Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 05, May - 2017

Pragmatic Approach of Data Mining in Wireless Sensor Networks:An Innovative Analysis

Use of Data Mining in WSNs

Mohd Muntjir
Department of Information Technology
College of Computers and Information Technology
Taif University
Taif, Saudi Arabia

Ahmad Tasnim Siddiqui
Department of Computer Science
College of Computers and Information Technology
Taif University, Taif, Saudi Arabia

***Abstract*—Wireless Sensor Network is a set of sensors that are integrated with a physical environment. Assortment of sensor nodes systematized into a network is also known as Wireless Sensor Networks .Now days, a wireless sensor networks have emerged as one of the most exciting fields in Computer Science research as well as the distributed systems. The main limitations of Wireless Sensor Networks are characteristics of sensor nodes and nature of sensor data generated by networks. In sensor networks data mining is the method of selecting application-oriented standards and patterns with acceptable accuracy from fast and continuous flow of data streams from sensor networks. Data mining methods has to be fast to process high-speed arriving data. Data mining, as the continuance of multiple intertwined disciplines, consisting statistics, machine learning, pattern recognition, database systems, information retrieval, World-Wide Web, visualization, and lots of application domains, has made great progress in the past decade. Data mining is one of the most imperative approaches by which useful patterns in data with nominal user intervention are known and available information of users and analysts to make decisions relayed on their vital organizations to adopt. To ensure that the advances of data mining research and technology will competently benefit the progress of science and engineering, it is important to scrutinize the challenges on data mining posed in data-intensive science and engineering and explore how to further develop the technology to assist new discoveries and advances in science and engineering. Main objective of this review paper is comparative analysis of various data aggregation; mining techniques are discuses with associate's advantage of accuracy, complexity, reduces energy consumption. Other factors such data mining techniques that affect the prediction are also discussed.**

Keywords*—Wireless Sensor Networks, WSN Partitioned WSN, Scheduling in WSN, Location-Based Scheduling, data aggregation,clustering,Data Mining.

I. INTRODUCTION

A Wireless Sensor Network (WSN) is a distributed network and it comprises a large number of distributed, self-directed, tiny, low powered devices called sensor nodes alias motes [13]. Wireless Sensor Network instinctively comprehends a large number of spatially distributed, diminutive, battery-operated, and embedded devices that are interacted to compassionately accumulate, procedure, and convey data to the users, and it has constrained computing and processing proficiencies. Motes are the small computers, which work jointly to form the networks. Motes are energy well organized, and multi-functional wireless device. [16]

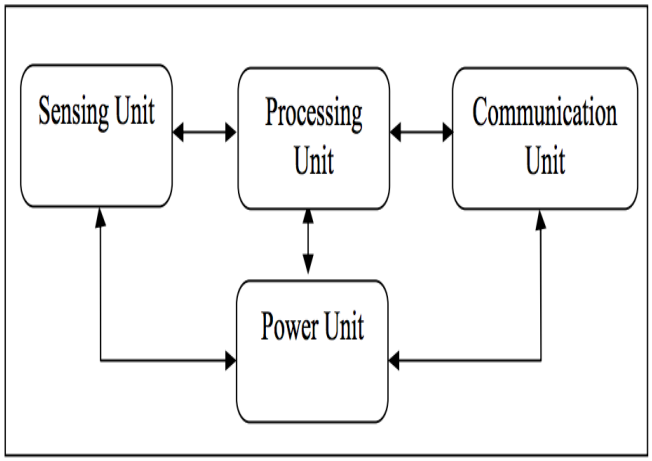


Fig. 1. Basic Building Blocks of Sensor Node

Wireless Sensor Networks device is rather basically applicable to a variety of fields. Basically, It is based on lesser nodes, radio transceiver, and battery. A Wireless Sensor Networks can overseer and assess roads automatically and continuously, with little human strength. A Wireless Sensor Networks can work during night even with poor weather circumstances, when there is fog or attendance of dust in the air. The wireless sensor networks (WSNs) perform function in autonomous manner in the spatial field to get precise values. Wireless Sensor Networks are a reasonably innovative application in the network, which specifies high quality monitoring for great environmental areas with relatively reasonable equipment [13]. Wireless Sensor Networks are collected of set of tiny sensor nodes, which can efficiently monitor their contiguous atmosphere. Due to the wide possible applications in battlefield observation, environmental observing, healthcare, weather forecasting, and disaster detection etc. Wireless Sensor Networks have concerned quite attention from both academic and industrial fields in current years [6] Wireless Sensor Network have a plentiful gains, but the vacant energy at each sensor nodes are treated as a constraint. Hence energy utilization is a major condition [9].

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 05, May - 2017

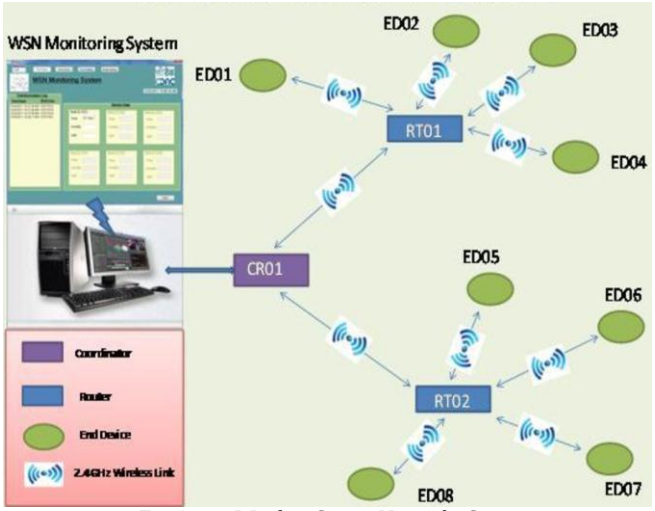
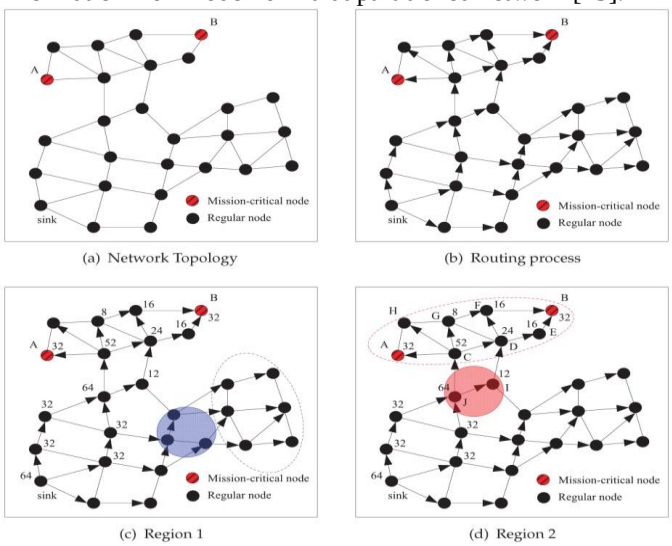


Fig. 2. Wireless Sensor Networks System

A. *Partitioned Wireless Sensor Networks*

The sensor network consequence with tree based topology which having one sink node and base station with four parent nodes. Every parent node is having link to one or more child nodes as shown in figure 4.

In the figure below Node 0 is sink node and parent node are3, 8, 11, and 21. Their consequent child nodes are shown through the link as shown in Fig 1. Consider power of node 11 is drain off completely which result in partitioned Wireless Sensor Networks, node 11, 16, 17, 18, 19 and 20 are package of sensor nodes in divided Wireless Sensor Networks. In segregated network sensed data from other nodes are temporary collected by earlier node of the parent node. Past mobile robot visits the parent node; MR gets the collecting information from that receptive node of parent node in partitioned system. Furthermore, for example consider partitioned network nodes 11, 16, 17, 18, 19 and 20. In this node 11 is being dull node. Hence, sensed data from 17, 18, 19 and 20 nodes are transitory composed by node 16, which is preceding node of the parent node. Further, past mobile robot visits the node 11; MR acquires the collected information from node 16 in that partitioned network [13].



(a) Network Topology (b) Routing process

(c) Region 1 (d) Region 2

Fig.3 . Partitioned Wireless Sensor Networks

B. *Scheduling in Wireless Sensor Networks*

Various scheduling method with different objective have been proposed for wireless sensor networks.

WSNs don't require a traditional scheduling problem; instead the aim is to optimize the data flow and processing, from source to sink.

1) *Time-Based Scheduling*

This is based on the time a node becomes the status of death node and near death node. Figure 1 indications that nodes 3, 8, and 11 are denoted as death node and the values of Death time for every node are in the order of node 8 < node 3 < node11. The acknowledging for above death time of parent nodes, node 8 is first lost node or going to dead node (close to death node) before node 3 and node 11. So, the mobile robot call node 8 first and then MR visit node3 and node 11 for data gathering method.

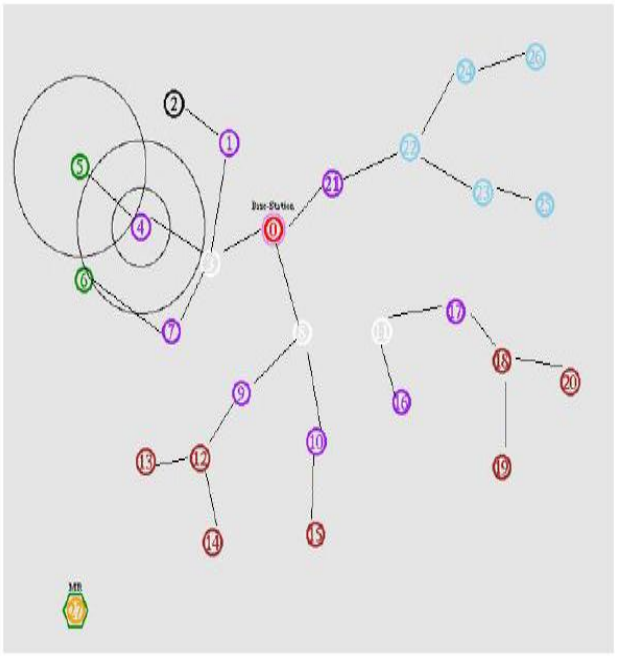


Fig. 4. Time Based Scheduling in WSN

2) *Location-Based Scheduling*

In most situations, the distance between the death node and the sink of the WSN is measured. Thus, location based scheduling (LBS), based on the distance amid the sink and the death node and near death node, is necessary. LBS is set to command the MR to visit the nearest the death node and near death node first to collect necessary information. The Euclidean distance formula calculates the distance between the sink and the location of the death node and near death node. For example, distance between sink and the dead node/near dead node as follows. Figure 5 shows that the distances of nodes 3, 8 and 11 linking the sink are in the sort of distance (3) < distance (8) < distance (11). Node 3 is little distance to the sink, when comparing distance between parent node and sink. So, the mobile robot visit node 3 first and then MR call node 8 and node 11 for data gathering process [13].

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 05, May - 2017

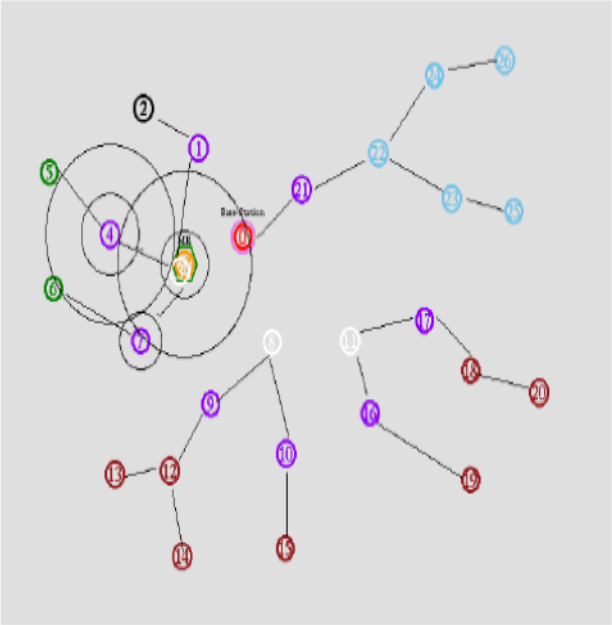


Fig. 5. Location Based Scheduling in WSN

C. Data Mining

Data mining is about processing data and recognizing patterns and trends in that information so that you can decide or review. Data mining ideologies have been around for many years, but with the beginning of big data, it is even more ubiquitous. Data mining is the process of determining actionable information from large sets of data. Data mining uses mathematical investigation to originate arrangements and trends that happens in data. Usually, these patterns cannot be revealed by traditional data investigation because the relationships are too complex or because there is too much data. The process of data exploration, discovery, and model building is frequently iterative as we objective and identify the distinctive information that can be extracted.

The rapid development of computer and information technology in the last twenty years has fundamentally changed roughly every field in science and engineering, transforming many disciplines from data-poor to progressively data-rich and calling for the improvement of new, data-intensive methods to conduct research in science and engineering [12]. Data mining is one of the most imperative methods by which useful patterns in data with minimal user intervention are known and accessible information of users and analysts to make decisions based on their vital organizations to adopt. Half-term data mining process involuntarily analyze large databases to find useful patterns can be applied. Data mining can be perform with relatively modest database systems and simple tools, comprising creating and writing our own or using off the projection software packages. Data Mining has emerged at the confluence of artificial intelligence, statistics, and databases as a technique for automatically discovering summary knowledge in large datasets. This presents to the process and main procedures in data mining, including, clustering, classification and pattern mining approaches [20]. Data mining systems and applications will also be covered, along with selected topics in current research. [17]

Fig. 6. Data Mining

Data mining is the process involves three steps are:

1. Initial excavation
2. Construction sculpts or gain credit with the help of pattern recognition / approval
3. Operation

Step 1: search. Usually this order with facts preparation will be done which may include data cleansing, data conversion and election the sub-set of fields is with enormous volume of variable (fields). Then according to nature analytical predictions, this order to precaution model simple or comments to recognize variables and determine the complexity of models for use in the next step requires.

Step 2: Production and verification of model validity.

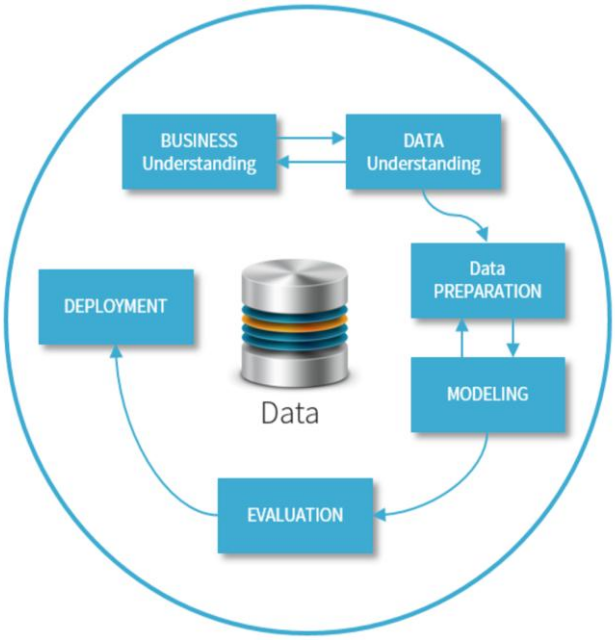


Fig. 7. Data Mining Process

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 05, May - 2017

This stage is to review different models and selecting the best model predicting the competence of the deals. Several techniques were developed to achieve this aim. And competitive estimation of models was called. For this purpose different models used identical data collection until be compared their competence, then the model that have the best performance, is chosen. This technique includes: Bagging, Boosting, Stacking and Meta-learning. Step 3: utilize. Last step before a model that has been selected, the work in new data until expectant`s precaution outs [10].

Knowledge Discovery in Databases is the process of examining for hidden knowledge in the huge amounts of data that we are precisely proficient of generating and storing that data. The Data is basically a collection of fundamentals, from which little knowledge can be collected. With the progress of data discovery techniques the value of the data is expressively enhanced. The basic task of KDD is to extract knowledge or information from lower level data or databases.

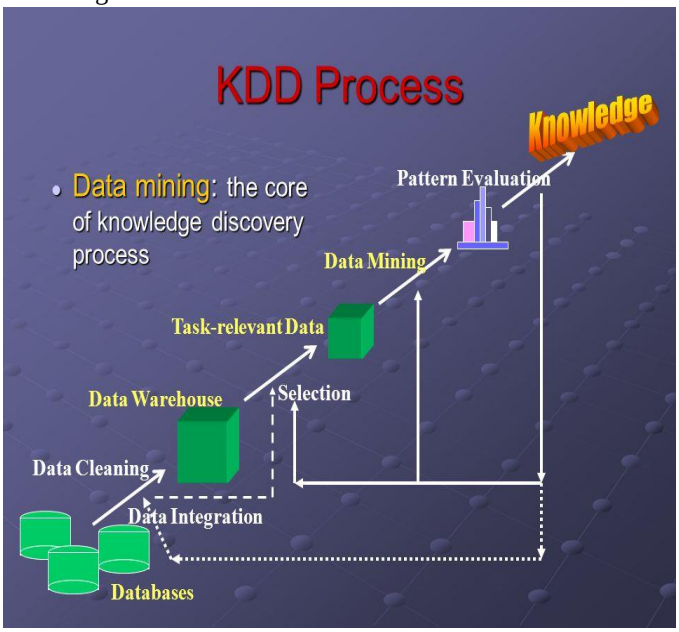


Fig. 8. KDD

1) *Cluster Analysis in Data Mining*

Determine the fundamental conceptions of cluster analysis, and then learning a set of typical clustering procedures, algorithms, and applications. Clustering analysis is generally used in many applications such as market research, data analysis, pattern recognition, and image processing. As a data mining purpose, cluster analysis operates as a tool to gain insight into the allocation of data to observe features of each cluster. Further, clustering can also help marketers discover distinct groups in their customer base. And they can characterize their customer groups based on the purchasing patterns. The process of clustering the facts objects into cluster is depended on the similarities and functionalities is data clustering. This procedure is much easier to gather and unlike sensors is located in the close clusters. The Alteration between the clusters is supported on the features of every cluster and they are unrushed by numerous distance functions like Manhattan distance and Eulidean distance [3].

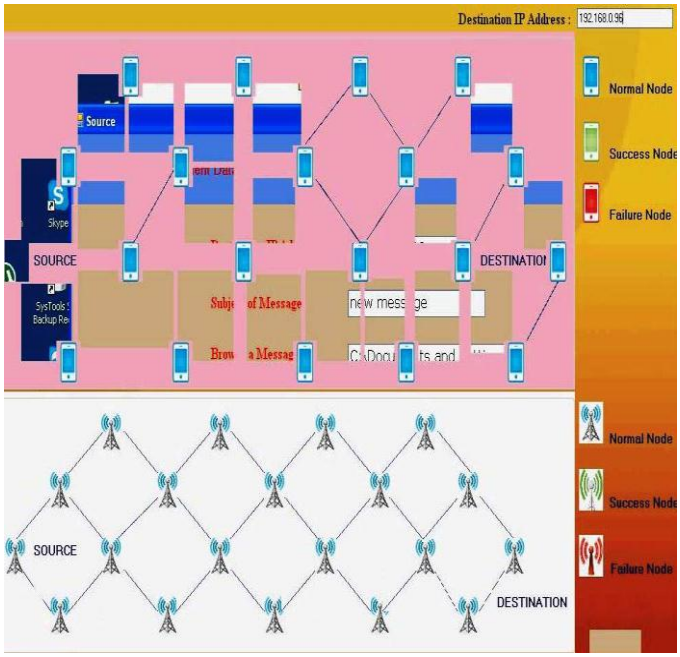


Figure 6. Data clustering for sensor networks

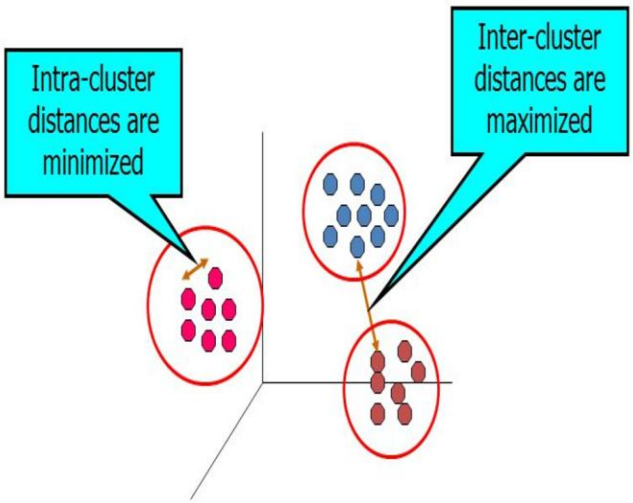


Fig. 9. Cluster analysis in data mining

2) *Data Aggregation in Wireless Sensor Networks*

WSN typically consists of a sink node referred to as a number of small wireless sensor network and a base station. Assumed the sensors nodes are to be unsecured with limited available energy while assumed to the base station are being secure with unlimited available energy. The sensor nodes control a geographical area and collect sensory information. Through wireless hop by hop transmissions the Sensory information is communicated to the base station. To conserve energy, by applying aggregation function on the received data this information is accumulated at intermediate sensor nodes. Aggregation reduces energy consumption on sensor nodes by reduce the amount of network traffic, It however complicates already existing security challenges for WSN and requires new security techniques. Providing security to aggregate data in WSN is known as secure data aggregation in Wireless Sensor Networks [21].

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 05, May - 2017

3) *Data Extraction Methodology*

In this section we present a distributed extraction solution to extract the required data from wireless sensor networks. Data extraction appears, to be a relatively straightforward component of a systematic review. Data extraction is the process of retrieving data out of data sources for further data processing or data storage. The network architecture consists of set of sensor nodes deployed in ad hoc fashion and reports the data to a well-equipped devise called the Sink that is linked to a database for storing the extracted epochs; each sensor will be attached with a flash memory for storing data. In this architecture, the extraction process starts by diffusing the mining parameters from the sink to all other nodes in the network. These parameters include the minimum support, time slot size, and the historical period. Upon receiving the mining parameters, each sensor will establish a local buffer that has a bit entry for each time slot in the historical period of the extraction. Initially, all the bit entries in the buffer are unset. After that, sensors keep track of the time and at the end of each time slot each sensor checks whether there is any detected event within this time slot [22]. If yes, the bit value corresponding to this time slot is set. At the end of the historical period each sensor traverses its local buffer, and if the number of set bits is greater than or equal to the minimum support, the node will establish a message (or a series of messages depending on the packet size) containing the sensor identifier and the time slot numbers where the corresponding bits are set. Then it sends this message to the sink. At the sink site, the sink will wait until it receives all the possible messages from the network and restructure the data in the messages in such a way that all sensors that report an event at the same time slot will appear in the same epoch. This epoch is then stored in the database. [19]

4) *Challenges*

According to the subsequent explanations [18] conventional data mining procedures for management sensor data in Wireless Sensor Networks are stimulating.

a) Resource Restriction

The sensor nodes are resource restrictions in terms of power, communication bandwidth, memory, and computational command. The foremost encounter faced by data mining techniques for Wireless Sensor Networks is to content the mining precision requirements while upholding the resource ingesting of Wireless Sensor Networks to a slightest.

b) Steadfast and Huge Data Entrance

The characteristic nature of Wireless Sensor Networks data is its high speed. Further, in many areas, data reaches faster than we are competent to mine. Moreover, spatiotemporal implanting of sensor data shows an important role in Wireless Sensor Networks application. This may basis many traditional data processing techniques to achieve inadequately on spatiotemporal sensor data. The contest for data mining techniques is how to manage with the incessant, speedy, and changing data becks and also how to integrate user communication during high-speed data influx.

c) Online Mining

In Wireless Sensor Networks, atmosphere data is physically dispersed, inputs attain unceasingly, and latest data items may alteration the consequences based on oldest data considerably. Most of data mining techniques that examine data in a disconnected manner do not chance the obligation of treatment allocated stream data. Therefore, a contest for data mining techniques is how to procedure dispersed flowing data online.

d) Modeling Variations of Mining Consequences throughout Time

While the data-generating occurrence is altering over time, the obtained model at any time should be conversant. Due to the steadiness of data streams, particular scholars have pointed out that apprehending the change of mining consequences is more significant in this area than the mining consequences. The research subject is how to perfect this change in the consequences.

e) Data Conversion

Since sensor nodes in WSNs are inadequate in terms of bandwidth, altering original data over the network is not realistic. Knowledge structure transformation is an imperative concern. After obtaining model and patterns locally from Wireless Sensor Networks data, the output are relocated to the base station. The experiment for data mining technique is how to competently characterize data and exposed patterns over network for communication.

f) Vigorous Network Topology

This topology in sensor network arrange in theoretically harsh, heterogenic, tentative and active atmospheres. Furthermore, sensor nodes may move among dissimilar positions at any point over time. Such heterogeneity upsurge the difficulty of scheming and suitable data mining technique for Wireless Sensor Networks. To discourse these contests, researchers have improved the conventional data mining techniques and also suggested new data mining algorithms to handle the data produced from sensor networks [18]

II. COMPARATIVE STUDY

Phiros Mansur & Sasikumaran Sreedharan et.al [1] represents some techniques which associate advantages of energy conservation and reduction of missing rate, accuracy of tracking. Moreover other factors like data mining techniques and clustering that affect the prediction.

Rouhollah Maghsoudi et al.in 2011[2] Data mining knowledge in reaction to technological advances in various Rmynh, base arena is built there. Data Mining face a different situation that the data size is large and want to make a small model and not too complicated and yet the data as well as describe. Requisite is to use data analysis to reduce the amount and the huge volume of information. One important and sensible facts in the world of machine intelligence and is robotics robots routing. Robot router has obstacle recognition and how to deal with the decision with barrier. For routing, algorithms including probabilistic methods (filtering particulate), evolutionary algorithms like as genetic, ants social and optimization particle mass, neural methods -

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 05, May - 2017

Fuzzy, inequality of matrix method relayed on gradient methods combined sensor information, etc. There are data mining schemes in the years 2010-2008 as a technique for routing and an absolute robot has been used and still is in progress. Overview of the methods in the paper mentioned in diverse articles since 2000 has so far. Although many data mining methods comprise, but mentioned in this article with precise literature data mining will deal with the routing problem.

Miao Zhao et al.in 2012[3] Recent study reveals that great benefit can be achieved for data gathering in wireless sensor networks by employing mobile collectors that gather data by short-range communications. To pursue maximum power saving at sensor nodes, intuitively, a mobile collector should traverse the transmission series of each sensor in the field such that each data sachet can be straight transmitted to the mobile collector without any relay. However, this approach may lead to considerably increased data gathering latency owed to the low moving velocity of the mobile collector. Fortunately, it is pragmatic that performing proper local aggregation via multi-hop transfers and then uploading the aggregated data to the mobile collector can effectively shorten data gathering latency. In such a scheme, the number of local transmission hops should not be randomly large as it may increase the energy consumption on packet relays, which would adversely affect the overall competence of mobile data gathering. Based on these observations, in this paper, author study the tradeoff between energy saving and data gathering latency in mobile data gathering by exploring a balance among the relay hop count of local data aggregation and the moving tour length of the mobile collector. Author first recommends a polling-based mobile congregation method and verbalizes it into an optimization difficulty, named bounded relay hop mobile data gathering (BRH-MDG). Exceptionally, a subset of sensors will be preferred as polling points that buffer locally aggregated facts and upload the data to the mobile collector when it arrives. In the meanwhile, while sensors are connected with these polling points, it is guaranteed that any packet relay is bounded within a given number of hops. Author then give two efficient algorithms for selecting polling points among sensors. The effectiveness of our approach is validated during extensive simulations.

Emad M. Abdelmoghith & Hussein T. Mouftah et.al [4] presents a amount of research work to minimize the volume of transmitted traffic by using data compression techniques and reducing power consumption levels in WSNs. In this paper, the present a data oriented approach called Model-based Clustering (MBC), which reduces, the flows of communication between sink node and sensor nodes.

Sherin Mathew et al.in 2013[5] lately, there has been a rapid growth in the wireless communication technique. Mobile sinks can be mounted upon urban vehicles with fixed trajectories supply the ideal infrastructure to effectively retrieve sensory data from such isolated WSN fields. Obtainable methods use either single-hop transfer of data from Sensor Networks that lie within the MS's range or intense involvement of network periphery nodes in data retrieval, processing, buffering, and delivering tasks. Our projected protocol aims at reducing the overall network overhead and energy expenditure associated with the multi-hop data recovery process while also ensuring balanced energy consumption among SNs and prolonged network existence. This is accomplished through building cluster structures contained of member nodes that route their dignified data to their allocated cluster head. Cluster heads implement data filtering upon raw data exploiting potential spatial-temporal data idleness and frontward the filtered information to appropriate finish nodes with sufficient residual energy, located in proximity to the MS's trajectory. Simulation consequences stipulate the superior concert of our projected algorithm to strike the appropriate performance in the power consumption and network lifetime for the wireless sensor networks (WSNs).

Daniele Apiletti & Elena Baralis et.al [6], Represents the complete design, validation, and implementation of the SeReNe framework. Given historical sensor readings, SeReNe acquire efficient sensor network data by discovering energy-saving models SeReNe exploits different clustering algorithms to discover temporal correlations and spatial which allow the identification of sets of sensor data streams and correlated sensors. Select the representative sensors from the given clusters of correlated sensors, to reduce the communication computation and power Costs, only the representative sensors are queried rather than directly querying all network nodes.

S. Anandamurugan & C. Venkatesh et.al [7], Presents several advantages of heterogeneous architecture for WSNs. It consists of some resource rich simple un-dynamic nodes and mobile relay nodes. The mobile relays have high energy as compare to un-dynamic nodes. The mobile relays help relieve sensors that are highly burdened by heavy network traffic and can dynamically move around the entire network, thus improving the lifetime.

M. Vijayalakshmiet al.in 2013[8] wireless Sensor Network (WSN) is an emerging technology. WSNs usually consist of a large number of small sensor nodes with limited onboard energy supply and deployed densely in a given region for information harvesting purposes. Since the sensor devices have limited memory and power capacity, the power consumption in WSN becomes as a major issue nowadays. So that, in the proposed framework, a scheme to reduce the power consumption in WSN is introduced. Proposed framework is clustering based. Clustering and Prediction procedures, which use sequential correlation among the sensor data, supply a chance for reducing the energy consumption of continuous sensor data collection. Thus it can attain stability and prolongs network lifetime. An adaptive scheme is presented which is used to control prediction, analyze the performance trade off between reducing communication cost and prediction cost, and design algorithms to take the benefit of adaptive scheme to

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 05, May - 2017

enable/disable prediction operations. Localized prediction scheme is performed which takes advantages over the previous dual-prediction scheme to minimize communication and computation cost thereby reducing the energy consumption. Sleep/awake scheduling can be applied. A practical algorithm designed for data aggregation will use faster and more efficient cluster-to-cluster propagation.

R.Sivaranjini et al.in 2013[9] Nowadays Wireless sensor networks playing vital role in all are Which is used to sense the environmental monitoring, Temperature, Soil erosion etc. Low data delivery competence and high-energy consumption are the inherent problems in Wireless Sensor Networks. Finding precise data is more difficult and also it will leads to added expensive to collect all sensor readings. Clustering and prediction techniques, which develop spatial and temporal correlation among the sensor data, provide opportunities for reducing the power consumption of continuous sensor data collection and to achieve network energy efficiency and stability. So as mention Dynamic scheme for energy utilization and data collection in wireless sensor networks (WSNs) by uniting adaptively enabling and immobilizing forecast arrangement, sleep or awake method with dynamic arrangement. Our agenda is clustering constructed. A cluster head demonstrate all sensor nodes inside the region and collects data values from them. Our outline is common enough to integrate many progressive features and show how sleep or awake scheduling can be pragmatic, which takes our framework method to designing a practical dynamic algorithm for data aggregation that prevents the need for rampant node-to-node broadcast of combinations, but rather it uses faster and more competent cluster-to-cluster dissemination.

S. Nithyakalyani et al.in 2012[10] Nowadays sensors are very essential for today life to monitor environment where human cannot get involved very a lot. Wireless Sensor Networks (WSNs) are used in many real world applications like environmental monitoring, traffic control, trajectory monitoring. It is more demanding for sensor network to sense and collect a large amount of data, which are continuous over time, which in turn must to be forwarded to sink for further decision-making process. Clustering of sensory data act as a center job of data mining. A clustering in WSN involves selecting cluster heads and assigning cluster members (sensors) to it for proficient data relay. The constraints in power supply, limited communication, bandwidth, and storage resources are the main challenges in Wireless Sensor Networks facing today.

S. Anandamurugan & C. Venkatesh et.al [11], Presents several advantages of heterogeneous architecture for WSNs. It consists of some resource rich simple un-dynamic nodes and mobile relay nodes. The mobile relays have high energy as compare to un-dynamic nodes. The mobile relays help relieve sensors that are highly burdened by heavy network traffic and can dynamically move around the entire network, thus improving the lifetime.

Onur Tekdas et al.in 2009[12] explore synergies among mobile robots and wireless sensor networks in environmental monitoring during a system in which robotic data mules gather measurements gathered by sensing nodes. A proof-of-concept implementation demonstrates that this approach considerably increases the lifetime of the system by conserving energy that the sensing nodes otherwise would use for communication.

Karthik in 2012[13] Data collection process is one of the significant aspects in the design consideration for future analysis in wireless sensor networks. Energy consumption by parent node increases owing to continuous forwarding of sensed data from their respective child nodes especially in the tree based topology. Once the power in the parent nodes was completely drained off, some of the child nodes get isolated/ partitioned from the sink node. The projected data collection method involves deployment of multiple mobile robots whose responsibility is to gather the data from the nodes whose energy is below the threshold value. Navigation of mobile robots to gather the data from partitioned nodes usually achieved by time and location based strategies. In projected hybrid scheduling, the navigation of mobile robots scheduled by both the combination of time and location based approaches with various region scheduling. In large network scenario, the mobile robot gets more burdens due to its extra responsibilities to visit all partitioned nodes. So the entire scenario is divided in to different regions and the deployment of multiple mobile robots is relayed on the requirements. Hence, the competence of sensed data comprised by the base station or sink node from partitioned or islanded WSNs is improved particularly using multiple mobile robots. Through simulation below the environment of NS-2 simulator, the results from various aspects show that proposed multiple mobile robots can develop the performance of collecting the sensed data in large-scale sensing fields and also it improves the lifetime of the sensor nodes.

Laxmi Choudhary in 2012[14] with the rapid improvement of computer and information technology in the last many years, an enormous amount of facts in science and engineering has been and will continuously be produced in massive scale, either being stored in gigantic storage strategy or flowing into and out of the system in the type of data streams. Moreover, such data has been made widely available, e.g., via the Internet. Such tremendous quantity of data, in the array of tera- to peta-bytes, has fundamentally changed science and engineering, transforming several disciplines from data-poor to increasingly data-rich, and calling for new, data-intensive methods to conduct research in science and engineering. In this paper, author discuss the research challenges in science and engineering, from the data mining perspective, with a center on the following issues: (1) information network analysis, (2) discovery, custom, and understanding of patterns and knowledge, (3) stream data mining, (4) mining moving object data, RFID data, and details from sensor networks, (5) spatiotemporal and multimedia data mining, (6) mining text, Web, and other unstructured data, (7) facts cube-oriented

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 05, May - 2017

multidimensional online analytical mining, (8) visual data mining, and (9) data mining by combination of sophisticated scientific and engineering domain knowledge.

Tzung-Cheng Chen & Tzung-Shi Chen et.al [15], Suggested from a wireless sensor network (WSN) a novel data-collecting algorithm using a mobile robot to get sensed data that possesses islanded/ partitioned WSNs is proposed in this paper. This algorithm allows the improvement of data collecting performance by the base station by identifying the locations of navigating a mobile robot and partitioned/islanded WSNs to the desired location. Two control approaches, a global- and local-based approach are proposed to identify the locations of the partitioned/islanded WSNs.

TABLE I. COMPARISONS BETWEEN SEVERAL TECHNIQUES

Ref No	Year	Findings	Techniques
1	2014	clustering and data mining	Make the prediction process more accurate that will help to reduce the missing rate.
2	2011	Solve the routing problems	Data mining methods
3	2012	Significantly increased data gathering latency due to the low moving velocity of the mobile collector	A polling-based mobile gathering approach and formulate it into an optimization problem, named bounded relay hop mobile data gathering (BRH-MDG).
4	2013	Model Based Clustering(data oriented approach)	Reducing energy power consumption level in WSN.
5	2013	Appropriate performance in the energy consumption and network lifetime for the wireless sensor networks.	Cluster structures consisted of member nodes that route their measured data to their assigned cluster head (CH
6	2010	energy-saving models to efficiently access sensor network data	SERENE prototype has designed to be stable, fast, able to manage a huge amount of sensor data stream and easy to access.
7	2010	AR (Aggregation Routing) Algorithm	Efficient Improve lifetime of the network. It has high energy then undynamic nodes.
8	2013	Stability and prolongs network lifetime	A practical algorithm designed for data aggregation will use faster and more efficient cluster-to-cluster propagation.
9	2013	Avoids the need for rampant node-to-node propagation of aggregate	Practical dynamic algorithm for data aggregation
10	2012	A clustering in WSN involves selecting cluster heads and assigning cluster members (sensors) to it for efficient data relay.	Clustering in WSN
11	2010	AR (Aggregation Routing) Algorithm	Efficient Improve lifetime of the network. It has high energy then undynamic nodes.
12	2009	Significantly increases the	Synergies among mobile

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 05, May - 2017

		lifetime of the system by conserving energy that the sensing nodes otherwise would use for communicatio n.	robots
13	2012	1.Improves the lifetime of the sensor nodes 2.Improves the performance of collecting the sensed data in large-scale sensing fields	Multiple mobile robots
14	2012	Fundamentally changed science and engineering, transforming many disciplines from data-poor to increasingly data-rich	The research challenges in science and engineering, from the data mining perspective
15	2011	GBA(global-Based approach LBA(local Based approach)	Improve sensed data collecting performance in apportion or islanded WSNs.

III. CONCLUSIONS

Applying Data Mining techniques reduces the size of data, deletes redundancy, improves the Wireless Sensor Networks speed and decreases the network traffic to extend the life-time of the network to guarantee short time of decision as early as possible. In this paper, we have introduced a comprehensive framework for mining wireless sensor networks, which consists of some interpretation for the sensors suggestions' rules problem. Our work is based on determining and relating real data from different sensors. Our methodology takes into account all features of a Wireless Sensor Networks (WSNs) that respects low energy capability, computing limitation, low memory capacity of sensor nodes, and conservational circumstances. Wireless sensor network technology has the prospective to enable main breakthroughs in the natural sciences by giving scientists the potential to collect high-fidelity data over large geographic regions and extended periods of time. In WSNs, since the sensor nodes are energy constrained and have limited lifetime, energy consumption of sensor nodes becomes as a major issue.
Two main approaches:
1) Clustering-based: sensor nodes form clusters and elect the cluster heads in such a way to improve energy efficiency, and
2) Prediction based: energy-aware prediction is used to find the slight trade-off among communication and prediction cost. Via performance evaluation, it is shown that it achieves energy efficiency even though the entity arrived from any random location and moves arbitrarily. K-Means Data Relay (K-MDR) clustering algorithm for Wireless Sensor Networks moderates the communication transparency and increases the absolute network life time by reducing the number of transmission between every sensor node to sink. Furthermore, the K-Means Data Relay algorithm declines the computational time and develops the enactment of the network when compared to K-Means algorithm. Hybrid moving based scheduling approach for data collection process is enhanced lifetime of the network. Our work is established on evaluating and conjoining real work from different sensors Concerning Data Mining techniques reduces the size of data, improves the WSN speed and deletes redundancy .We also suggest to assured the message altercation among nodes in order to have a appropriately dwindling network. For large handling area and more number of nodes catastrophe in the network, several mobile robots used for dispersed partition in order to gather data.

ACKNOWLEDGMENT

We would like to gratefully and sincerely thank to The Dean and Vice Dean of our College (College of Computers and Information Technology) and Chairman of our Departments for his guidance, understanding, patience, and most importantly, his friendly nature during this research paper. We would also like to thank my friends and colleagues, and the university who provided me an efficient support to work on this atmosphere and good infrastructure. We would also like to thank to all the previous researchers who worked very hard and helped others to comprehend the subject of Data Mining and Wireless Sensor Networks System.

REFERENCES

[1] Phiros Mansur, Sasikumaran Sreedharan, "Survey of Prediction Algorithms for Object Tracking in Wireless Sensor Networks", IEEE International Conference on Computational Intelligence and Computing Research, 2014

[2] RouhollahMaghsoudi,SomayyeHoseini,YaghubHeidari,"Surveying Robot Routing Algorithms with Data Mining Approach",Vol .2 No.2 (2011) 284-294

[3] Miao Zhao, Member, IEEE, and Yuanyuan Yang, Fellow, IEEE," Bounded Relay Hop Mobile Data Gathering in Wireless Sensor Networks", Vol. 61, No. 2, Febuary 2012

[4] Emad M. Abdelmoghith, Hussein T. Mouftah, "A Data Mining Approach to Energy Efficiency in Wireless Sensor Networks", IEEE 24th International Symposium on Personal, Indoor and Mobile Radio Communications, 2013.

[5] Sherin Mathew, S. Nandhini," A Novel Approach for Sensory Data Collection in Wireless Sensor Networks with Mobile Sinks",Volume 2, Issue 2, March 2013

[6] Daniele Apiletti · Elena Baralis · Tania Cerquitelli, "Energysaving models for wireless sensor networks", London Limited : Springer-Verlag, april 2010

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 05, May - 2017

[7] S.Anandamurugan, C.Venkatesh, "Increasing the Lifetime of Wireless Sensor Networks by using AR (Aggregation Routing) Algorithm", IJCA, 2010

[8] M. Vijayalakshmi, V. Vanitha,"Cluster Based Adaptive Prediction Scheme For Energy Efficiency In Wireless Sensor Networks", Vol 04; Special Issue; June 2013

[9] R.Sivaranjini, E.Dinesh,"An Adaptive scheme for energy consumption and data collection in Wireless Sensor Networks",Volume 2 Issue 3 ǁ March. 2013 ǁ PP.50-55

[10] S. Nithyakalyaniand S. Suresh Kumar,"Data Relay Clustering Algorithm for Wireless Sensor Networks: A Data Mining Approach", 2012

[11] S.Anandamurugan, C.Venkatesh, "Increasing the Lifetime of Wireless Sensor Networks by using AR (Aggregation Routing) Algorithm", IJCA, 2010.

[12] Onur Tekdas And Volkan Isler,University of Minnesota Jong Hyun Lim and Andreas Terzis,Johns Hopkins University,"Using Mobile Robots To Harvest Data From Sensor Fields",2009

[13] V. Karthik,"Region Based Scheduling With Multiple Mobile Robots for Data Collection Strategies in Wireless Sensor Networks",Vol. 2, Issue 7, July 2013

[14] LaxmiChoudhary," CHALLENGES FOR DATA MINING",Volume 2, Issue 2 (February 2012)

[15] Tzung-Cheng Chen, Tzung-Shi Chen, "On Data Collection Using Mobile Robot in Wireless Sensor Networks", IEEE Transactions On Systems, November 2011.

[16] http://shodhganga.inflibnet.ac.in/bitstream/10603/22912/7/07_chapter_01.pdf

[17] Neville, "Data Mining ",https://www.cs.purdue.edu/homes/neville/courses/CS57300.html

[18] [AzharMahmood, Ke Shi, ShaheenKhatoon, and Mi Xiao," Data Mining Techniques for Wireless Sensor Networks: A Survey", Hindawi Publishing Corporation International Journal of Distributed Sensor Networks Volume 2013, Article ID 406316, 24 pages, http://dx.doi.org/10.1155/2013/406316

[19] AzzedineBoukerche and SamerSamarah," A Performance Evaluation of Distributed Framework for Mining Wireless Sensor Networks ∗", Proceedings of the 40th Annual Simulation Symposium (ANSS'07) 0-7695-2814-7/07 $20.00 © 2007 IEEE,

[20] Mohd Muntjir, Ahmad TasnimSiddiqui, Mohd Rahul"An Improved Data Mining Technique and Online Mining for Energy Efficiency in Wireless Sensor Networks: A Comparative Review", International Journal of Computer Science and Telecommunications [Volume 7, Issue 7, October 2016], Page-19-24

[21] Muntjir, Mohd. "Implementation of Data Mining in Wireless Sensor Networks: An Integrated Review." International Journal of Computer Science and Network 5.04 (2016): 692-698.

[22] Muntjir, Mohd. "An Augmentation of Data Collection Functioning With Implementation by Pattern Variation Discovery in Wireless Sensor Networks: Novel Overview with Applications.", Asian Journal of Technology and Management Research,Vol 6,Issue 1,June 2016,Page-75-84

Author's Info

Mohd Muntjir is working in College of Computers and Information Technology Taif University at Taif Saudi Arabia. He received his M.C.A. degree from H.N.B. Garhwal University Uttarakhand India and Ph.D. degree in Computer Science from OPJS University, Rajasthan India. He is a member of professional societies like ACM, IEEE, VAS, IJETAE, and CSTA. His interested are mainly in Wireless Sensor Networks, Cloud Computing Systems, and Database Management Systems. The author has published many research papers in distinctive journals, conferences and books/book Chapters.

Ahmad Tasnim Siddiqui is working in College of Computers and Information Technology Taif University at Taif Saudi Arabia. He did M. Phil. In Computer Science, Madurai Kamaraj University, Madurai, TN, India .He has completed his Ph.D. degree in Computer Science from OPJS University; Rajasthan India. He is a member of professional societies like ACM, IEEE, etc. His interested are mainly in E-Commerce, E-Business, E-Learning, and Database Management Systems. The author has published many research papers in distinctive journals, conferences and books/book Chapters.