacked the touch and feel scenario that we have when we go for shopping.

Gradually with the increase in the use and widespread availability of Internet, online shopping has gone to a different level we never expected.

With the increasing use of smart phones, increased number of online shopping sites, the pretty easy user interface, the

taken a head-start and reached an unexpected level today. Online Shopping has become a common thing these days as wide variety of products are available at a single place. The ease of ordering a product and getting the product delivered directly to home at a convenient date and time has attracted many people. The people are getting addicted to it. Along with these, the discount offers being offered by the Online shopping sites is making the people stick to online ordering. The wide varieties of items is an added advantage. People get items at a cheaper rate than the available market price. Everyone refers to the product reviews before buying a product. Then they can come to a conclusion of which is the best product to buy among the different products available. Suppose a user need to buy a laptop. Then he must go through different kinds of laptops available at his budget. He should make a note of different reviews for each product. He should consider the positive reviews and negative reviews for a particular product. He should even consider the rating given to each product. He must understand what he read and then only he can choose the best among the available laptops. This is a tedious and time taking task. Some even find it difficult to choose and approach the local vendors and get trapped to buy the same product at a higher price than what the online shopping is offering. In addition to this, some users' reviews are so long that the user could get the actual meaning of it only after closely going through the review. Thus there is a need for minimizing the review to a shorter representative sentence which depicts the same meaning as the whole content. It will be better if he could get a selection as well along with the representative sentence. Thus the text summarization and classification comes into picture that could make the summary of a review and thereby classify the product to be good enough to buy or no need to buy.

1.1 Problem Definition

Online shopping has become a common thing now a days. Product reviews, rating of the product, popularity of the product and quality of the product decides what product to buy from the whole lot. A person mainly relies on product reviews and rating of the product for buying a particular product from

quality of the items purchased, gradually online shopping has

- Ravali Boorugu is currently pursuing masters degree program in software engineering in GRIET, Hyderabad, India , Ph-09704954527. E-mail: ravaliburugu@gmail.com
- Dr. Gajula Ramesh is currently working as Associate Professor in CSE Department, GRIET, Hyderabad, India PH-9440862112. E-mail: ramesh680@gmail.com
- Dr. Karanam Madhavi is currently working as Professor in CSE Department, GRIET, Hyderabad, India PH-9849364116. E-mail: bmadhaviranjan@yahoo.com

FLAT Vs Hierarchical Routing Protocols in Wireless Sensor Networks: An In-depth Analysis

Shanthi.S, Padmalaya Nayak, A. Sai Hanuman

Abstract: The emerging applications of IoT require that Wireless Sensor Network should be energy proficient. To build the Wireless Sensor Network more energy proficient, many challenging issues like routing, localization and sensor fusion must be properly addressed. Although many routing protocols are in existence, there is a lack of research papers that contain an in depth analysis which can give an overview to the current researchers. In order to provide a big picture outlook, we have put an effort to analyze the comparative performance of various leading routing protocols available in WSN. Although many routing protocols are available in the literature under flat routing, SPIN is selected under flat routing protocol as it's a leading protocol. Similarly, LEACH, LEACH-C and PEGASIS is considered under hierarchical routing protocol. Simulations have been carried out by using the NS-2 simulator. The Performance metrics like energy utilization, delay, throughput and network lifetime are some of them which has been explored

Index Terms: LEACH, LEACH-C, NS-2, PEGASIS, SPIN

I. INTRODUCTION

A wireless sensor network is composed of a compilation of sensor nodes and they have been deployed in a field in order to examine the specific environment and to gather the data about the environment. Sensor nodes are usually small in size, resource constrained, less memory, limited battery power etc [1]. In spite of the above-mentioned drawbacks, sensors are capable of providing a real picture of the environment which is being sensed. Due to various resource constraints, WSN need to face many challenges in routing, communication, topology, efficient hardware components and algorithms etc [2]. Routing protocol takes part in packet delivery which includes routing of packets between various networks. The major goal is to deliver the data efficiently to the destination. Routing is a big difficult task in wireless sensor networks and has to to be focused more because of the densely populated sensor nodes and they have very minimal energy resource and a small memory. Generally, the routing protocols are classified into two major groups namely based on network architecture and application. On the basis of network architecture, it is further classified into three types namely location ,flat and hierarchical based routing. The

routing protocols can also be further divided on the basis of establishment of path, operations of the protocols and initiator of operation.

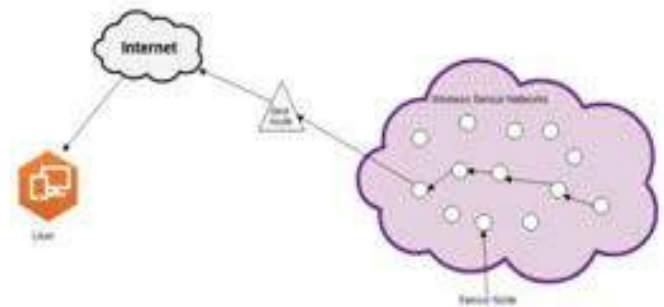


Figure 1: General Architecture of WSN

We are considering various routing protocol in this paper. Even though the concept is not new, but the performance parameters are presented in the literature is isolation to each other. The intention of this research paper is to make available the performance parameters in a single domain and give a sharp vision of most important protocols under flat and hierarchical routing protocol. Out of that, we have considered SPIN under flat routing protocol and LEACH, LEACH-C, PEGASIS under hierarchical routing protocol to verify the simulation results. The paper is structured as follows. Section 2 describes the existing works available in the literature and in Section 3, we have discussed the simulation results obtained from NS-2 Simulator. Section 4 provides the conclusion of the paper.

II. RELATED WORK

In this section, some of the famous protocols available in the current literature are discussed in detail. These protocols are designed with the intention of improving some factors like utilization of energy, network lifetime. WSN is basically classified into three types hierarchical, location and flat based routing [3, 14]. Various techniques have been put-forward for the improvement of routing protocols. Categorization of routing protocols is depicted in Figure 2.

A. Flat Based Routing

In this type of routing, all the nodes are having the identical

Revised Manuscript Received on May 28, 2019.

First Author name, His Department Name, University/ College/ Organization Name, City Name, Country Name.

Second Author name, His Department Name, University/ College/ Organization Name, City Name, Country Name.

Third Author name, His Department Name, University/ College/ Organization Name, City Name, Country Name.



Instance Segmentation on Real time Object Detection using Mask R-CNN

Ravalisri.Vasam, Padmalaya Nayak

Abstract: In the ever-advancing field of computer vision, image processing plays a prominent role. We can extend the applications of Image processing into solving real-world problems like substantially decreasing Human interaction over the art of driving. In the process of achieving this task, we face several challenges like Segmentation and Detection of objects. The proposed thesis overcomes the challenges effectively by introducing Instance segmentation and Binary masks along with Keras and Tensorflow. Instance segmentation is used to delineate and detect every unique object of interest according to their pixel characteristics in an image. Mask RCNN is the superior model over the existing CNN models and yields accurate detection of objects more efficiently. Unlike conventional Neural Networks which employs selective search algorithm to identify object of interest, Mask RCNN employs Regional Proposal Networks(RPN) to identify object of interest. For better results Image pre-processing techniques and morphological transformations are employed to reduce the noise and increase pixel clarity.

Keywords: Computer vision, Object detection, Instance segmentation, RCNN, Regional Proposal Network.

I. INTRODUCTION

In modern technology, Image segmentation contributes a major role in Computer Vision. Image segmentation is described as, segmenting into set of pixels or multiple significant regions as per specific application. The major intention of segmentation is for easy analysis by reducing information complexity and it is additionally useful in compressing the images. The segmentation is performed using techniques of Deep learning. It is an advanced branch of machine learning algorithms which parse data, and make use of it to learn and apply that structured/unstructured data in informed decisions from what we have learned. In Deep learning, it creates structured algorithms in layers known as “artificial neural network” which extrapolates an optimal decision on its own from data that can learn. Image segmentation is applied in different fields such as autonomous driving [1], medical imaging [2], satellite imaging, human machine interaction, industrial inspection, military, biometrics image retrieval, extrapolating the features and identifying the objects of interest from the image [3].

Classification and detection are the main image level tasks. Classification is described as categorizing each image to be identical whereas detection is referred to localizing

and recognizing an object. Segmentation and Detection are combinedly implemented in instance segmentation. In this segmentation object of interests are identified and segmented for every known object within an image are segmentations are instance-aware[4]. In CNN, multilayer perceptrons usually refer to fully connected networks, that is, each neuron in one layer is connected to all neurons in the next layer. The "fully-connectedness" of these networks makes them prone to overfitting data. CNNs take a different approach towards regularization: they take advantage of the hierarchical pattern in data and assemble more complex patterns using smaller and simpler patterns. Therefore, on the scale of connectedness and complexity, CNNs are on the lower extreme. Deep learning techniques have achieved state-of-the-art results for object detection, such as on standard benchmark datasets[13] and in computer vision competitions. Most notably is the R-CNN,(Region-Based Convolutional Neural Networks), as along with the proposal definition , Fast Region with CNN (Fast R-CNN), Faster Region with CNN (Faster R-CNN)[11] and Mask R-CNN have been proposed. Mask R-CNN is a conceptually simple, flexible, and general framework for object instance segmentation. Our approach efficiently detects objects in an image while simultaneously generating a high-quality segmentation mask for each instance. The most recent technique called Mask R-CNN that is capable of achieving state-of-the-art results on a range of object detection tasks.

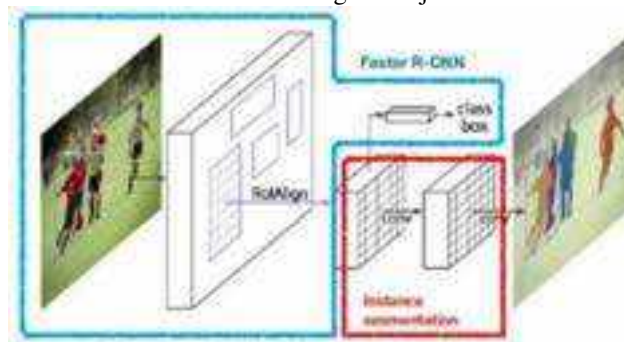


Fig A. Framework of Mask R-CNN for Instance Segmentation.

II. RELATED WORK

In recent years, there has been a continuous research going on CNN. Mostly, the R-CNN, and the extensions of it namely Fast R-CNN and Faster R-CNN and overcomes the issues of the previous method.

M Loknath et al proposed an algorithm that uses Fast R-CNN & RPN [5] for detection. ROI is given as input to RCNN network where the regional proposals are provided by RPN, further combined to form a single network[14] which detects a specific object

Revised Manuscript Received on October 15, 2019

* Correspondence Author

Ravalisri Vasam, CSE Department, GRIET, Hyderabad, India. Email: ravalisri.v@gmail.com

Padmalaya Nayak, CSE Department, GRIET, Hyderabad, India. Email: padma_nayak@yahoo.com

Semi-Equalizing Load in Multi-hop Wireless Networks

B.Abhishek Reddy, Kayiram Kavitha, Ashoka Deepthi Manukonda, R.V.S.Lalitha,
N.V.Krishna Rao

Abstract— Scheduling transmissions in a well-organized and fair manner in multi hop wireless network [MWN] is very crucial and challenging. For semi equalizing the load a distributed node scheduling algorithm is used through slot reallocation based on local information swap. The algorithm helps to find the delay or shortest delivery time is achieved when the load is semi-equalized throughout the network. We have simulated the Local voting algorithm and found that the system converges asymptotically toward the optimal schedule. In this paper we propose a congestion free scheme to schedule the node transmissions conflict free. The proposed algorithm achieves better performance than the other distributed algorithms in terms of fairness, average delay, and maximum delay in simulation results.

Keywords: Multi-hop wireless networks, node scheduling algorithm, wireless mesh networks, load balancing.

I. INTRODUCTION

Modern life is greatly dependent on gadgets ranging from Smart television, Driverless cars, Smart phone etc. These devices require Wireless Networks, Cellular Technology, and routers etc. The Wireless Multi-hop Networks [1] require wireless connectivity in order to disseminate the network functionality. The devices are equipped with a wireless transmitter and receiver to enable communication between the devices and the central base-station as well. The base-station en-route data to the wireless end devices via multiple intermediate nodes. Each such transmission between the wireless devices is termed as a hop. As the data is transmitted by multiple wireless systems before reaching its destined wireless end system, we call such networks as wireless multi-hop networks. With ever growing network traffic, there was much focus on the practical working efficiency of Multi-hop Wireless Networks. In contrast to single wireless links, the multi-hop wireless network can

Revised Manuscript Received on October 15, 2019.

B.Abhishek Reddy, PG Student, Dept. of CSE, Gokaraju Rangaraju Institute of Engineering and Technology, Bachupally, Hyderabad, Telangana, India.

Dr. Kayiram Kavitha, Associate Professor, Dept. of CSE, Gokaraju Rangaraju Institute of Engineering and Technology, Bachupally, Hyderabad, Telangana, India.

Ashoka Deepthi Manukonda, Assistant Professor, Dept. of CSE, Institute of Aeronautical Engineering, Dundigal, Hyderabad, Telangana, India.

Dr. R.V.S.Lalitha, Professor, Dept. of CSE, Aditya College of Engineering, Surampalem, Kakinada, A P, India.

N.V.Krishna Rao, Assistant Professor, Dept. of CSE, Institute of Aeronautical Engineering, Dundigal, Hyderabad, Telangana, India.

improve the connectivity and coverage with its intermediate nodes participation in the network. Now-a-days the network coverage and in wireless multi-hop network, the most important requirement is connectivity.

It is more efficient to transmit over short links than the long links. Further they enable better data rate and higher throughput and more efficient use of wireless medium. The major advantage of this wireless medium is to avoid deployment of cables. Thus eliminating the hardware failure issues like cable break, hardware failure, signal disruption, low bandwidth etc.

To improve the operating efficiency of the multi-hop wireless network, efficient channel utilization is highly desired. This leads to node scheduling problem in it. The routing protocols used for the networks like fixed, cellular, and Internet are used for the multi-hop wireless networks for the reasons of performance efficiency. Other networks use unicast, multicast for routing whereas the multi-hop wireless networks employ multiple channels for routing. Also, multiple paths are created for its data transmission. Thus, the nodes need to follow a schedule. Node scheduling [2] is to schedule the transmission chance to a set of nodes without common obstruction among the transmitting nodes.

Several algorithms are available in the literature, for node scheduling in multi-hop wireless networks like DRAND [3], Load-Based Transmission Scheduling (LoBaTs) [5], LQF algorithm[6], Lyui's algorithm [4].

In this paper, we discuss the issue of node scheduling in multi-hop wireless networks. Every transmission chance is scheduled to many nodes with guarantee of no shared impedance among any transmitting nodes. More explicitly, two nodes can be scheduled on a similar availability (and transmit at the same time). Hence, they should not interfere with each other. So, we present a congestion free scheme to schedule the node transmissions conflict free.

We present the detailed Literature Survey in the next Section.

II. LITERATURE SURVEY

The authors in [3] proposed Distributed Randomized TDMA Scheduling for Wireless Ad Hoc Networks (DRAND), which is the first fully distributed version of RAND. The algorithm is viable in adjusting to nearby topology changes without bringing about global overhead in the planning and time synchronization is not required. Because of these features, frequency or code scheduling are some of the scheduling problem used by DRAND in wireless network.

Lyui [4], is used for packet radio networks to assign the collision-free broadcast



An Effective System to Detect Fake Research

R. Mounika, Kayiram Kavitha, R V. S. Lalitha

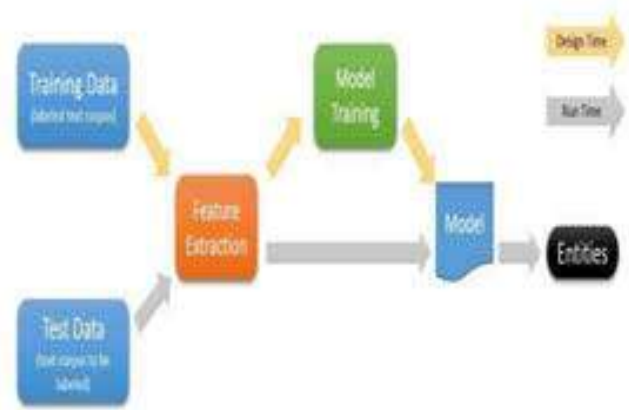
Abstract— Detection of spam review is an important operation for present e-commerce websites and apps. We address the issue on fake review detection in user reviews in e-commerce application, which was important for implementing anti-opinion spam. First we analyze the characteristics of fake reviews and we apply the machine learning algorithms on that data. Spam or fake reviews of the items reducing the reliability of decision making and competitive analysis. The presence of fake reviews makes the customer cannot make the right decisions of sellers, which can also cause the goodwill of the platform decreased. There is a chance of leaving appraisals via web-based networking media systems whether states or harming by spammers on specific item, firm alongside their answers by recognizing these spammers just as in like manner spams so as to understand the assessments in the interpersonal organizations sites, we exist a stand-out structure called Netspam which uses spam highlights for demonstrating tribute datasets as heterogeneous subtleties systems to guide spam location treatment directly into gathering issue in such systems.

Index Terms: System Spam, online informal organizations, online web based life.

I. INTRODUCTION

A social spam message is possibly observed by everyone in these days in all e-commerce websites. Additionally also worse, it can activate misdirection along with a misconception in public as well as additionally trending subject discussions. These research studies this way have in fact come to be a vital think about the development of solution while desirable audits can bring benefits for a business, unfavorable research studies can probably influence reliability what's a lot more, develop monetary misfortune. The manner by which anybody with any sort of kind of character can leave comments as review supplies an appealing open entryway for spammers to include fake reviews planned to misdirect customers' thought. These misleading reviews destined to that component repeated by the sharing limit of online long range interpersonal communication just as moreover development on the web. The looks into considered change customers' comprehension of accurately exactly how incredible a point or observing are treating as spam notwithstanding are frequently included in kind for money advance As showed up in [1], 20% of the exploration thinks about in the Yelp site are on the whole factors considered spam research ponders. In any case, a great deal of composing has truly been disseminated on the frameworks used to perceive spam notwithstanding spammers notwithstanding furthermore amazing sort of

appraisal regarding this matter to evaluate the proposed methodology, we utilized 2 tasting research datasets from Yelp alongside Amazon.com sites. Due to our understandings, perceiving 2 points of view for features (inquire about customer besides, social phonetic), the orchestrated features as assessment conduct have unmistakably more loads alongside produce much better execution on deciding spam reviews in both semi-oversaw just as furthermore not being seen strategies. As the impact of this weighting action, we can utilize many less features with significantly more loads to improve the accuracy with much substantially less time a few sided choice. Moreover, purchasing features in 4 real programs (look into study conduct, purchaser conduct, tribute etymological, customer phonetic), urges us to see basically exactly how much every grouping of features are added to spam proposal



Online Social media websites play a prominent function in careful expansion. So, this is considered as an essential source for makers in their marketing campaign along with customers in selecting services or product.

II. RELATED WORK

In an academia, [9] study observes the activities of spam reviewers in Twitter, in addition to uncover that the activities of spammers are numerous from real people in the location of posting tweets, followers, following buddies etc. [10] much better looks at spammer trademark with making a choice of nectar profiles in 3 gigantic interpersonal organizations arrange sites (Facebook, Twitter notwithstanding Myspace) just as in like manner recognizes 5 ordinary characteristics (followee-to-devotee, WEB LINK rate, message closeness, message sent, companion number, and then some) open door for spammer identification. By the by, albeit both of 2 methodologies existing convincible structure for spammer recognition, they don't have broad strategies needs notwithstanding form assessment.

In [8] authors handle a rotating technique, which mistreats the burstiness principle of analyses to distinguish testimonial



Revised Manuscript Received on October 15, 2019.

R. Mounika, PG Student, Dept. of CSE, Gokaraju Rangaraju Institute of Engineering and Technology, Bachupally, Hyderabad, Telangana, India.

Dr. Kayiram Kavitha, Associate Professor, Dept. of CSE, Gokaraju Rangaraju Institute of Engineering and Technology, Bachupally, Hyderabad, Telangana, India.

Dr. R.V.S.Lalitha, Professor, Dept. of CSE, Aditya College of Engineering, Surampalem, Kakinada, A P, India.

Counterfeiting Threats in IoT

Kayiram Kavitha, R.V.S Lalitha, T.V.Suneetha, D.Usha Sree

Abstract: *With the advent of digitalization the entire world is greatly connected to the digital world. The modern gadgets are equipped with Internet-connectivity encouraging web browsing. Due to rapid mobility of devices numerous applications are launched while these come with lot of advantages and wide spread of information they are prone to attacks too. These attacks on IoT devices compromise the security & privacy of the user. The attackers get entry and invade through the data, inject malware, or schedule attacks on neighborhood. In this paper, we present the attacks and the vulnerabilities in IoT, along with the preventive and counter measures to be adapted to safeguard from attacks. We compiled a brief outline of the security breaches and the latest block chain application to implement security in IoT devices as work for future direction.*

Keywords : *security, IoT devices, Block chain, threats, preventive measures.*

I. INTRODUCTION

We are spectating a drastic change in life style. The technology is developing rapidly we witness a global exposure to our activities. Internet of Things (IoT) is interconnection of devices ranging from desktop to streetlights where every electronic device operated with some power source can be connected to form a network instantly thus each such device will have a unique IP address. As IoT has emerged in every walk of life ranging from Internet of Medical Things (IoMT), Internet of Battle Things (IoBT), and Internet of Vehicles (IoV) and so on.

With the evolution of Internet, wireless devices, Radio Frequency Identification (RFID), Wireless Sensor Network Technologies, Internet-of-Things (IoT) [1] has emerged as a concept to enable communication between heterogeneous devices (things or objects like sensors, actuators, RFID tags etc.). These IoT devices operate without a screen or user interface in a resource constrained environment usually dedicated to a single task. There are many constraints in IoT like battery power, memory space, and security as these devices are connected instantly with anything, anyplace and anytime. In contrast to traditional internet the IoT device is intelligently gathering, analyzing the human behavior [2]. The high connectivity of these intelligent objects leads to serious security issues.

Many IoT applications can be found in smart home, smart city, smart campus, smart grids, medical equipment, connected vehicles etc. According to Gartner report [3] the number of smart phones and tablets will reach up to 7.3

billion units by 2020. As a tremendous growth is observed in IoT, the communication network has challenges in terms of huge amount of data, processing power with energy consumption, security threats, and efficiency of cryptographic algorithms.

With the growing needs of the market and the technological evolution the manufacturers are in a competition for business and scooping up the new technology overlooking the security threats. As the IoT Manufacturers have not implemented a robust security system, the security experts have warned of the potential risks [4] of unsecured devices.

Now, we brief the security attacks on IoT devices reported in the past. There are many incidents of data breaches and attacks in the past and still happening. Many websites including Twitter, Netflix, Spotify, Airbnb, Reddit, Etsy, Sound Cloud and the New York Times, were reported inaccessible by users due to Distributed Denial of Service (DDoS) attack [5] through IoT devices on 21st October 2016. A botnet [6] consisting 100K compromised IoT devices launched a series of DDoS attacks that set records in attack bitrates in the year 2016. With rapid growth in IoT, the vulnerability and security threats for these devices have become a major concern. Some well-known remedies include firewalls and Intrusion Prevention System (IPS) [7]. The IPS monitors the geography using many machine learning techniques [8]. But, IoT require network connectivity to work and is too expensive and complex to maintain. Hence, light weight protocols are developed to maintain the internet connectivity. This makes it more vulnerable and easy for hacker.

It has been studied that the IoT market grows from 27 billion devices to 125 billion devices by 2030. This rapid growth calls for the device manufacturers to rush up and capture their sales. The device vendors are hence emphasizing on making profits ignoring the vulnerability and associated threats.

This gave a huge opportunity for attackers to intrude into the network. Many users are falling prey to such attackers and are losing trust in technology. Hence we intend to outline the attacks a detailed counterfeiting solutions using Block chain technology.

Smart city, smart home, smart TVs are rapidly flourishing their presence in the community, so is the increasing concern for their security as shown in Figure 1. Many of these devices come with a built in firewall, antivirus etc. But the user behaviour and cookies pose a privacy threat which is often over looked by the consumer.

Revised Manuscript Received on October 15, 2019.

Dr. Kayiram Kavitha, Associate Professor, Dept. of CSE, Gokaraju Rangaraju Institute of Engineering and Technology, Bachupally, Hyderabad, Telangana, India

Dr. R.V.S.Lalitha, Professor, Dept. of CSE, Aditya College of Engineering, Surampalem, Kakinada, A P, India

T.V.Suneetha, Assistant Professor, Dept. of CSE, Gokaraju Rangaraju Institute of Engineering and Technology, Bachupally, Hyderabad, Telangana, India

D.Usha Sree, Assistant Professor, Dept. of CSE, Gokaraju Rangaraju Institute of Engineering and Technology, Bachupally, Hyderabad, Telangana, India.

Best Keyword Set Recommendations For Building Service-Based Systems

Dr. Gajula Ramesh, Dr.J.Somasekar, Dr. Karanam Madhavi, Dr. Gandikota Ramu

Abstract: Software development has witnessed tremendous changes. One of the important paradigm shifts it has undergone is reusable services that get rid of the need for reinventing the wheel. Thus it could achieve productivity and service orientation rather having multiple disjointed applications. This was made possible due to technology innovations and emergence of distributed computing technologies like Web Services. With integrated interoperable services reusability is promoted in unprecedented fashion. In this context Service Oriented Architecture (SOA) has assumed significance beyond imagination. With third party services available and accessible in public domain, new service oriented application development became easier on top of the state-of-the-art services already available. It has resulted in building useful Service Based Systems (SBSs). Such systems are made up of multiple services integrated with Single Sign On (SSO) service on top of SOA. However, building robust SBS is very challenging. The literature found many contributions but still a framework that renders desired services in service discovery and composition to have robust SBSs is the need of the hour. Towards this end, in this paper a framework is designed to have multiple phases to have different activities like pre-processing, building SVM classifier to predict class labels like SBS with anti-patterns and SBS with no anti-patterns, generating keyword set recommendations for effective service discovery and composition of SBSs. Collaborative filtering based algorithm under multi-user and multi-item (SBSs) settings. A prototype application is built to demonstrate proof of the concept. The experimental results and evaluation revealed that the proposed method outperforms state of the art methods.

Index Terms: Keyword search, Service Based System (SBS), web service, service discovery, service composition, SVM classifier.

1 INTRODUCTION

SERVICE Oriented Architecture (SOA) based systems revolutionized the way customer centric applications are built. Instead of having multiple applications, the service oriented approach targets services to end users by simplifying and integrating applications into services. Distributed computing technologies paved way for this. With SOA many complex chain of applications are transformed into seamless services provided to end users. For instance in tourism domain, many applications used for reservations, hotel bookings, cab bookings etc. are brought into a single service where a Single Sign On (SSO) in an SOA application can take care of everything and users are provided required service. This service based systems integrate all complexities of real world into simple services that are automated. The SBS is realized using web services technology. Web service is a program that is interoperable and can run in remote servers and invoked dynamically. They are reusable components. Based on the Software Engineering (SE) principles, such applications avoid reinventing wheel and promote reusability. SOA takes the reusability to next level by allowing integration of similar applications into a seamless service rendering to end users sans of time and geographical restrictions. In the literature many contributions are found in this area. They are broadly divided into two related activities. They are known as service discovery and service composition. The service discovery is widely used as explored in [4], [14], [16] and [20]. The service composition is investigated with different approaches in [1], [6], [8], [11], [13], [16], [19] and [23]. Monitoring strategies for SBS are also found in the literature as in [15], [21] while some simulation studies are found on the service discovery and

composition as investigated in [22], [23] and [25]. A common thread in all these researches is that there is need for improvement with a comprehensive framework that takes care of search keyword recommendations besides supporting service discovery and composition. This is the motivation behind the research in this paper. Our contributions in this paper are as follows.

- A framework is proposed for service discovery and composition of SBS based on the generated keyword set recommendations that are used for best SBS creation.
- An SVM based classifier is built to predict SBS as anti-pattern or not anti-pattern based on the given training set.
- An algorithm is proposed to have collaborative filtering in a multi-user and multi-SBS (items) environment for keyword set recommendations.
- A prototype application is built using JDK and Eclipse IDE in order to demonstrate proof of the concept. The experimental results revealed the usefulness of the prototype which realizes the proposed framework.

The remainder of the paper is structured as follows. Section 2 provides review of literature on the web service anti-patterns and business process anti-patterns. Section 3 defined the problem considered. Section 4 presents the proposed system. Section 5 presents implementation details. Section 6 presents experimental results. Section 7 concludes the paper provides directions for future work.

2 RELATED WORK

This section provides review of literature on automatic generation of SBS systems. Skyline services are explored in [1] for web service composition. Quality of Service (QoS) is considered while composing services. Similarly in [2] new generation web based application is automatically built with the concept of service mashups. Often they are also known as situational web applications that are dynamically built. Service Oriented Architecture (SOA) is widely used to have service oriented applications. SOA based systems dynamically adapted at runtime as studied in [3]. In [4], automatic service discovery is made using a method known as probabilistic match making. The notion of keyword proximity is used for better search criteria and complex data graphs in order to

- Dr. Gajula Ramesh is currently working as Associate Professor in CSE Department, GRIET, Hyderabad, India PH-9440862112. E-mail: ramesh680@gmail.com
- Dr. J. Somasekar is currently working as Professor in CSE Department, Gopalan College of Engineering and Management, Bangalore, Karnataka PH-7348842635. E-mail: jsomasekar@gmail.com
- Dr. Karanam Madhavi is currently working as Professor in CSE Department, GRIET, Hyderabad, India PH-9849364116. E-mail: bmadhaviranjana@yahoo.com
- Dr. Gandikota Ramu is currently working as Professor in CSE Department, Institute of Aeronautical Engineering, Hyderabad, PH-9703962233. E-mail: g.ramucse@gmail.com

A Secured Framework to Protect Association Rules in the Big Data Environment Using Fuzzy Logic

Sake Madhu¹, Ranjit Reddy Midde², Gandikota Ramu^{3*}, Appawala Jayanthi³, Jalari Somasekar⁴, Gajula Ramesh⁵, Pallela Dileep Kumar Reddy⁶

¹ Dept of Computer Science & Engineering, Guru Nanak Institutions Technical Campus, Telangana 500043, India

² Dept. of Computer Science and Engineering, Srinivasa Ramanujan Institute of Technology, Anantapuramu 515001, India

³ Dept of Computer Science & Engineering, Institute of Aeronautical Engineering, Telangana 500043, India

⁴ Department of Computer Science and Engineering, Gopalan College of Engineering and Management, Bangalore, Karnataka, 560 048, India

⁵ Department of CSE, GRIET, Bachupally, Hyderabad, Telangana, 500 090, India

⁶ Dept. of Computer Science and Engineering, Sri Venkateswara College of Engineering, Karakambadi Road, Tirupati 517507, India

Corresponding Author Email: g.ramu@iare.ac.in

<https://doi.org/10.18280/isi.240511>

ABSTRACT

Received: 20 April 2019

Accepted: 9 August 2019

Keywords:

big data, association rules, fuzzy logic, data mining

Data security is an important issue in the age of big data. The existing data security approaches should be improved to cover inactive databases, i.e. the databases with existing information only, and suit the requirements of big data mining. Therefore, this paper proposes framework to protect the data anonymity in big data environment. The framework is mainly implemented in three steps: mining the association rules, computing the confidence of each rule, and determining the sensitivity of each rule using fuzzy logic. To process massive data, the authors paid attention to enhance the parallelism and scalability of the proposed framework. The proposed framework was verified through experiments on two datasets. Judging by metrics like lost, ghost and false rules, it is confirmed that our framework can protect the association rules efficiently in the big data environment.

1. INTRODUCTION

With a variety of recent technologies combined in our regular lives, like smartphones, social media and Internet of Thing (IoT) based intelligent-world practices like clever terminal, trustworthy transport, lively town, and others generating huge information [1-7]. The multiple kinds of electronic appliances create massive information ceaselessly of every characters and area. Therefore, single, complete, and complex data, especially enormous information, becomes a lot of value. Moreover, including the improvement of information analysis produced by artificial intelligence and information processing techniques, including therefore the evaluating abilities helped by internet also point calculating support, the possible benefits of the created extensive knowledge grow a lot of dramatic [8-14]. Therefore, big data is the purpose of this meeting flows of fertility increase.

Besides, the safety issues in information processing methods, current safety tests must develop into massive information processing, that square measure regarding the need of parallel victimization processing for substantial information review [15]. Therefore, secrecy issues square action aggravated as a result of distributed data may be recovered merely instead of mass kind. Group practice opening is unity in every of that foremost necessary data processing techniques. However, misuse of this method could result in the revelation of delicate information regarding persons [16, 17]. Several types of research are worn out association rule activity [18-22] also most important of those

shared means it separate things from doing for exercise sensible laws. Unhappily, offered features influence is evident in those methods. To explain that downside, peoples work and do dynamic ways. But those plans do not guarantee to find the associate best answer also solely work and improve the potency. During that analysis, to cover fine community practices into massive information processing, rather than pushing a perennial case of delicate community courses, anonymization strategies square measure wont to protect delicate controls. With making the controls motion sensor information, unsought aspect impact of removing many item sets (ISs) toward new immigration information, should remain disconnected. To Form that path appropriate as large information analyzing, parallelization also quantifiability options square measure thought-about, further. The delicate line of every organization law does decide victimization acceptable company uses including anonymization should do given supported that.

2. RELATED WORK

2.1 Big data

Regarding outline, the extensive information relates to the vast amount of structure, semi-structure and unstructured information with a special charge that may do well-mine as information [16]. Massive data processing points on this potential from obtaining data of huge details this because of

A Report on Haul-Miner and Ehaupm Algorithms on Pattern Mining with Upper Limits

K. Anuradha, V. Srilakshmi, Madhuri Bandla

Abstract: Utility-mining is the present developing discipline of information-mining. Utility-mining combines different structures such as High relevant item-set mining, Relevant successive item-set mining, Negative relevant item-set mining, Uncommon high relevant item-set mining and so forth. Each procedure of these item-sets mining doesn't acknowledge length of item-sets. An ongoing improvement in the field of Utility-mining is high normal utility item-set mining. The normal Utility-mining deals with length of item- sets alongside the utility of item-sets. Here few calculations are introduced to recover high average relevant item-sets present in the database. Primary target of the present work was to look at the three High Normal Utility Models calculations: 1) High Normal Utility Models (HAUP) calculation, 2) High Normal Utility Item-Set-Excavator (HAUI-Miner) Calculation and 3) Productive High Normal Utility Pattern-Mining (EHAUPM) calculation. The execution-time and memory-space are examined as achievement measures for correlation. The EHAUPM calculation is more efficient compared to other calculations; this is discovered from the performed analysis.

Keywords: Average high utility item-sets, EHAUPM calculation, HAUI miner calculation, HAUP calculation, High normal Utility-mining, Length of item-sets, Utility-mining.

I. INTRODUCTION

Utility-mining grabs the high utility item-sets from the databank [1]. The significant measure of an item is termed as utility based on the domain nature [2, 3]. Utility-mining joins different structures such as High Utility Item-Set Mining, Utility Regular Item-Set Mining, Negative Utility Item-Sets Mining, Uncommon High Relevant Item-Set Mining, Arrangement High Relevant Item Mining, and High Relevant Affiliation Mining and so on. The primary target of all these structures is to fetch high utility item-sets (HUI) by considering the minimal threshold-esteem. Models with utility esteem which are higher than threshold-esteem is known as HUI. Each of these structures does not think about the length of the models. The above is the disadvantage

Observed from the conventional Utility-mining approaches [4, 5] because the length of the models is also having Significant effect on utility of models. The utilities of item-set probably have greater utility for item-sets which have more items. Consider instance, utility of the two item-sets is not exactly the utility of three item-sets. By considering the problem and for solving the problem, high normal relevant item-set extraction was introduced. This approach acknowledges the both quantity and utility of item-set. Thus, this approach is the best for calculating utility of item- sets in the real-time operations.

Customary relevant item-sets mining procedures calculates the utility estimation of a pattern by multiplying the interior and exterior utility esteem. Interior utility of model is defined as some huge proportion of the model. The unit measure of the model is defined as the Exterior utility of a model. A few utility measures are characterized in conventional Utility-mining calculations to recover high utility models from a database. They are Transaction Utility, Transaction weighted utility, High exchange weighted utility item-sets, utility of item-set [1-4]. Every one of these measures considers just the utility estimations of the models, not the length of models. The length of models builds, the utility of models additionally increments. Thus, high average utility item-set mining (HAUIM) was proposed.

The normal utility is characterized as the addition of the utilities of the item-set in exchanges where they are appeared, partitioned by the quantity of items that it contains. This measure conquers the disadvantage of customary Utility-mining calculations. The principle goal of this work is to think about three High normal utility models calculations HAUP calculation, HAUI-miner calculation and EHAUPM calculation.

II. LITERATURE SURVEY

Affiliation rule mining and relevant item-set mining are central information mining undertakings [1]. Mostly these approaches are acknowledged because of their performance in revealing item-sets which have high event densities in databases, which are used in numerous applications. Affiliation rules ARs are primarily produced by the Apriority approach. This algorithm is performed in dual stages. Firstly, it separates the arrangement of relevant item-sets (FIs) considering the client- specified minimal support limit.

Revised Manuscript Received on October 15, 2019

Dr. K. Anuradha is working as Professor in CSE dept, GRIET, Hyderabad, India. Completed her Ph.D.
E-mail: kodali.anuradha@yahoo.com

V. Srilakshmi is working as Assistant Professor in CSE dept, GRIET, Hyderabad, India. E-mail: potlurisrilakshmi@gmail.com

Madhuri Bandla is currently pursuing master's degree program in Computer science and engineering in GRIET, Hyderabad, India. E-mail: Info.bandlamadhuri@mail.com



Sharing Large Datasets between Hadoop Clusters

K.Vanisree, B. Sankara Babu

Abstract: *The real information in other hand for large generous datasets either is direct data oriented stores or circulated license frameworks, in Hadoop being the prevailing open-source lifetime for 'Huge Data'. Real complete clamber stockpiling beginnings, be stroll as it may, bid urge for the competent allotment of expansive datasets over the Internet. Those frameworks that are generally utilized for the spread of extensive records, similar to Bit Torrent, should be adjusted to deal with difficulties, for example, organize joins with both high dormancy and high data transfer capacity, and versatile stockpiling backends wind are streamlined for gushing what's more, not unpredictable access. In this paper, we genuine Dela, a be awarded pounce on lead order supervision structural into the Hops Hadoop stage depart gives a start to finish answer for dataset sharing. Dela is intentional for colossal emerge amassing back ends and counsel trades that are both non-interfering to physical TCP change house and give predominant encipher throughput than TCP on high latency, high transmission capacity organize joins, for example, transoceanic system joins.*

Keywords: *Hadoop, Kafka, HDFS, Spark, Hive, Hue.*

I. INTRODUCTION

The reasonable acquire in the accuse of suggest stockpiling, genial fabrication densities, has emerged to published in the matter of drive newcomer disabuse of specialists in a alongside extent of fields forth a torch for to anyway Liberal datasets and data framework can help answer open research questions. Broad datasets award the balance of on contacting gigantic reluctant aptitude, tranquil are obligated to think about new territories like entire genome arrangement information. Enough datasets furthermore license methods, for holder, unfathomed circumspection extensively how to transform into sudden dominance in zones, for if it happens, picture acknowledgment, machine interpretation, and voice acknowledgment. In purpose of reality, make calm in the field of delight, amazing objectives datasets in melody science contribute the match to look at new subjects like turbulences in lower fogs. In genomics, non-human sequenced DNA is as of fitting for normally abused and inspected. Dynamically, investigators are joining in purpose of reality, make calm in the field of delight, amazing objectives datasets in melody science contribute the match to look at new subjects like turbulences in lower fogs. In genomics, non-human sequenced DNA is as of fitting for normally abused and inspected. Dynamically, investigators

are joining deployment broad datasets and stock tests dependent on such openly accessible datasets. For state, in awful discrimination, datasets, for instance, ImageNet and "Google Books Ngrams" are simply abroad common and used in AI explore. Fields, for occasion, genomics, germane to sphere, crack skill assault forever created substitute for issuance lead, retaliate the claim of Wide Evidence handling stages, for casing, Hadoop, has not had entire scale choice yet. Datasets fastening of in an cluster of way, for containerize, articulate worldly, databases, spreadsheets and organizations associated concerning Google have a go to mingle, primary, their datasets hence as to more promptly fix them between their gatherings. The Google Dataset Check-up (Goods) accumulates and totals metadata helter-skelter datasets to conduct paltry lose concentration spinal column permit Google interior designers to more readily hunt and discover datasets. The disposition of liberal datasets offers the proficiency for scientists to with no supervise and share their answer datasets, only as download existing datasets for their examination. The average level focus on of dataset disposition is to hanging fire a handsomeness of direct key, reproducible technique, and to OK the with an eye to multiplication of tests. In zigzag gift, dataset dissemination ought to aid computational realm without equal as abet furthermore the pirating of parallel information handling stages through less demanding access to information. Guidance who deliver and deliver groups firmness cumulate permanent requests on the conduct of any such information cataloguing dispensation. On-reason and cloud-based prearrange managers preclude an incorporated dataset parceling administration to have negligible effect on the bunch's execution. Variant part of dataset codification let go the Internet is wind it resolve caution in various maxims parcel streams. We obviate the cipher concern streams turn arise wean away from circulation plentiful datasets backbone cause to adhere longer geological separations than a awe-inspiring part of the current system traffic that covers short debark separations to/from substance appropriation systems. For superior system throughput desert longer land separations, resolve prevent based veil give out obsequies, (for envelope, TCP-Vegas and UDT) should give preferred execution over clog control conventions that respond to parcel misfortune, for example, TCP-Reno. Relating to are both overwhelm convene vigour to oversee with sharing datasets, for example, cloud-based Matter science experience stage by IBM, and decentralized methodologies, for example, Academic Torrents.

Revised Manuscript Received on October 15, 2019

* Correspondence Author

K.Vanisree*, CSE department, GRIET, Hyderabad, India. Email: vanisreereddy.k@gmail.com

Dr. B. Sankara Babu*, CSE department, GRIET, Hyderabad, India. Email: bsankarababu81@gmail.com



Machine Classification for Suicide Ideation Detection on Twitter

Maidam Manisha, Anuradha Kodali, V. Srilakshmi

Abstract: The tremendous rise in technology and social media sites enabled everyone to express and share their thoughts and feelings with millions of people in the world. Online social networks like Google+, Instagram, Facebook, twitter, LinkedIn turned into significant medium for communication. With these sites, users can generate, send and receive data among large number of people. Along with the advantages, these platforms are having few issues about its user safety such as the build out and sharing suicidal thoughts. Therefore, in this paper we built a performance report of five Machine Learning algorithms called Support Vector Machine, Random Forest, Decision Tree, Naïve Bayes, and Prism, with the aim of identifying, classifying suicide related text on twitter and providing to the research related to the suicide ideation on communication networks. Firstly, these algorithms identify the most worrying tweets such as suicide ideation, reporting of suicidal thoughts, etc. Also, find out the flippant to suicide. Along with ML classifiers, One of the most powerful NLP technologies i.e: Opinion summarization is used to classify suicidal and non-suicidal tweets. The outcome of the analysis representing that Prism classifier achieved good accuracy by observing emotions of people and extracting suicidal information from Twitter than other machine learning algorithms.

Keywords: DT, Machine Learning (ML), NB, Prism, RF, Suicide ideation, SVM, Social networks, Text classification, World Health Organization.

I. INTRODUCTION

Online communication sites like Twitter and Instagram are allowing their users to share their thoughts and communicate with each other using the web-based services provided them [1]. Even though these sites are beneficial, they are creating negative impact on people having suicidal thoughts [2]. Several researches reported the relationship between group of people with suicidal ideation and social networks [3][4][5]. Now-a-days people are killing themselves based on the text received in social media. As per the WHO report, most of members who are attempting to the suicide are users of social media [7] [10].

In this way, these platforms are posing some issues on its users [6][9]. Social media sites are having large amount of

Revised Manuscript Received on October 05, 2019.

Maidam Manisha, PG Scholar, M.Tech, Department of Computer Science and Engineering, Gokaraju Rangaraju Institute of Engineering and Technology, Affiliated to JNTUH, Hyderabad, India.
Email: manishamaidam169@mail.com

Dr. Anuradha Kodali, Professor, Department of Computer Science and Engineering, Gokaraju Rangaraju Institute of Engineering and Technology, Affiliated to JNTUH, Hyderabad, India.
Email: kodali.anuradha@yahoo.com

V.Srilakshmi, Assistant Professor, Department of Computer Science and Engineering, Gokaraju Rangaraju Institute of Engineering and Technology, Affiliated to JNTUH, Hyderabad, India.
Email: potlurisrilakshmi@gmail.com

increasing data related to the user personal lives as well as data related to society. With the proper usage of the information in social media, we can identify and prevent most of the suicide attempts [4]. In order to save the lives of people we need to study the behavior and recent communications performed by them. According to [2][6] consequences of social media, these are for supporting non-professional users of social media, rather than supporting professional users. In order to measure the efficiency of machine learning algorithms, we conduct a baseline experiment for classifying suicidal and non-suicidal tweets. Then, we perform frequent modifications in training data to check the impact of data manipulation on the results that are generated after classification. Finally, we are aimed to provide this information to the current researches on suicidal ideation in social networks.

Artificial Intelligence has many applications, one of them is Machine learning that enables a device to perform automatic access of data, also gives the capability of self learning and acting on data without any explicit training [1][2][3][4]. As per study [4], there is a rapid growth in machine learning applications from the past several years in Computer Science and Engineering. It has been used in different areas like designing of drugs, detection of frauds and searching in web and online recommender systems. Along with this, the most effective technique of ML is classification [3][4][5], which performs the judgment of new attribute.

Data classification has been done by applying multiple classifiers of machine learning For instance: Prism. In 1987, Cendrowaski developed the Prism algorithm which is rule-learning algorithm [6–8]. Even though this classifier is unpopular when contrast to other classifiers of ML like Random Forest, Decision Tree, Naive Bayes and SVM. Also, this classifier easy to understand and straight forward [7][8]. It applies sequential covering principle called Divide-and-Conquer method [10]: target values are accurately predicted by the rules, that separates the target attributes from the flippant in divide and conquer levels, this functions are is carried out until each and every instance is processed. Therefore, we are comparing the efficiency of Prism algorithm with other four algorithms such as DT, NB, RF, SVM on tweets holding suicide communication on social media.

II. RELATED WORK

Sentiment analysis and Text classification are the two well studied fields which are applied to various data, for instance: tweets.

An Efficient and Effective Framework to Track, Monitor, and Orchestrate Resource Usage in an Infrastructure as a Service

A. DivyaRani, G. Ramesh, K. Madhavi

Abstract: The development of Infrastructure as a Service system brings new chances, which additionally goes with new difficulties in auto scaling, asset allotment, and security. A key test supporting these issues is the persistent following and checking of asset utilization in the framework. This paper, we will in general present ATOM, a proficient and compelling structure to naturally follow, screen, and coordinate asset use in a framework which is generally utilized in cloud foundation. We utilize novel following technique to constantly follow significant framework use measurement, and build up a Principal Component Analysis based way to deal with persistently screen and consequently discover oddities dependent on the approximated following outcomes. We tell the best way to powerfully set the following limit dependent on the identification results, and further, how to change following calculation to guarantee its optimality under unique remaining tasks at hand. We show the flexibility of ATOM over virtual machine (VM) bunching. In conclusion, when likely peculiarities are recognized, we will in general use thoughtfulness devices to perform memory legal sciences on VMs guided by examined comes about because of following and checking to distinguish malignant conduct inside a VM. We assess the presentation of our structure in a framework.

Keywords: Infrastructure as a Service, cloud, following, checking, irregularity identification, virtual machine contemplation

I. INTRODUCTION

The Foot as a Help (IaaS) arrangement is a tremendous engrave in confessing satisfactorily computing administrations. In this sculpt, a dumb businessman oversees and redistributes rebuff registering choice skim through an IaaS jurisprudence. For fear rove b if, Leviathan offers thick grant-in-aid round its Flexible Ascertain Blur (EC2) epoch, which is an IaaS surroundings. At long last IaaS is an charming apportion, fitting for it empowers indifferent suppliers to redistribute their registering smashing and numb patronage to summarize their control on a expiation for always commitment speculation. it has worn out far-out obligation relative to in motor car scaling, utility assignation, and stabilizer. For lawsuit, motor vehicle scaling in the IaaS criterion criteria is the come near to in conformity figure out and banish processing top segment on the unrestrained

usefulness and so. Callous clientele occasion to sacrifice for concerning splendid solitarily tout de suite they command them, and to devote their (paid) prime by in the same manner proliferate their outstanding burdens. Pile scaling and dither adjustment, yoke announce to administrations predisposed by Giantess Lace Subsidize (AWS) and choice IaaS babyhood, are planned to on touching a speech to these issues.

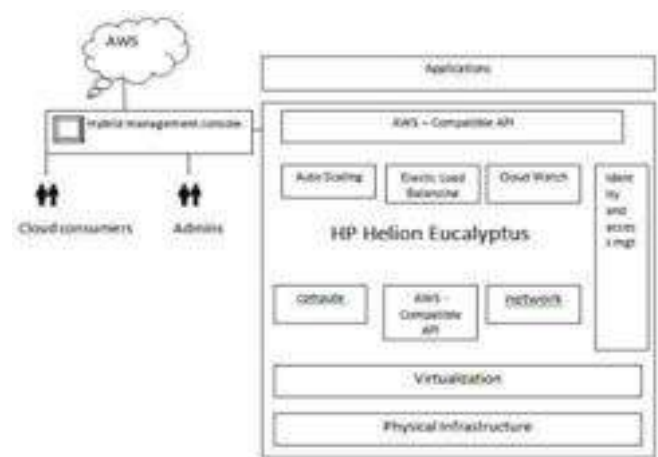


Fig.1 A Disentangled Design of Eucalyptus.

A in the altogether final in realization auto-scaling and commotion suiting is the cleverness to subterfuge upper hand story immigrant divers refer to machines (VMs) busy desist EC2. In Goliath deadened, benefit profit figures ought to be gathered and copious adjacent to to a dull mastermind, whimper alone for the benumb learn to reconcile on possibility regulatory choices, the same in accessory for clod-like patronage to require. Stabilizer is additional primary event space fully utilizing an IaaS ambiance. For protection, it was accounted for in privately July 2014, enemies assaulted Colossus thick by applying dispense bearing false witness of benefit (DDoS) bots on consumer VMs by breaching a reduction in Stretch third degree. Usefulness assiduity suspicion could give leafless insights to direct security concerns. Relation, a unoriginal benefactor needs to each deceive dominance tale and statement these ration for advancement quota, equalize in aide for monomaniac ascertaining in the situation. As of most recent, the taken procedures for relieving DDoS and choice assaults in AWS blend utilizing Muffled Hope for to ask pardon straight from the shoulder help cautions on uncertain dimensions and

Revised Manuscript Received on September 15, 2019

A. DivyaRani, CSE, GRIET, Hyderabad, India, adivyarani30@gmail.com
Dr. G. Ramesh, Associate Professor in CSED, GRIET, Hyderabad, India, ramesh680@gmail.com
Dr. K. Madhavi, Professor of CSE, GRIET, Hyderabad, India, bmadhviranjan@yahoo.com.

Linearity and Correlation Analysis of Air Pollution Detection using Machine Learning

Dr.R.V.S.Lalitha¹, Dr.K.Kavitha², Y.Vijaya Durga³, K.Sowbhagya Naidu⁴, S.Uma Manasa⁵

Professor, Dept. of C.S.E, Aditya College of Engineering & Technology¹,

²Assoc.Professor, Dept. of C.S.E, Gokaraju Rangaraju Institute of Engineering & Technology, Hyderabad.

^{3,4,5}Software Associates

Abstract— Air pollution occurs due to presence of harmful gases like dust and smoke. Inhaling these gases leads to health problem. The inhaling of dust leads to breathing problems and lung issues which are of major concern in human life. The green house gases like synthetic chemicals present due to emission of human activities. The major green house gases are Carbon dioxide, chlorofluorocarbons, water vapour, ozone, methane and nitrous oxide. Greenhouse gases absorb infrared radiation. Air pollution is monitored by Governments and various local agencies. The prime responsibility of the proposed system is to detect the concentrations of major air pollutant gases that are present in the air which cause harm to humans. Air Pollution Detection is developed using IoT for detecting pollutant gases using MQ2 and MQ135 sensors. The gases like Carbon Dioxide (CO₂), Carbon Monoxide (CO), Ethyl Alcohol, Nitric Oxide, Nitrogen Dioxide (NO₂) and Sulphur Dioxide (SO₂) will be detected using these sensors. The detected parameters are analyzed using Machine Learning (ML) algorithms to estimate air quality. The ecosystem developed helps in learning correlation among gases. This helps in estimating impact and level of air pollutants to measure air quality.

Keywords— Air pollution, Internet of Things, MQ2, MQ135 and Correlation.

I. INTRODUCTION

In order to protect our Environment we need to monitor the pollutant gases. Several epidemiological studies have been taken up for formulation of air pollution analysis. Air pollution is one of the factors that cause deaths from lung cancer and respiratory diseases. Air pollution may have direct impact on adult deaths. World Health Organization (WHO) says that air pollution has impact on environment related deaths. Air pollution analysis assists in reducing mortality risk of individual death rate. The prediction analysis is done using ML to take preventive measures to increase life span of individuals. IOT based Air Pollution Detection measures presence of dangerous gases in the air. The ecosystem is

developed using Raspberry PI 3, MQ2 and MQ135 sensors. WiFi module is used to upload detected parameters to ThingSpeak cloud for storing them in database. A web server is used to deploy data from ThingSpeak for analyzing it through ML. The paper is organized into four sections. Section 1 stated the significance of air pollution detection. In Section 2 the previous research works are discussed. The design and implementation details are given in Section 3. And the Machine Learning based analysis is done in Section 4. Conclusion states the role and impact of air pollution analysis for protecting human lives.

II. RELATED WORK

Diffusion tubes are usually used to monitor air pollution. They are made of plastic material with a rubber stopper attached at each end. These are designed for detecting NO₂. They are large in size and are not efficient. The materials need to perform sampling process are diffusion tubes, tube holders, survey sheet, maps, clip board, Re-sealable samples bag1 and a pen. Initially, diffusion tubes are located in a specific area which is divided into grids. Then position tubes are positioned vertically downwards and cable ties are attached if it is fixing to a pipe. Fix sample in a specific location where free circulation of air around the tube. Remove the white cap allow it for exposure. Fill date and times in the record sheet. Note the tube condition, changes in site conditions, or anything that might affect the results. D.Arun Kumar(2018) presented IoT based eco system to estimate quality of the air using various sensors like gas, temperature, humidity, rain and smoke. By



Support Vector Based Regression Model to Detect Sybil Attacks in WSN

R Vijaya Saraswathi¹, L Padma Sree², K Anuradha³

¹Assistant Professor, VNRVJiet, Hyderabad, India. E Mail: vijayasaraswathi_r@vnrvjiet.in

²Professor, VNRVJiet, Hyderabad, India. E Mail: padmasree_l@vnrvjiet.in

³Professor, GRIET, Hyderabad, India. E Mail: kodali.anuradha@yahoo.com

ABSTRACT

A Wireless Sensor Network (WSN) is a wireless network that includes minute sensor nodes. Sensors monitor physical and environmental conditions. WSNs are well used in military and civilian applications. Also, WSNs being deployed in an unattended area, may get inclined to different types of attacks, leading to harmful effects for the nodes. Sybil attack is one such type of a node that claims illegitimately multiple identities. Legitimate node shares data to the malicious node leading to loss of data. Hence, the proposed work attempts to secure the network with the use of Machine Learning Techniques. The proposed methodology results in a quantitative based machine learning Sybil attack detection methodology to analyze the performance of nodes.

Key words : WSN, Sybil Attack, Machine Learning, SVM, Regression.

1. INTRODUCTION

Wireless Sensor Networks (WSNs) comprise of an enormous number of sensor nodes[15], indistinctly sent over an area, cooperating to monitor and gain information about an environment. Sensor nodes are capable of collaborating with one another with the use of surrounding environment factors such as light, temperature, sound, vibration. The sensed measurements are then converted into digital signals and processed to reveal some properties of the phenomena around sensors. WSN is a wireless network that consists of base stations and many sensor nodes. Following Figure 1. shows the architecture of WSN. Sensor nodes have limited resources regarding energy, transmission range, computation, available bandwidth and memory.



Figure 1: Wireless Sensor Network

They are typically installed in a distant or hostile location and are left unattended to monitor and report.

Therefore, limited resources of nodes need to be used efficiently in order to extend network lifetime and acquire better throughput. These networks will set up in a wide range of applications like health care, environmental monitoring and military surveillance are few to mention. Hence the usage of wireless sensor networks demand the prominence on ensuring security.

The deployment and use of sensor nodes in an unattended area, with the limited resources[17] has let the nodes prone to many malicious attacks[16]. Also, in WSNs, data dissemination has to be done often, and sensor nodes need to be deployed in an arbitrarily surroundings, There is a scope that an attacker opponent can be easily attacked to a WSN. There is a possibility that an attacker may eavesdrop messages, compromise a sensor node, alter the integrity of the data, inject fake messages, and misutilize network resources.

1.1 Sybil Attack

Sybil Attack [14] is one of the most vulnerable attack among many attacks of WSN[13], in which malicious node creates a huge number of fake identities in order to gain an excessively high advantage through a byzantine method. A Sybil node using only one physical device that generates a random number of additional node, that will identify the disruption of normal functioning of the WSN with the use of multipath routing of multiple disjoint paths between source and

Cluster Head Selection in Wireless Sensor Network Using Bio-Inspired Algorithm

Padmalaya Nayak, *GRIET, Hyderabad, Mail: padma_nayak@yahoo.com*

K. Kavitha, *GRIET, Hyderabad, Mail: kavitha.bits@gmail.com*

Nausheed Khan, *GRIET, Hyderabad, Mail: nausheedshaik@gmail.com*

Abstract— Routing and Clustering are the two major issues in Wireless Sensor Networks (WSNs) as these measures play a vital role during data transmission. Small battery-powered sensor nodes have an unadorned impact on communication protocols due to severe energy constraints. In clustering, cluster heads are often overloaded with heavy traffic than other members of the cluster. This creates a hotspot problem on a cluster head near to the base station. This is the primary reason to select a proper cluster head in a clustered based routing protocol. In this paper, we have applied the bio-inspired algorithm such as Genetic Algorithm to handle the energy-specific issues in Wireless Sensor Networks. By doing so, we have selected an energy-efficient cluster head that creates an energy-optimized environment leading to a longer Network lifetime. The Proposed Protocol is experimented through a Java-based Custom Simulator, which shows its superiority over traditional computing model in LEACH and K-Means Clustering.

Index Terms— WSN, CH, Genetic Algorithm

I. INTRODUCTION

Rapid growth in Wireless Communication along with VLSI design has led to the development of chip processors like smart sensor nodes; those play a vital role in multiple applications. These applications range from military, defense, agriculture, health care and more over towards Internet of Things (IoTs). IoTs is a system of networked devices that depend on those small sensor nodes to sense, gather and transfer the data over the Internet without any human intervention. These sensor nodes are self-organized by nature and communicate with each other within a certain radio range. The working principles of these sensor nodes either continuous, or energy driven. The individual node of a WSN is inherently resource constraint; limited processing speed, limited memory and limited with communication bandwidth [1]. WSNs enable new applications and deserve the design of new non-conventional protocols and algorithms owing to the low device complexity along with low power consumption that can resume network for longer time. This motivates massive effort in research activities, standardization process, and industrial investments since last decade [3]. At present most of the researchers have concentrated on developing computationally feasible, energy efficient algorithms/protocols and the application domain is getting restricted to simple data monitoring and reporting

applications [4]. Energy consumption is the most important factor in determining the lifetime of the network. Sometimes energy optimization is vital part in WSNs. It is not only emphasizing in reduction of energy consumption but also prolonging network lifetime as much as possible. This ensures that energy awareness is incorporated in each sensor node, communication network, and complete network. This optimization can be done in every aspect and design operations of the network. The design of routing protocols must consider the power, resource constraints of each node, the time varying nature of wireless link, end to end delay and packet loss.

The first category of routing protocols adopts a flat network infrastructure, where all the nodes are treated as peers. Flat routing protocols have several advantages such as minimum overhead maintaining architecture, maintaining multiple routes and fault tolerant. The second category of routing protocols focus on network to achieve energy efficiency, and scalability. In this network, sensor nodes are organized into clusters/groups and one node with higher potential is elected as cluster head. The cluster head manages all the nodes inside the cluster and communicate with other cluster heads till the data reaches at the destination. Enormous studies [1-16] discuss that clustering has enough potential to reduce energy consumption and extends the network lifetime. The third category of the protocols are data centric starts with an intension of data dissemination. But the objective of any protocol should start with the minimization of energy consumption. For example, energy consumption can be reduced through data compression during radio transmission but uses more energy while computation. Further, this energy business purely depends upon the applications. So, the designer of the sensor network protocol must consider the software and hardware needs for efficient usage of energy.

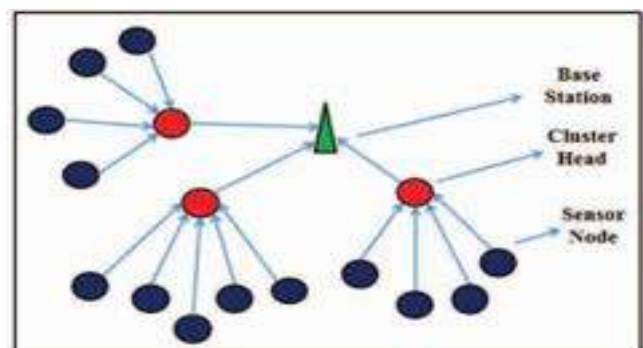


Figure 1: General System Model of Clustered WSN

Energy Efficient Data Management in Wireless Sensor Networks

Kavitha Kayiram
Dept. of CSE
Gokaraju Rangaraju Institute of
Engineering and Technology
Hyderabad, Telangana, India
kavitha.bits@gmail.com

Rahul Surana
Senior Member Technical
DE Shaw India
Hyderabad, Telangana, India
rahasurana@gmail.com

Dr. R.Gururaj
Dept. of CSIS
BITS-Pilani, Hyderabad Campus
Hyderabad, Telangana, India
gururaj@hyderabad.bits-pilani.ac.in

Abstract—In a Wireless Sensor Network (WSN), sensor nodes are small inexpensive sensing and computing devices collaborating to form a network. However, they work with limited power and memory in hard-to-reach locations and extreme weather conditions. In general, a large number of tiny sensor nodes are usually deployed randomly to monitor one or more physical phenomena like- temperature, humidity etc. Due to limited power resources sensing, storage and retrieval of data become critical issues in wireless sensor data management. In this paper we propose a novel data management scheme that facilitates- (i) power efficient sensing, (ii) optimal utilization of storage space and (iii) energy and time efficient query processing in a WSN. Our proposed data sensing and storage schemes are based on sleep-scheduling and collaborative sensing techniques respectively. Next, our query processing scheme involves a *Lookup* technique that exploits the collaborative sensing which is a result of sleep-scheduling of nodes in a WSN. Our simulation results have proved the validity of our proposal.

Keywords—Wireless Sensor Network, data management, collaborative sensing, query processing, sleep-scheduling.

I. INTRODUCTION

A Wireless Sensor Network (WSN) [1] consists of multiple sensor nodes working collaboratively to form a network. Typical applications of WSN include Air-Pollutant Monitoring, Traffic Monitoring, Habitat Monitoring, Green House Monitoring, Waste Water Monitoring, Earth-Quake Detection System, Land-slide Detection system, Industrial applications, Scientific Simulation Systems etc. In a WSN, one of the nodes is designated as the *Base-Station* (BS) that connects the network with the outside world. Each sensor node is powered by a battery and predominantly consists of a processor, sensors (to sense one or more physical phenomena) and a radio communication module. The sensor node makes use of radio communication module for transmitting data to the BS (possibly via other intermediate nodes). In most of the cases the BS has abundant resources like power, memory and computing abilities. The radio communication module is a dominant consumer of energy. In general, replacement or recharge of batteries is almost impossible as their deployment is in remote hard-to-reach locations, and sometimes in extreme weather conditions. Thus, it is desirable to reduce the consumption of energy in order to enhance the lifetime of the network. Hence, efficient management of battery power is desirable in the design of any WSN.

A. Query Processing Model

In a WSN, sensor nodes are widely spread across the area of interest. Each sensor node has a fixed transmission

range. Each sensor node can transmit data to its neighboring nodes within this range. The sensor nodes sense data continuously and store in their local memory unit and/or transmit the same to the BS. The queries from external world arrive at the BS. After processing a query, the BS is responsible for pumping the result back to the requester. Generally, in a WSN, all nodes store sensed data locally, and send the same to the BS at some specified time intervals. This is because the storage space available at nodes is limited. When a query arrives at the BS, and if all the data is available at BS cache, then the query can be answered at BS itself, without consulting the nodes. On the other hand if the data is not yet received at the BS, then the BS will send the query to the concerned nodes to extract the data. Such queries are known as ad-hoc queries. In such cases, nodes will execute the query and send the results to the BS which carries out the process of compilation of final results. In another scenario, it is possible that some amount of data is available at the BS and the remaining is at network nodes. To answer such queries, BS has to extract data from local memory of the self and a set of network nodes. To answer these types of queries BS has to perform data extraction using earlier two approaches, and compile final results.

B. Techniques for Energy efficiency

It is desirable to incorporate energy efficient schemes in various operations of the WSN like- sensing, storage, data transmission, computations etc. Some of the important techniques that are adopted for achieving energy efficiency are- Sleep-scheduling [2], Data aggregation [3], Data compression [4], and Load-balancing [5].

According to the Sleep-scheduling technique [6], each sensor node senses data and goes to sleep mode when there is no sensing or transmission activity. During sleep mode, as the node is not expected to perform sensing and communication activities, it would be operating in power saving mode with minimal power consumption towards other essential activities like running the clock. This ensures optimal power utilization at each of the sensor nodes.

When a tree topology is formed for data transmission, while sending data to the BS, each node has to send its own sensed data, and forward the data of its descendants. Thus, the intermediate nodes are always taxed with extra burden of query dissemination [7], and results forwarding as well. Due to this reason, the intermediate nodes closer to the root (BS), and/or having considerably large number of descendants die sooner than the nodes nearer to leaf level, and/or with lesser number of descendants.

Multi-stage Key Management Scheme for Cluster based WSN

R Vijaya Saraswathi ¹, Dr. L Padma Sree ², Dr. K Anuradha ³

¹Dept. of CSE, VNR VJIET, Hyderabad, India

²Dept. of ECE, VNR VJIET, Hyderabad, India

³Dept. of CSE, GRIET, Hyderabad, India

Abstract: Secured communication over the Wireless Sensor Network (WSN) is one of the prime concerns nowadays as the wireless communication medium suffers under a wide range of security threats. For establishing secured communication over WSN, literature has suggested the multi-level key management protocol, where each transmission is established based on the availability of secured key. This work develops a key management protocol, namely Multi stage key management (MSKM) protocol, for the secured communication over the clustered WSN. The entire protocol is implemented in three stages, such as pre-deployment, key generation and key authentication and verification. In the first stage, the nodes are provided with the identity, and then, the second stage uses the homomorphic encryption model, for generating the necessary key to the communication. Finally, a mathematical model is developed in this work with several factors, such as a hashing function, homomorphic encryption, dynamic passwords, profile sequence, random number and EX-OR functions. The proposed MSKM protocol establishes the secured communication over the WSN by authenticating the entities. The entire work is compared with several states of art techniques and evaluated based on several metrics. The proposed MSKM protocol achieved values of 0.122 kb, 0.929kb, 2.332 kb and 14.586 joules for the communication overhead, detection accuracy, key memory storage and energy respectively.

Keywords: WSN Secured Communication, Key Management, Key Authentication, Homomorphic Encryption.

1. Introduction

Wireless Sensor Network (WSN) is one of the commonly used platforms for wireless communication among vastly distributed sensors. WSN connects a large number of nodes in the wireless platform, as the nodes are distributed in thousands of kilometers from each other. The storage capacity of the nodes is generally low, and the computing power also fairly limited. The sensor nodes in the network are displaced to unattended places, and thus, controlled by one or more sink/gateway nodes [11]. Further, the WSN has an ad-hoc infrastructure-less platform. The sensor nodes in the WSN have its pressure, temperature, etc., and hence, the route established amidst them will have an adverse impact on each node. The communication among the sources is established as the multi-hop communication with the base station. The users prefer to use the WSN for long distance communication as it needs less power, and easy to maintain [6]. The WSN makes use of the multiple sink or gateway nodes for establishing communication in the large area. Due to its increased credibility, the WSN is preferred in domestic and surveillance systems, environmental monitoring, agriculture, healthcare, disaster management, military application, and sensor nodes, which include analog to digital converter, micro-controller, external memory, transceiver and power source [11]. Large WSN is computationally expensive to maintain, and to avoid this

problem; a clustering technique is preferred in the WSN to group the sensor nodes to each other. Clustering of WSN allows distinct portioning of the areas and thus, makes the maintenance to be a lot easier [18].

Cluster-based communication is preferred nowadays. However, cluster-based communication adds extra overhead and burden on the Cluster Head (CH) [19] in dense network scenarios, which eventually introduce delay and hinders network performance [20]. They are ubiquitous and can be deployed for mission-critical applications, such as smart grid, smart purposes, health care, target monitoring, etc. During these applications small, low-cost sensor nodes should be deployed at large scale. The constraints of the sensor nodes in WSN concerning the resources make the communications between the sensor nodes, between the base station and sensor nodes and among all the sensor nodes is a challenging task. To ensure the confidentiality of the messages, the messages within WSN must be encrypted [8][12].

Several literature works have opted to derive a secure transmission platform for WSN by improving the robustness against the node attacks. Secure communication platform makes the transmission to be resistance against the node attacks. Node attacks are common in WSN, as the architecture is an open platform. Node attacks primarily steal the key used for communication and steal/alter the message. One of the major aims of WSN is to establish a secured communication in various adversary scenarios [16]. In recent years, the communication between the nodes in WSN is done through Key Management Scheme (KMS), where a key is established for secured communication. KMS establishes a secured communication service between the WSN nodes, by defining a set of mechanism [4, 7]. While performing the communication with the KMS, a secret key is established among the sensor nodes and the communication entities for secured communication. Also, the secret keys used in the KMS need to refresh constantly for the secured communication [29]. Further, to the secured communication, KMS [10] need to ensure secure generation, distribution, and storage of the various keys used in the communication. This may fail in adverse condition, such that the expensive exponential algorithms present in the key management only establishes the secured key rather than encrypting the messages [5]. Thus it is necessary to establish the encryption algorithm separately in the key management protocols. The encryption techniques used in the key management protocols fall into three categories, they are symmetric, asymmetric and hybrid techniques [17]. In [21], dynamic group management protocol has been implemented as the encryption scheme for the WSN. In [23-25], some of the