
Aditi Kansal  
Department of Computer Science  
Jaypee University of Information Technology  
Waknaghat, Solan, India  
E-mail: aditi.kansal4@gmail.com

Yashwant Singh  
Department of Computer Science  
Jaypee University of Information Technology  
Waknaghat, Solan, India  
E-mail: yashwant.singh@juit.ac.in

Abstract: Wireless Sensor Networks are very useful now a day. They are used to sense data. Nodes can detect patterns and match them. Wireless Sensor Network plays an important role in target tracking, military surveillance, traffic management and weather forecasting. It works on the basis of collection of data or information from different network and machine learning techniques are applied on them. It may be discovery of patterns and pattern matching but there are some limitations of communication capability due to limited bandwidth. Thus, devices are needed and used to improve the performance of sensors. Data mining techniques also use machine learning. Thus, devices help to increase reliability and get proper feedback and results. Time series data are result of software failure.

Keywords: WSN, Machine Learning, Artificial Neural Network, SVM, Fuzzy Logic, Rough Set

I. INTRODUCTION

Wireless sensor networks are networks in which sensors are deployed. Sensors are of less expensive. There are multifunctional sensors which can do many functions i.e., it can note the temperature and pressure. These help to track object and also used in monitoring objects. Sensor nodes are integrated and used for low power consumption units, integrated memory, and radio transceiver and energy source [1]. Wireless sensor networks have been used in various applications such as target monitoring, target tracking, home animation, sales tracking etc. Target tracking, vehicle tracking, forest fire detection, earthquakes etc are done in WSN but information about the location is not collected. So, by the use of localization techniques we estimate and compute the position of sensor nodes. Nodes are confined to low power embedded processor, sensors, memory storage, radio transceiver and power source. Cost of hardware techniques used are reduced and computational power has increased [2]. Decision taking and pattern matching play an important role in WSN. Only relevant information is sent and used in monitoring of event detection.

Unlike traditional wired and wireless networks, WSN has its own design and resources which include energy range of communication which is short, bandwidth which is low and the computation time is less. The ideal sensor node is smart, whose power consumption is less, is capable of fast data attainment, well grounded, economical needing minor care. Event detection mechanism is recognized as an indispensable component of most of the applications which facilitate efficient sensing of physical world using wireless sensor network. Event is classified as a pattern of infrequent or abnormal occurrences. But due to harsh environment there are faults in pattern matching and pattern monitoring domain. These have limited battery life. There are various real-life applications [3].

- Environment monitoring
- Event detection
- Habitat monitoring
- Health and medical monitoring
- Target tracking
- Surveillance monitoring

Event detection is one of the applications of data observation in wireless sensor networks. A large amount of sensor data is semantically processed and only relevant information is sent to the user. Event is classified as a pattern of infrequent or abnormal occurrences in the area of interest. Event processing runs on wireless sensor network with streaming data so the hardware constraints and the nature of data need to be considered [4]. The primary aim of event detection is to create perceptions by processing and mining physical world data captured by the sensors and adapting background knowledge to relate it to possible events. An event can be a simple (or atomic) or a composite event [14]. The events such as temperature > 90 or light > 50 come under simple events. Events like explosion detection which is a combination of two or more simple events come under composite event. These events are detected by taking readings from multiple sensors present on the sensor node. Simple events require only participation of a single sensor (say, light sensor) and composite events require two or more sensors for their detection.

Machine learning is a branch of artificial intelligence that concerns the construction of the system from which learning can be done from data [5]. This is very attractive to manually construct them. Machine learning is used in Web search, spam filters, recommender systems, stock trading and many other applications. Event patterns are found with the help of machine learning and data mining techniques. Historical data sets are used for comparing data (sensor data) and for pattern matching. A common task in machine learning application domains involves monitoring routinely collected data from many interesting events. The main focus of this special issue
of machine learning is on event detection tasks and machine learning algorithms that address these tasks [6]. Basically machine learning is a technique which is subfield of artificial intelligence that concerns a question that how we can improve the performance of computer programs with experience. The family of machine learning algorithms is very efficient in WSN communication. In this section we will use various machine learning algorithms and techniques for efficient communication in WSN. We will use machine learning in WSN for event detection. Various algorithms will be used here. In this paper, we will present event detection in wireless sensor networks using machine learning techniques and various algorithms based on those techniques. The sensors nodes detect various events in our system. A number of different events like object detection, target detection and many more events can be detected. There will be no formal specification or expert knowledge needed. Communication cost is reduced to a minimum as the raw data is processed.

The remainder of the paper is organized as follows. In section I, we discuss about the clusters in the network. In section II, we discuss about various machine learning classifications.

II. ARTIFICIAL NEURAL NETWORKS

The Artificial neural networks are kind of learning based algorithms. These basically work on principle of neuron. The first model of neuron contained two inputs and one output. Both the inputs should be active for correct output. The weights for both the inputs were equal and output was binary. The mathematical model functions of these are F: X → Y. The data flows from input nodes to the output nodes through the whole network. Ability to learn is important property of artificial neural networks and to adjust input/output weights to reflect the exactly learned function. Thus, for training an artificial neural network, a set of data for training are needed in which inputs are already mapped to get the possible output. For example classification of different numbers, the pictures can be considered as inputs and the numbers can be considered as outputs. In contrast to decision trees, inputs cannot be described in terms of characteristics. The neural network is also known as supervised offline learning algorithm. This consists of a training set, which has already been classified. Offline is how much necessary is the training set which will be used for classification. There are also unsupervised and online learning neural networks. For example a network which is used for learning the data model for sensor readings. Neural networks are very suitable for such problems where features or attribute-values pairs are not available. However, they have large memory and processing requirements like decision tree learning. There are also some techniques which are applicable in WSN for static classification problems such as data models or link quality estimation. They can be efficiently implemented even on standard sensor nodes because of their low requirements. An attribute selection measure is a heuristic for selecting the splitting criteria that best separates the given data partitioning criteria, D of the class labeled training tuples into individual classes.

III. DECISION TREE

The Decision Tree classifiers are usually successful in many diverse areas such as radar signal detection, remote sensing, expert systems and speech identification. It has capability to break down complex decision making process into simpler one, thus providing proper solutions which are easier to interpret. They can be easily represented in form of graphs or in form of rules. Chi square (C4.5) and ID-3 are the most widely used decision tree algorithms. The C4.5 computes the information gain and finds which feature splits clusters best. C4.5 takes feature of highest information gain and puts that information in the root node of the tree and then recursively computes the information gain for the subclasses until all the samples from training set are classified. In wireless sensor networks various problem can arise that how to classify links as good or bad based on the data such as signal strength or delivery rate.

IV. SUPPORT VECTOR MACHINES

In machine learning, support vector machines are supervised learning models with associated learning algorithms that analyze data and match patterns [15]. The basic SVM takes a set of input data and estimates, which of two possible classes forms the output for each input value, making it a non-probabilistic binary linear classifier. Consider example of a training value set, which are marked as belonging to one of two categories, an SVM training algorithm which builds a model that assigns new examples into one category or the other. An SVM model can be described as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as large as it can be. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on.

V. FUZZY LOGIC AND ROUGH SET THEORY

Rough set theory can be used for classification to discover structural relationships within imprecise or noisy data. It means the values used are discrete. Therefore, the attributes with continuous values attributes must be discretized before using them. Rough set theory is based on the establishment of the equivalence classes within the training data. All the samples of data forming an equivalence class are undetected and the samples are identical with respect to attributes describing the data. Rough sets can also be used for feature reduction and relevance analysis. The problem of finding the minimal subsets of attributes that can describe all off the concepts in the given data set is NP-hard. However, algorithms to reduce the computational intensity have been proposed. In one method, for example, a discernibility matrix is used that stores the differences between attribute values for each pair of data samples. Rather than searching on the entire training set, the matrix instead searched to detect redundant attributes. Fuzzy logic is useful for machine learning as it provides advantage of working at high level of abstraction. The use of fuzzy logic in rule based systems involves the following: Attribute values are converted to fuzzy values. Fuzzy logic comes in when conventional logic fails. Fuzzy logic can deal
VI. REINFORCEMENT LEARNING

Reinforcement learning (RL) is biologically considered as one of the machine learning technique. In this technique, learning agents gain its knowledge by directly interacting with its environment. This can be explained by example. A mouse trying to find cheese in a maze and it has to select a direction to move [7]. This is a reinforcement learning technique in which agents select actions and receive rewards from the environment.

Agent states- A set of learning agents was considered which consist of finite set of possible states T and c to be considered as the coordinate representing the state of the agent at time step t. The current state of the mouse is in maze. Actions- Q-Learning basically associates a different set of actions and to each of the states in T. Let the actions of them mouse are left, right, backward, forward. Immediate rewards. There is an associate with each example and in our example, the changes in the states not leading to the goal have immediate rewards of 0(no cheese) and the ones leading to the goal state have an immediate reward of 1(cheese reached). Only the moves with their associated rewards from its current state can be seen by agent. The global knowledge about the environment is not there but only the states information and the rewards. Action costs- There are also costs in addition to rewards which are associated with each action in each state. This is a scalar value which represents the cost of the action. In this example the cost is one unit of energy of one cheese bite for the movement of the mouse. The negative reward costs are directly subtracted from immediate reward. Value function- The value function represents the expected total accumulated reward, in contrast to immediate rewards which are associated to each action in each state. These are easily observable; the value function represents the expected total accumulated reward. The goal of the agent is to learn a sequence of actions with a maximum value function, such that the reward on the taken path is maximized.

Updating a Q-Value: A simple rule exists to update a Q-Value after each step of the agent: Q(c+1, T) = Q(c, T) + γ(R(c, T) − Q(c, T)). The new Q-Value of the pair {c+1, T} in state c+1 after taking action at in state C is computed as the sum of the old Q-Value and a correction term, which includes the received reward and the old Q-Value. γ is the learning constant. It prevents the Q-Values from changing too fast and thus oscillating.

The total computation of the received reward is as follows: R(c, T) = r(c, T) + c(c, T) (2); where r(c, T) is the immediate reward as defined above and c(c, T) is the cost of taking the action at in state c.

Exploration strategy (action selection policy): Learning is performed in episodes, e.g., the mouse takes actions in its environment and updates the associated Q-Values until reaching the cheese. After completion, a new episode begins, repeating until the Q-Values no longer change. The question is how to select the next action. Always taking the actions with maximum Q-Value (greedy policy) will result in finding locally minimal solutions. On the other hand, selecting always random (random policy) will mean ignoring prior experience and spending too much energy to learn the complete environment.

VII. RELATED WORK

A. Distributed Event Detection in Wireless Sensor Network for Forest Fires

The approach progresses by dividing the training observations into constant size groups of sample vectors [8]. These techniques work in incremental manner and hence considered useful and use only the fractional processed data at each step. In DFP-SVM or distributed fixed partitioning, we find the estimates of the hyper planes by series of incremental steps that occur at each cluster. Only the current estimation value is send and not all the past data as is there in the previous approaches. By following such a technique we achieve reduction in energy consumption followed by other advantages, some of which may include-reduction in complexity and increased efficiency of the system.

The setup involves deploying sensors which have the ability to sense carbon dioxide, carbon monoxide, temperature and other parameters that can predict fire. Each sensor will detect the requisite parameter and send it. This data will a large stream data. The clusters are created dynamically and this large stream of data is then passed and clusters are made which contain data that needs large storage value. The data is send from the cluster head to the base station. We use clustream algorithm which is basically used for clustering of large data streams. This algorithm generates such data about data termed as metadata along with some decision attributes which as generally estimated and predicted from the previous stored observations. The meta data that is in the tabular form is analyzed and recognized by which we can predict or detect the event.

B. Forest Fire Detection In Wireless Sensor Network Using Fuzzy Logic

Fuzzy Logic System is a technique by which we assign values ranging from 0 to 1 to all the variables involved. FLS structure consists of the following processes namely inference,
defuzzification and fuzzification. Crisp inputs and converted into fuzzy inputs by the technique of fuzzification. A function called membership function (MF) evaluates the truth degree for every input and output which are taken into consideration. The value of membership function is between 0 and 1 which ranges between the intervals of the chip. Usually triangle, trapezium and bell shaped curves are shapes which are used. This paper has taken five parameters into consideration in detection of fire [9]. These are smoke, temperature, carbon monoxide, ionisation and photoelectric effect. For the output, five attributes are considered: low, medium, high, very high, and very low. For temperature, smoke and carbon monoxide there are three variables: high, medium, and low and MF for distance has three variables as close, average and far. The inputs of the fuzzy logic are considered which are based on fuzzy rules and give a fuzzy output. The fuzzy rule can be written as IF a1 is T1 and a2 is T2 …and an is Tn THEN b is b n . There are various steps for detection of fire using fuzzy rules. For each crisp input to be taken, there are a lot of variables which are defined. These are temperature having values from 0 to 120°C and can have values V, M and H, smoke having values from 0 to 100ppm(categorized as L, M, H), humidity value ranges from 0 to 100% (categorized as L, M, H), light values ranging from 0 to 1000lux (categorized as L, M, H) and distance ranging from 0 to 80m categorized as Close, Medium and Far. By plotting the special input parameter along the X axis and projecting it on the vertical side on the side where the MF lies. The outputs are achieved by studying the relations of the different input parameters. A very good example is that the fire is low when we assume that the light, smoke and temperature is low, distance is far.

C. SWATS: Wireless Sensor Networks for Streamflood and Water flood Pipeline Monitoring

The system used has the main objective to permit the low cost and high accuracy monitoring of the water flood and the steam flood system. It identifies major problems (such as leakage, blockage, outside force that might cause the flow) and obstacles that would take place when developing pipeline networks so that systems with high reliability can be established [10]. The system is made so that there is no false alarm that may take place due to various reasons such as pressure, humidity change, environmental effects or the phase change. To detect the anomalies in pipeline networks of steam and water is difficult and sometimes be erroneous because these sensors have inherent problems and the other environmental effects give an impetus to it thereby reducing the accuracy level of these sensors and sometimes generating false alarms thus making this field a challenging field in sensor detection.

The problems are mainly due to the different size and the shapes of the pipes thus varied pressure that is applied on the pipes and complexity of the pipe topological properties such as merge, split etc. Water and steam is transient so a single sensor is not enough for the detection. All these problems are taken into account and they are solved by making a multisensory algorithm that employs multimodal sensing capacity. The sensor accuracy is increased by combining the inputs that are received from multiple sensors and also studying the data correlations that is present among the different attributes. This technique applies a novel method for application in water flood system and also includes the localization and the identification of the steam and water.

D. Wireless Sensors and Neural Networks for Intruders Detection and Classification

The architecture used in this technique consists of both multi homing and multitiier approaches. The lower tier comprises of the sensor nodes that are spread over a large region in the space where the event is to be monitored. The sensor nodes propagate the data to gateways which comprises the middle tier of the architecture. The lower tier consists of the architecture of the system utilizes both multitiier and multi homing techniques. These gateways also deliver the packet to the central server. The middle and the upper tier communicate with each other by using the wireless fidelity (Wi-Fi). The cluster gateway, WSN cluster, the monitoring Client and the interface module are the basic networking modules. There are 20 modules in the WSN cluster. The sensor board is equipped with light and temperature sensor. An activation function as given in Eq.1 is applied to the ANN which is dot product so as to get the output of the system. The system involves a set of input values, a bias value plus synaptic weight.

\[ U_j(x) = \sigma (x \cdot w_j + \theta) \]  \hspace{1cm} (Eq.1)

Where \( u_j \) is the output of the \( j \)th neuron, \( \sigma \) the activation function, \( x \) the input vector, \( w_j \) the synaptic weight vector of the \( j \)th neuron, and \( \theta \) the bias associated with the \( j \) neuron. The activation function is usually a nonlinear function, e.g. hyperbolic-tangent or the logistic function.

E. Application of Wirless Sensor Networks in Forest Fire Detection under Uncertainty (Rough Set Approach)

Attribute reduction mechanism and efficient feature selection process is established by rough set approach. In this approach the granularity and the aggregation is established by the equivalence relations sometimes also termed as indiscernibility relations on the set of objects. We require the formation of a system in the data representation of rough set which is a pair

\[ S = (U, A) \text{, where } U \text{ is a nonempty.} \]

Finite set called the universe and \( A \) is a nonempty, finite set of attributes [11]. \( A \) is the decision attribute or class label. Using this technique the most critical aspects of forest fire is studied and rules are developed for its detection by building a robust model which also incorporates the missing values which might sometimes occur in real life situation .It has been also ensured that the performance will not go below a certain threshold in this system despite that the individual nodes fail. The sensors in a cluster are equipped with domain specific function procedures or lookup tables with limited computing capability. For each cluster \( S \) is \( (U,A) \) is generated dynamically by the cluster and it basically consists of the observations that are acquired from the sensor nodes. These values are sent to the cluster head.

F. Distributed Event Detection in Wireless Sensor Networks for Disaster Management

A tree which uses discrete and continuous values as input to form a graph is known as decision tree. It is a greedy approach [12]. There are two phases in construction of decision tree: training phase and testing phase. A set of data is taken as input
in training phase and minimum depth of the tree is found which will reduce the time complexity and memory space.

In above fig, every sensor is involved in detection of events. The results of all the values sensed by all the nodes are collected and sent to the voter which works on the basis of reputation based voting. A conclusion is made by finding the reputation of each and every node and the node with highest reputation value is selected. Firstly, it is assumed that all the events detected by sensors are correct. The detected values of all the nodes are sent to the neighbour's detection value table and every node decides according to the value of its neighbour sensors. The difference between these values is calculated i.e. the sensor node value and the values of other nodes. The comparison of the difference is then compared with the threshold values which are predefined. The voter gives a positive vote if the value is smaller than the threshold value and vice versa. To make a decision the values of the NDVT tables are sent to the voter. The most difficult task is to choose the node with highest reputation value which will decide whether event has occurred or not. This can be done by the use of voting techniques.

E. Reputation Technique 1
It evaluates the values detected by each and every sensor. Then these values are taken into consideration by comparing with the neighbourhood. Then average of all the values is calculated and reputation value of each sensor node is multiplied which is taken as weight. It can be calculated as

$$W_s = R_s \times A_{vs}$$

Where, $R_s$ is the reputation value of the sensors, $W_s$ is weight of sensors and $A_{vs}$ is average of all the values.

2) Reputation Technique 2
In this technique, two threshold values are predefined as $Q_1$ and $Q_2$ and these values can be assigned manually and the reputation value is compared with these threshold values which helps to decide whether decision is perfect or not. If $R_s \geq \theta_1$; decision is perfect. The decision is Ok if the value lies between both threshold values and poor if $R$ is less than $\theta_2$. 
Table I- COMPARISION OF ALL THE TECHNIQUES

<table>
<thead>
<tr>
<th>S.no</th>
<th>Technique</th>
<th>Machine Learning Technique Used</th>
<th>Advantage</th>
<th>Event Detected</th>
<th>Basic Methodology</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>Forest Fire Detection in Wireless Sensor Network Using Fuzzy Logic</td>
<td>1)Fuzzy Logic</td>
<td>1) Deeper analysis can be done to observe effect of each input on entire system. 2)Deals with linguistic variables</td>
<td>Fire</td>
<td>1) Uses five membership function (temperature,smoke,light, humidity and distance) as input.</td>
</tr>
<tr>
<td>D</td>
<td>Wireless Sensors and Neural Networks for Intruders Detection and Classification</td>
<td>1)Neural Network</td>
<td>1) Economical as it uses small cheap nodes that are self powering and self configuring. 2) Detect and classify different type of intruders.</td>
<td>Assets and national borders</td>
<td>1)Multi-homing technique allows for better and reliable identification of intruders. 2) The implemented networking modules are the WSN cluster, the cluster gateway, the monitoring server, the monitoring client, and the interface module</td>
</tr>
<tr>
<td>E</td>
<td>Application of Wireless Sensor Networks in Forest Fire Detection under Uncertainty</td>
<td>1)Rough set theories</td>
<td>1) Offers an attribute reduction algorithm and the dependency metric for feature selection. 2) Deal with data redundancy and inconsistency in a resource constrained environment</td>
<td>Forest fire</td>
<td>1)Unlike traditional compression algorithms, the rough set based dynamic feature selection algorithm allows the compression of the data stream without altering the underlying data semantics.</td>
</tr>
<tr>
<td>F</td>
<td>Distributed Event Detection in Wireless Sensor Networks for Disaster Management</td>
<td>1) Decision Tree.</td>
<td>1) Voting mechanism to reach consensus. 2)High interpretability</td>
<td>Forest Fire</td>
<td>1) Based on detecting events using decision tree classifiers running on individual sensor nodes and applying a voting to reach a consensus among detections made by various sensor nodes</td>
</tr>
</tbody>
</table>

VIII. CONCLUSION

The functionality of detecting the event is helpful and important when we need to detect the natural calamity such as fire or earthquake. It has been observed that many real world activities exhibit certain pattern which can be detected by applying certain machine learning techniques. In this paper we have studied and discussed variety of methods applying machine learning techniques used for event detection and a comparative analysis of all of them are summarized.
REFERENCES


ABOUT THE AUTHOR

Aditi Kansal received B.Tech from Maharshi Markandeshwar University, Mullana(Haryana), India in Computer Science in the year 2011. She is currently working as teaching assistant and is pursuing M.Tech. Degree in Computer Science and Engineering from Department of Computer Science, Jaypee University of Information Technology, Waknaghat, Solan, HP (India). His research area is Wireless Sensor Networks, Machine Learning and its applications.

Dr. Yashwant Singh is Assistant Professor (Senior Grade) in the Department of Computer Science and Engineering and ICT at JUIT, Solan. He has completed his Ph.D. in Computer Science from Himachal Pradesh University Shimla, year 2011. M. Engg from Punjab Engineering College Chandigarh (Now PEC University) in year 2006 and B.E. from Sant Longowal Institute of Engineering and Technology, Sangrur, Punjab in year 2004. He has published a number of papers in international journal and conferences.