



ISSN 2047-3338

An Outlier Detection and Rectification Method in Cluster Based Wireless Sensor Network

Dr. Laxman Sahoo, Alok Ranjan Prusty and Sanjaya Kumar Sarangi

Abstract— The evolution of Wireless Sensor networks (WSNs) is a demandable, efficient and emerging area of Computer Science Engineering which has been currently employed in various field of engineering particularly in communication system to make it effective and reliable. For successful application of WSN it is important to maintain the basic security level, both from external and internal attacks else entire network may collapse. In the field of WSNs, if the sensed patterns that significantly deviate from the normal pattern are considered as outlier. The possible source of outlier includes noise, errors, events and malicious attack on the network. WSNs are more likely to generate outlier due to their special characteristics like constrained available with the resources causing frequent physical failure and harsh deployment area. In this paper we have investigated and experimented different outlier attack in WSN and how these attacks are efficiently detected and rectified. The usual outlier detection techniques are not directly applicable to wireless sensor network due to the nature of sensed data, specific requirements and limitations of the WSNs. Our proposed method will train and test data set for detection as well as rectification of outlier due the malicious attack and identify the affected sensor nodes in a largely deployed cluster based wireless sensor network under a common outlier detection framework which is designed by some well specified supervised learning and classification based data mining technique.

Index Terms— Classifier, Cluster, Machine learning, Outlier, Outlier detection and rectification and Wireless Sensor Networks

I. INTRODUCTION

WIRELESS sensor networks are potentially one of the most important technologies ever made. Recent advancement in wireless communications and electronics has enabled the development of low-cost, low-power, multifunctional miniature devices for use in remote sensing applications. A sensor network is composed of a large number of sensor nodes which consist of sensing, data processing and communication capabilities.

Sensor networks are predominantly data centric rather than address-centric. So sensed data are directed to an area

containing a cluster of sensors rather than particular sensor addresses. Aggregation of data from cluster increases the level of accuracy, network scalability, robustness, efficient resource utilization on the other hand reduces data redundancy and power consumption. In the uncontrolled environments with dynamic topology there is always a chance at any instance any sensing node of the network become a victim of malicious attack and a major source of outlier for the entire network. Outlier, also known as anomaly, originally stems from the field of statistics. In WSNs, outliers can be defined as, “those measurements that significantly deviate from the normal pattern of sensed data”. Potential sources of outliers in data collected by WSNs include noise, errors, events and malicious attacks.

Outliers in WSNs can be analyzed and identified at different nodes in the network. Depending on the scope of data used for outlier detection, outlier may be either local or global. Local outliers are identified at individual sensor nodes. For detection, nodes should identifies the anomalous values only depending on its historical values or the alternative is that in addition to its own historical readings each sensor node collects readings of its neighboring nodes to collaboratively identify the anomalous values on the other hand the identification of global outliers can be performed at different levels in the network [5]. The sort of outliers caused by malicious attacks is very much harmful and concerned with the issue of wireless network security. So outlier due to these attacks should be detected and rectified before user access which is the main scope of our work and we will describe it through out this paper.

Section II, deals with the different types of malicious attacks that create outlier in WSN. We presented related works so far in section III. In section IV we have proposed and implemented successfully the rule based classifier method for detection and rectification of outlier in a cluster based wireless sensor network. By using this method we observe it is very easy to detect and rectify the outlier and the method can be used in practical applications of WSN. The experimental outcome of our proposed method is shown in section V. The conclusion and future scope of our work is reflected in section VI and VII respectively.

II. DIFFERENT MALICIOUS ATTACK IN WSN

One of the major challenges wireless sensor networks face today is security. While the deployment of sensor nodes in an unattended environment makes the networks vulnerable to a

Professor, School of Computer Engineering., KIIT University, Odisha, e-mail: laxmansahoo@yahoo.com.

Asst. Prof., Dept. of Computer Sc. and Engg., Koustuv Institute of Technology (KIT), Odisha. e-mail: alokprusty87@yahoo.co.in

Asst. Prof., Dept. of Computer Sc. and Engineering, Koustuv Institute of Self Domain (KISD), Odisha. e-mail: sanjaya_sarangi@yahoo.com

variety of potential attacks, the inherent power and memory limitations of sensor nodes makes conventional security solutions unfeasible. To achieve a secure system, security must be integrated into every component, since components designed without security can become a point of attack. The networked nature of large, ad-hoc, wireless sensor networks raises new threats and significant challenges in designing security schemes and five of the most pronounced challenges in Wireless Sensor Networks are in [11].

Attacks on the network system or network can be broadly classified as interruption, interception, modification and fabrication [12]. Security attacks in WSN can be classified into two major categories according to the interruption of communication act, namely Passive attacks and Active attacks [11]. The monitoring and listening of the communication channel by unauthorized attackers are known as passive attack.

TABLE I
TYPICAL THREATS IN WSN

Threat	Layer	Defense Mechanism
Jamming	Physical	Spread spectrum , Lower duty cycle, Tamper proof, key management scheme.
Tampering		
Exhausting	Link	Rate limitation
Collision		Error correcting code
Route information manipulating	Network	Authentication Encryption
Selective forwarding		Redundancy Probing
Sybil		Authentication
Sink hole		Authentication, Monitoring, Redundancy
Warm hole		Flexible Routing, Monitoring
Hello flood		Two way authentication, Three way handshaking.
Flooding	Transport	Limiting connection number, Client puzzle
Clone Attack	Application	Unique pair wise key

The Attacks against privacy is passive in nature, a passive attack it is said that the attack obtain data exchanged in the network without interrupting the communication. The unauthorized attackers monitors, listens to and modifies the data stream in the communication channel are known as active attack [13]. Again Network threats can also be classified into outside and inside treats. An outside attacker has no access to

most cryptographic materials in sensor networks, while an inside attacker may have partial key materials and the trust of other sensor nodes. Inside attacks are much harder to detect and defend .Typical attack and layer wise defence mechanisms in WSNs are summarized as in TABLE I [16].

III. RELATED WORK

Studding security threats as a main source of outlier in WSNs is a broad area. As compared to wired, in wireless networks it is a major challenging work. A detail description of WSNs architecture, applications, key design challenges, sensor network deployment, different localization algorithms, WSN characteristics, medium-access and sleep scheduling algorithms, energy efficiency and robust routing protocols, data centric wireless networking, different security mechanism are given by Bhaskar Krishnamachari in [3]. A volume of content about present status in sensor network security and research issues is presented by Perrig, J. Stankovic, and D. Wagner al, in [1]. Some of the security framework concerns with flexible routing, safe communication, and electronic and physical node destructions. The Sybil attack was given in [17],

Newsome et.al, it shows several variants in data aggregation, misbehavior and voting for cluster head. They have given effective security mechanisms against these different attacks for variants. Hu et al. examine the wormhole attack and suggest packet leashes to prevent an attacker from maliciously passageway packets to different areas in a WSN given [14]. Out of these attacks Sybil attack [8] and the wormhole attack [15] are very harm in nature. In [10] and [9] had discussed about two security protocols, SNEP and μ TESLA. In [7], Karlof and Wagner, review on sensor network routing protocol weakness and defense technique against several electronic attacks these protocols indemnify data discretion, authentication, purity and authenticated broadcast in severely resource constrained background like WSNs. Their model provide defense to sybil, wormhole, eavesdrop attack [1], spoof, respond and message modify attacks [18].

Attackers do traffic examination for determining locations while transmitting messages to the base station is discussed in [9]. In [9], J. Deng et al. have discussed for the protection of the base station from different attacks. Protection from Denial-of-service (DoS) attacks is a key challenge for researchers in WSNs. Wood and Stankovic study the attacks at different protocol layers in the network [2]. Where they have designed a time factors constraint which reduces network defenselessness to DoS attacks. In [9], A. D. Wood et al. have discussed about the radio frequency jamming DoS attack and presented a method to route around the jammed area of the network.

IV. THE RULE BASED CLASSIFIER METHOD

In a classification problem, we typically have historical data i.e., labeled examples and unlabeled examples. Each labeled example consists of multiple predictor attributes and one target attribute (dependent variable). The value of the target attribute is a class label. The unlabeled examples consist of the predictor attributes only. The goal of classification is to construct a framework using the historical data that accurately

predicts the label (i.e., class) of the unlabeled examples. A classification task begins with build data also known as training data for which the target values or class assignments are known. Data with known target values, to compare the predictions to the known answers such data is called test data. The application of a classifier to the new data is called scoring the data.

A. Bayesian Approach

We used Bayesian approach that is Naive Bayesian Algorithm for rule based classification of outlier for detection and rectification. Bayesian logic is applied to decision making and inferential statistics that deals with probability inference using the knowledge of prior events to predict future events. The essence of the Bayesian approach is to provide a mathematical rule explaining how you should change your existing beliefs in the light of new evidence.

B. Naive Bayesian Algorithm

This is a well established Bayesian method primarily formulated for performing classification task. Given a set of object each of which belong to a known class and each of which has known vector of variables, our aim is to construct a rule which will allow us to assign future objects to a class, by only the given vectors of variable describing the future object. Problem of this kind is called supervised classification. The naive Bayes method is easy to construct and may be readily applied to huge data set. This method is usually robust and does quite well.

For understanding the algorithm, let us consider [20] that there are two classes, labeled as $i = 0, 1$. Aim is to use the initial set of object with known class membership also called the training set to construct a score such that larger score is associated with class 1 object and smaller score with class 0 objects. Classification is then achieved by comparing this score with a threshold value t . If we define $p(i|x)$ to be the probability that an object with measurement vector $x = (x_1, x_2, \dots, x_p)$ then any monotonic function of $p(i|x)$ would make a suitable score. In particular the ratio $p(1|x)/p(0|x)$ would be suitable. According to elementary probability $p(i|x)$ can be decomposed as $f(x|i)p(i)$, where $f(x|i)$ is the conditional distribution of x for class i objects and $p(i)$ is the probability that an object will belong to class i if we know nothing further about it.

Therefore,

$$\frac{p(1|x)}{p(0|x)} = \frac{f(x|1)P(1)}{f(x|0)P(0)}$$

To understand the above classification we need to calculate the $f(x|i)$ and $p(i)$. To estimate the $f(x|i)$ the naive Bayes method assumes that the components of x are independent,

$$f(x_j|i) = \prod_{j=1}^p f(x_j|i)$$

Where, $j = 1, \dots, p$ and $i = 0, 1$

Naive Bayes algorithm used for binary and multi class classification. Naive Bayes makes prediction using Baye's theorem which derives a prediction from the underlying evidence. It tell that if A and B are two events probability of event A occurring given that event B has occurred is equal to the probability of event B occurring given that event A has occurred multiplied with probability of event A occurring and divided by the probability of event B occurring.

Mathematically,

$$P(A|B) = (P(B|A)P(A)) / P(B)$$

$$Prob(A \text{ given } B) = Prob(B \text{ and } A) * Prob(A) / Prob(B)$$

Here, each attribute is conditionally independent of other. Given a particular value of the target the distribution of each predictor is independent of other.

C. Outlier Detection and Rectification

Outlier detection is an important part of computer security. It provides an additional layer of defense against computer use after physical, authentication and access control. Outlier detection is used in the networks by comparing the set of baselines of the system with the present behavior of the system. Thus, a basic assumption is that the normal and abnormal behavior of the system can be characterized. Outlier detection describes the abnormal patterns of behavior, where abnormal patterns are defined beforehand. Fig. 1 shows the clustered based wireless sensor network architecture consisting of sensor nodes and cluster heads. All cluster heads gather the sense data send to the sink or base station. It is a two-tier hierarchical cluster topology [3].

The deployment of nodes with this topology make it easy for the multiple nodes of their local region to report to cluster head and in the heterogeneous settings of the network the cluster-head nodes seems to be more powerful in terms of computation and communication. Main advantage of this two-tier hierarchical cluster based approach is that it decomposes a large network into separate zones within which data processing and aggregation can be carried out locally. Data reach a cluster head they would then be routed through the second tier network formed by cluster-heads to another cluster-head or a gateway. Second tier network may utilize a higher bandwidth radio or it could even be a wired network.

This topology consists of two types of sensor nodes:

- (1) Forwarding nodes or simple sensor nodes which sense the activity and ability to send it to the base station.
- (2) Cluster head (CH) or simple data gathering point node, where all sensed data from the nodes are collected. As shown in Fig. 1, we have three clusters. Each cluster selects a cluster head which is responsible for collection of data from the sensor nodes of their zone and send to base station (BS) or sink. A clustering [6] based routing protocol called the Base station Controlled Dynamic Clustering Protocol (BCDCP) [7], [18], which distributes the energy dissipation evenly among all sensor nodes to improve network lifetime and average energy savings with the use of a high energy base station to set up cluster, cluster heads selection, cluster head to cluster head (CH-to-CH) routing path formation, schedule creation for each cluster and achieve other energy rigorous tasks. United Voting

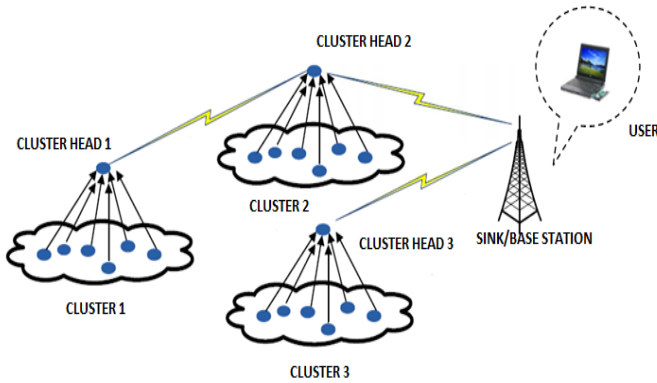


Fig. 1: A two-tier hierarchical three cluster based wireless sensor network architecture

Dynamic Cluster (UVDC) routing algorithm based on lingering of energy which periodically select cluster head according to residual energy among the nodes located in the event area in wireless sensor networks [19].

Our work is purely based on rule based classification method to detect and rectify the outlier patterns in the sensor nodes of the respective clusters. This rule based method can handle the security problem related to attack in cluster based wireless sensor network. Here in this method we use distributed data mining, machine learning and rule based classifier for providing solution against wireless sensor network. In wireless sensor network, initially, sensor nodes sense the action, and then report to their corresponding cluster head. All information is processed at cluster heads. Then, cluster heads send sensed data file to base station.

While gathering data file at cluster head, it may collect some erroneous data, or it is possible that some sensor nodes may send wrong information to CH. These data is causes outlier. In Fig. 2 we have integrated our method for rule based classification of outlier in the cluster based wireless sensor network. Fig. 3 deals with the skeletal structure of our proposed method. Before sending data file to base station cluster heads need to detect or remove the outlier. In the data file, we will detect the abnormal event information. For that

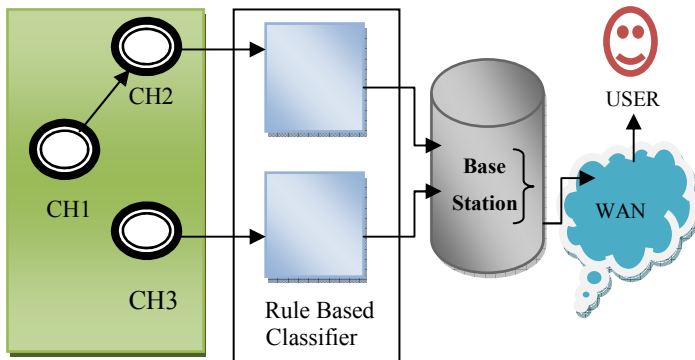


Fig 2: The Rule based outlier detection and rectification method is embedded in the cluster based wireless sensor network

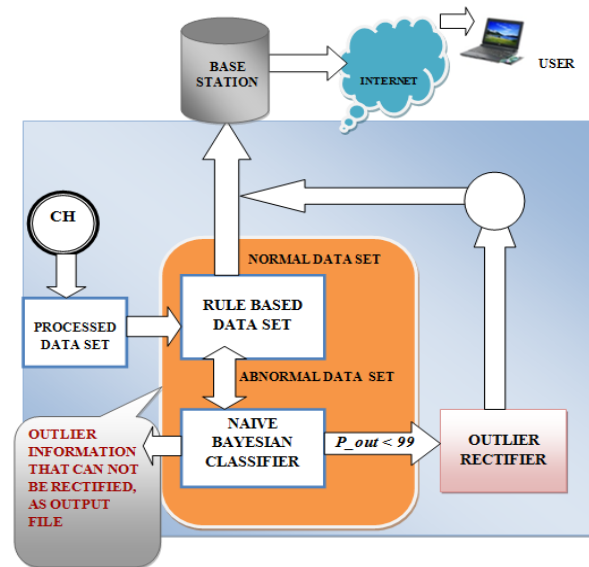


Fig. 3: Skeletal structure of our proposed method, here we have given the architecture of one CH, Similarly all CHs have their own method.

we have to execute our classifier based outlier detection and rectification method in-between cluster head and base station. Cluster heads ensures all outliers present in the data to be removed before it sends to base station.

In this scenario, each cluster head have its classifier for training the data. At first the information of all sensor nodes for a cluster has taken as processed file from a cluster head (CH). The file is then processed using rule based data set, if all processed trained data and processed test data are normal then it will pass the file to base station otherwise the data is abnormal and the file is not to be sent to the base station. We then apply naive Bayesian classifier to find the average value of outlier patterns from the abnormal data set, we have taken a threshold value for abnormal behavior of data collected.

If the average value of the outliers are more than the threshold then it signifies that majority of the sensor nodes with in that cluster is unhealthy, hence instead of forwarding these unhealthy data to the user via base station it will be stored as output file and if average value is less than the threshold then the rectifier module of our method will repair and rectify them to become normal data to be send to base station for user access via Internet which signifies that some the sensors not all in that particular cluster are creating outlier patterns. In each time the number of outliers detected, rectified and either stored as output file or sent to the base station. This process will continue until all outliers are detected for the data coming from all the cluster heads.

By implementing our method if an abnormal node sends any erroneous data to CH, the classifier finds out the abnormal nodes and, if an abnormal node is detected, either it will filter the individual node from the global networks or repair and rectify the erroneous data. The proposed LAS Algorithm illustrates how our method works.

D. LAS Algorithm

Input: n data files from n clusters.

Output: Number of outlier files for rectification and discard in percentage.

```

1. Outlier_Detection (Input: n data files)
2. {
3.    $j := 1, r := 1, d := 1, p\_out, th := 99;$ 
4.   for ( $i := 1$  to  $n$  data files)
5.   {
6.     read  $i^{th}$  data set file from  $i^{th}$  cluster head;
7.      $temp[j] = Naive\_Bayesian(i^{th}$  data file)
8.     if ( $temp[j]$  is an outlier)
9.     {
10.     $j := j + 1;$ 
11.     $p\_out := percentage(\%)$  of outlier
12.    if ( $p\_out < th$ )
13.    {
14.     $r ++;$ 
15.    rectifier ( $temp[j]$ );
16.    }
17.  else
18.  {
19.     $dis[d] := p\_out$ 
20.     $d ++;$ 
21.    return  $dis[d]$  as output file;
22.  }
23. }
24. }
25. }
```

E. Analysis

The LAS Algorithm takes as n numbers of processed data files. Each file is independent for each Cluster Head. In Line No. 3 we have taken 5 variable j, r, d, p_out, th and j, r, d initialized to 1 respectively, th is initialized to 99 as threshold value. In Line No. 4 for processing n data files we have taken for loop with a range of value for variable $i = \{1, \dots, n\}$. This loop will successfully executed for all n data files for n Clusters Head's of n clusters. Line no 6 will read i^{th} data set file as many times the loop will be executed. Line No. 7 will call the *Naive Bayesian classifier* program to train the classifier for detection of outlier for i^{th} data file from i^{th} Cluster Head, after this the result (i.e. average value of the outlier) stored in a test file that is here in the form of temp array $temp[j]$. In Line No. 8 the test file $temp[j]$ is tested with the classifier model, if the test file from i^{th} cluster is outlier then in Line No. 10 the counter j will be incremented, in Line No. 11 the percentage of outlier will be calculated and value stored in p_out . In Line No. 12 the percentage calculated in the Line No. 11 is compared with the threshold value th with a conditional code. if p_out is less than th then increment the counter r in Line No. 14. In Line No. 15 we call the *rectifier()* module and pass the value that we got in Line No. 7. In Line No. 17 else (i.e., p_out is greater than th) will be executed, then in Line No. 19 store p_out in $dis[d]$ and in Line No. 20 increment the counter d which tells that these no. of files can not be rectified with the rectifier hence these data file can to be send to the base station

for user access. In Line No. 21 we are display these outlier files as output file.

The percentage of outlier is calculated as:

$$\text{Percentage of outlier file} = \frac{(\text{total no. of outlier} \times 100)}{(\text{total no. of traces data})}$$

The Time complexity of the above algorithm is $O(n)$ as execution occurs for n data files with the help of for loop as a key operation.

V. EXPERIMENTAL OUTCOME

The simulation was based on the sensor network running NS2 (version 2.34) in ubuntu10.10 platform. We used 300 sensor nodes, three clusters. Each cluster head was elected using united voting dynamic cluster routing algorithm based on lingering energy in wireless sensor networks. All sensor nodes are constant bit rate transport protocol; we used Ad hoc On-Demand Distance Vector (AODV) as routing protocol. The movement of all sensor nodes was randomly generated over a $1100m \times 1100m$ field, with a maximum speed of 75m/s and an average pause of 10ms. Each simulation runs for a time period of 10,000 simulation seconds. We run this simulation for three types of malicious attacks. Using our method we calculate the percentage of outlier detected and percentage of outlier need rectification for respective cluster. The result was shown in Table II and Fig. 4.

TABLE II
PERCENTAGE OF OUTLIER DETECTED AND RECTIFIED

Percentage (%) of detection rate	Types of attack			Avg.	Rectify
	Hello flood	Sybil	Sink hole		
CH2	98.14	98.19	99.00	98.44	Yes
CH3	99.45	98.30	99.76	99.17	No

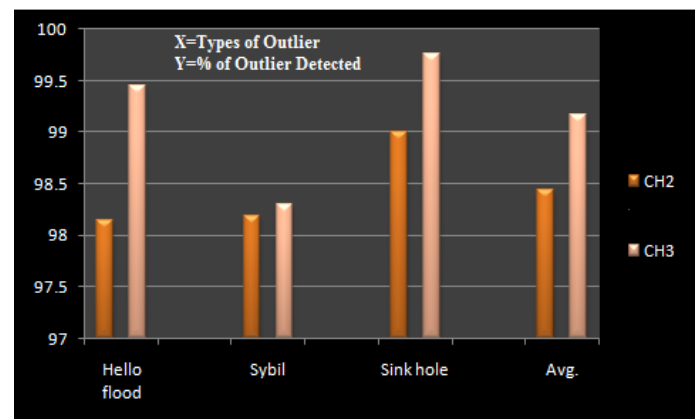


Fig. 4: Percentage of outlier detected and need to be rectified for CH2 and CH3 respectively

VI. CONCLUSION

In this paper we have proposed and implemented successfully the classifier based outlier detection and rectification method. We have used the Naive Bayesian technique as well as a rectifier module to detect and rectify the outlier files and each time send healthy and non erroneous data files to the user through the base station. Throughout the work the simulation result shows the success stories our proposed method for detecting and rectifying outlier patterns in a well clustered wireless sensor network. Here in our experiment we have taken the threshold value as 99%, the average value of all three class of outlier files for all two cluster heads compared with the threshold value. As per the LAS Algorithm flexibility is given if average value can not reach the threshold value then rectification of outlier is allowed other wise store as unhealed output data file. In our experiment CH2 with avg. value 98.44 need rectifications on the contrary CH3 with avg. value 99.17 shows a highest level of outlier hence stored as output file. To achieve accuracy we apply the classifier to a perfect training set of data with known classification.

VII. FUTURE SCOPE

As a future work the functionality of the LAS method can be upgraded by detecting the affected Clusters with different outliers as well as the affected individual nodes can be identified, for this we have to make little change to the architecture of every individual sensor node i.e., Each sensor node may have a series of flag bit with possible values 1/0 for particular outlier or malicious attack. If the data set file containing outlier information from a particular cluster then by training and manipulating the data with the reference or pre processed data we can able to track the malicious vulnerable nodes and healthy nodes under a affected cluster as well as we can calculate which node is how much percentage affected with which outlier. But the constrained is consumption of more energy, more memory requirement and powerful cluster heads. Efficient optimization of these constraints can be very helpful for the future work. Other direction of the future work will be optimization of the rectifier module for different class of outlier.

ACKNOWLEDGMENTS

We like to thank our parents, other family members and specially to the founder of KIIT University, Dr. Achyuta Samanta for their continuous encouragement and support.

REFERENCES

- [1] Adrian Perrig, John Stankovic, David Wagner, "Security in Wireless Sensor Networks" Communications of the ACM, Page53-57, year 2004.
- [2] A. D. Wood and J. A. Stankovic, "Denial of service in sensor networks," in IEEE Computer, October 2002, pp. 54–62.
- [3] Bhaskar Krishnamachari "Networking Wireless Sensors". published in the USA by Cambridge University Press, New York in 2005.
- [4] Dr. G. Padmavathi and Mrs. D. Shanmugapriya, "A Survey of Attacks, Security Mechanisms and Challenges in Wireless Sensor Networks". (IJCSIS) International Journal of Computer Science and Information Security, Vol. 4, No. 1 & 2, 2009.
- [5] D. Janakiram, A. Mallikarjuna, V. Reddy, and P. Kumar, Outlier Detection in Wireless Sensor Networks using Bayesian Belief Networks, Proc. IEEE Comsware, 2006.
- [6] E. Shi, A. Perrig, "Designing Secure Sensor Networks," IEEE Wireless Communications, pp.38- 43, December, 2004.
- [7] F. Zhao and L. Guibas, Wireless Sensor Networks, Elsevier, 2004, pp.1-20
- [8] H. Chan and A. Perrig. Pike: Peer intermediaries for key establishment in sensor networks. In IEEE Infocom, 2005
- [9] J. Deng, R. Han, and S. Mishra, "Intrusion tolerance and anti-traffic analysis strategies for wireless sensor networks," in Proc. of the 2004 IEEE International Conference on Dependable Systems and Networks (DSN), June 2004.
- [10] J. Deng, R. Han, and S. Mishra, "Insens: Intrusion-tolerant routing in wireless sensor networks," in University of Colorado, Department of Computer Science Technical Report CUCS- 9393-02, 2002.
- [11] K. Casey. Security in wireless sensor networks. In Department of Computer Science and Software Engineering, Auburn University.
- [12] Stallings, W., (2000) *Cryptography and Network Security Principles and Practice*, Cryptography Book, 2nd Edition, Prentice- Hall, 0-13-869017-0.
- [13] T. G. LUPU "Main Types of Attacks in Wireless Sensor Networks" International Conference in "Recent Advances in Signals and Systems". in 2009, ISSN: 1790-5109, ISBN: 978-960-474-114-4.
- [14] Y. C. Hu, A. Perrig, and D. B. Johnson, "Packet leases: A defense against wormhole attacks in wireless ad hoc networks," in Proc. of 21st Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM), April 2003.
- [15] Y. C. Hu, A. Perrig, and D. B. Johnson, "Wormhole detection in wireless ad hoc networks," in Tech. Rep. TR01-384, Department of Computer Science, Rice University, June 2002.
- [16] Zoran S. Bojkovic, Bojan M. Bakmaz, and Miodrag R. Bakmaz "Security Issues in Wireless Sensor Networks"INTERNATIONAL JOURNAL OF COMMUNICATIONS Issue 1, Volume 2, 2008
- [17] J. Newsome, E. Shi, D. Song, and A. Perrig. The Sybil Attack in Sensor Networks: Analysis & Defenses. In Proc. Intl Symp on Information Processing in Sensor Networks, 2004.
- [18] C. Karlof and D. Wagner, "Secure routing in wireless sensor networks: Attacks and countermeasures," in Proc. of 1st IEEE International Workshop on Sensor Network Protocols and Applications, May 2003.
- [19] Guo Bin,Li Zhe, "United voting dynamic cluster routing algorithm based on residual-energy in wireless sensor networks". Journal of Electronics & Information Technology. 2007, 29(12).pp:3006-3010.
- [20] XindongWu · Vipin Kumar · J. Ross Quinlan · Joydeep Ghosh · Qiang Yang · Hiroshi Motoda · Geoffrey J. McLachlan · Angus Ng · Bing Liu · Philip S. Yu · Zhi-Hua Zhou · Michael Steinbach · David J. Hand · Dan Steinberg "Top 10 algorithms in data mining" Knowl Inf Syst (2008) 14:1–37 DOI 10.1007/s10115-007-01142.



Prof. Dr. Laxman Sahoo, Professor in School of Computer Engineering at KIIT University, Odisha. Being a product of Indian Institute of Technology (IIT), Kharagpur (1987) he served as Faculty member/ H.O.D Computer Sc. Engineering/Coordinator/ Director at Birla Institute of Technology (BIT), Pilani, BIT, Mesra, BIT, Jaipur etc. He has undertaken as Principal Investigator and co-investigator of many Govt. of India sponsored projects. He has guided many M.Tech and Ph.D Scholars including B.Tech, MCA, BMI, ECE students. He has number of publications in national and international journals, conferences and authorized several books. His area of interest is Object Oriented Fuzzy Database, Artificial Intelligence, Expert systems, Mobile Networks, Performance Evaluation of Computer System, Knowledge Based system, Wireless-Network, Software Engineering etc.



Prof. Alok Ranjan Prusty, Asst. Prof. in Computer Science and Engineering at Koustuv Institute of Technology (KIT), Odisha. He has received his B.Tech and M.Tech in Computer Sc. and Engineering from College of Engineering Bhubaneswar under Biju Patnaik University of Technology (B.P.U.T), Govt. of Odisha. He has attended many national work shops, conferences, seminars etc. He is a member of IE(India) and also ISTE certified. His area of interest is Wireless Sensor Network, Real Time System (RTS), Ad-hoc Network, Embedded System, Advanced Computer Architecture and Organization, Microprocessor, Data Communication and Networking System, Operating System etc.



Prof. Sanjaya Kumar Sarangi, Asst. Prof. in Computer Science and Engineering at Koustuv Institute of Self Domain (KISD), Odisha. He has number of publications in national conferences. He has received his M.Tech from SOA University, Odisha. His area of interest is mobile Ad-hoc and Wireless Sensor Network, Data communication and networking etc.