Adaptive Data Aggregation with Mobile Agents and Evolutionary Computing based Clustering in Sparse Wireless Sensor Networks

Naladala Mounika
Dept. of Computer Science & Engg
UCEK, JNTUK, Kakinada
Andhra Pradesh, India
e-mail: naladala2@gmail.com

Ratna Kumari Challa
Dept. of Computer Science & Engg
UCEK, JNTUK, Kakinada
Andhra Pradesh, India
e-mail:ratnamala3784@gmail.com

Kanusu Srinivasarao
Dept. of Computer Science & Engg
Yogi Vemana University, Kadapa
Andhra Pradesh, India
e-mail:kanususrinivas@gmail.com

Abstract—The Information processing based on Data mining in WSN is at its starting stage, when compared to traditional machine learning in WSN. In order to solve a particular problem in WSN the researchers now a day were mainly focused on applying machine learning techniques. The Different researchers will have different assumptions, application scenarios and preferences in applying machine learning algorithms. These differences will result to a major challenge in allowing researchers to build upon each other's work so that research results will accumulate in the community. Thus, a common architecture across the WSN machine learning community would be necessary in order to overcome these differences. The improvement or optimizing of the performance of the entire network in terms of energy conservation and network lifetime will be one of the major objectives in wireless sensor network. This paper will survey the Data Mining in WSN applications from two perspectives, namely the network associated issue and application associated issue. In the network associated issue, different machine learning algorithms applied in WSNs were used in order to improve network performance will be discussed. In application associated issue, machine learning methods that have been used for information processing in WSNs will be summarized.

Keywords*-Evolutionary Computing, Data Aggregation, Clustering, Wireless Sensor Networks*

*****

I. INTRODUCTION

Wireless sensor network (WSN) is widely considered as one of the most important technologies for the twenty-first century. In the past decades, it has received tremendous attention from both academia and industry all over the world. A WSN typically consists of a large number of low-cost, low-power, and multifunctional wireless sensor nodes, with sensing, wireless communications and computation capabilities. These sensor nodes communicate over short distance via a wireless medium and collaborate to accomplish a common task, for example, environment monitoring, military surveillance, and industrial process control.[2] The basic philosophy behind WSNs is that, while the capability of each individual sensor node is limited, the aggregate power of the entire network is sufficient for the required mission.

NETWORK DESIGN CHALLENGES AND ROUTING ISSUES

The design of routing protocols for WSNs is challenging because of several network constraints. WSNs suffer from the limitations of several network resources, for example, energy, bandwidth, central processing unit, and storage. The design challenges in sensor networks involve the following main aspect:

A.The capacity of energy limited:

The nodes which are battery powered will have less energy capacity. For the network designers in hostile environments this Energy will be an important challenge. For example, a battlefield, where it is impossible to access the sensors and recharge their batteries. The sensor nodes when they reaches a particular threshold, will be faulty and will not function properly, which will affect the network performance. In order to extend the network lifetime and performance of the network, the routing protocols that were designed for sensors nodes should be energy efficient.

B. Sensor locations:

Another challenge that faces the design of routing protocols is to manage the locations of the sensors. Most of the proposed protocols assume that the sensors either are equipped with global positioning system (GPS) receivers or use some localization technique to learn about their locations [2].

C. Limited hardware resources:

In addition to limited energy capacity, sensor nodes have also limited processing and storage capacities, and thus can only perform limited computational functionality. These hardware constraints present many challenges in software development and network protocol design for sensor networks, which must consider not only the energy constraint in sensor nodes, but also the processing and storage capacities of sensor nodes.

D. Massive and random node deployment:

Sensor node deployment in WSNs is application dependent and can be either manual or random which finally affects the performance of the routing protocol. In most applications, sensor nodes can be scattered randomly in an intended area or dropped massively over an inaccessible or hostile region [3]. If the resultant distribution of nodes is not uniform, optimal clustering becomes necessary to allow connectivity and enable energy efficient network operation.

NETWORK CHARACTERISTICS AND UNREALIABLE ENVIRONMENT

A sensor network usually operates in a dynamic and unreliable environment. The topology of a network, which is defined by the sensors and the communication links between the sensors, changes frequently due to sensor addition, deletion, node failures, damages, or energy depletion. Also, the sensor nodes are linked by a wireless medium, which is noisy, error prone, and time varying. Therefore, routing paths should consider network topology dynamics due to limited energy and sensor mobility as well as increasing the size of the network to maintain specific application requirements in terms of coverage and connectivity. [4]

A. Data Aggregation:

Since sensor nodes may generate significant redundant data, similar packets from multiple nodes can be aggregated so that the number of transmissions is reduced. Data aggregation technique has been used to achieve energy efficiency and data transfer optimization in a number of routing protocols.

B. Diverse sensing application requirements:

Sensor networks have a wide range of diverse applications. No network protocol can meet the requirements of all applications.

II. Related Work

S. R. K. Joel B. Predd, and H. Vincent Poor[1] had found out the problem in distributed learning in WSNs.They also discussed how parametric methods for distributed signal processing were inappropriate for applications where data is sparse and prior knowledge is limited. They also discussed distributed learning in networks with fusion center and distributed learning in adhoc networks with in-network processing.

R. G. C. Intanagonwiwat, and D. Estrin [3] had introduced Directed-diffusion and this is used for designing distributed sensing algorithms. This Directed-diffusion is data centric dissemination and also the reinforcement based adaptation for the best path and also for caching and in-network data aggregation.

A. C. W. R. Heinzelman, and H. Balakrishnan [4] had introduced LEACH, a cluster based routing protocol which reduces the global energy usage by distributing the load to all nodes at a time. LEACH also reduces communication energy. By distributing energy among all nodes will reduce the energy dissipation and enhancing system lifetime.

K. Sohrabi et al,[5] had presented a set of algorithms for establishing and maintaining connectivity in wireless sensor networks. These algorithms exploit low mobility and abundant bandwidth. They also developed different algorithms for setting up sub-networks to perform signal processing functions.

I. E. C. Chien and C. McConaghy [6] present a modem architecture that when the low power was dissipating it will provides a robust communications for wireless sensor networks. The FPGA implementation has been verified and said that this modem architecture can dissipates only 33 mW for both the transmission and the reception. The implementation can be directed to an ASIC technology with an estimated power performance of less than 1 row.

A. S. a. A. Chandrakasan [7] proposed an OS-directed power management technique to improve energy efficiency of sensor nodes.DPM is an effective tool which is used to reduce system power consumption without significantly degrading performance.

M. M. Yong Wang and Li-Shiuan Peh [8] had discussed that using supervised learning technique which is used to address some challenges in wireless networks. Using this technique we can provide routing optimizations in planning and deploying real-world sensor networks.

C. P. a. K. J. R. Liu [9] had proposed learning algorithms called Reinforcement Learning (RL) in order to obtain the near optimal policy in point-to-point communication and good transmission strategy in multimode scenario.

R. C. a. V. V. T.Srinivasan [10] had proposed a technique for determining the capability of transfer named packets based on the type of the data that was forwarded in wireless sensor networks with the usage of fuzzy logic architecture at each node in the network.

III. Existing System

The existing system is useful for applications where sensor nodes were deployed in dense manner. When data is transmitted to the sink node this system does not consider the moderated energy consumption.If the network is sparse then packet delivery ratio will be reduced as well as more energy will be consumed. The data conveyed to the sink node will also be decreased.

IV. Proposed System

In this we introduce mobile agents who are mainly used to implement the sparse network in order to improve packet delivery ratio as well as accuracy of data aggregation. This is also used to find out the optimized value of the network parameter using evolutionary computing technique for clustering of sensor nodes.

Information processing in WSNs has three major steps namely pre-processing, data aggregation and inference. [8] Pre-processing is the first step of information processing. It includes simple actions performed on raw data such as signal conditioning (cleaning, compression, scaling and etc.), noise filtering and etc. Data aggregation is the process of aggregating data to the fusion center or inference center in WSN. Inference is a process of using machine learning techniques to extract hidden information out of the aggregated data. Most of current researches focus on applying machine learning algorithms for making inference (step three of information processing in WSNs), such as classifying a moving object in a surveillance WSN based on data gathered by the sensors, abnormal environmental event identification in an environment monitoring.

A. Cluster Evolutionary Algorithm:

Input: population size, expecting similarity threshold, mutation probability

Output: Best fit thresholds

1. Randomly generate similarity thresholds between [0,1] with population size
2. Find the Fitness of each Array in population
3. Select the top most half best fit arrays and clone them
4. Do mutation in all of the arrays to get offspring and find fitness again
5. Select top most half best fit arrays from this offspring

6. Merge previous half and current half population to get the next generation
7. Repeat the process till given fitness value > = excepted similarity threshold

One more enhancement is making the clustering as hierarchical clustering to clearly differentiate the generalization from specialization. Child level cluster is crisper in nature than parent level and exhibits more specific nature and behaviour. To do this, same Adaptive Leader follower algorithm is applied to individual cluster to get multi-level. Here adaptive is more significant term because proposed algorithm can handle concept drifts, conceptual similarity, dynamic building of clusters, and dynamic selection of similarity thresholds.

B. Data Mining in Wireless Sensor Networks:

One of the major objectives of many WSN research works is to improve or optimize the performance of the entire network in terms of energy-conservation and network lifetime. Most of the research activities focus on the design of efficient routing protocol at the network layer selection of low-power modulation scheme at the physical layer or adoption of power-saving modal of operation at data link layer to achieve energy-awareness in WSNs [8].

To illustrate how learning is relevant to decentralized inference and to discuss the challenges that WSNs pose, it will be helpful to have a running example at hand. Suppose that the feature space X models the set of measurements observable by sensors in a wireless network. For example, the components of an element $x \in X = IR3$ may model coordinates in a (planar) environment, and time. Y = IR may represent the space of temperature measurements. A fusion center, or the sensors themselves, may wish to know the temperature at some point in space-time; to reflect that these coordinates and the corresponding temperature are unknown prior to the network's deployment, let us model them with the random variable (X, Y). A joint distribution P X Y may model the spatiotemporal correlation structure of a temperature field. If the field's structure is well understood, i.e., if P X Y can be assumed known a priori, then an estimate may be designed within the standard parametric framework. However, if such prior information is unavailable, an alternative approach is necessary.

C. Model for Data Mining in WSN using Distributed Learning:

Now let us pose a general model for distributed learning that will aid in formulating the problem and categorizing work with in the field. Suppose that in a network of m sensors, sensor i has acquired a set of measurements, i.e., training data, $Si \subset X \times Y$. [1]

Every sensor of the network can read a single value at time and send the data to the Fusion center using network back bone. Later, Distributed learning in WSNs with a fusion center would like to utilize the data which was collected locally to build the overall estimate of the continuously varying field. To achieve this goal divide the network into different clusters and elected the cluster head which is used to collect the data from its members and send the aggregated or summary information to the Fusion Center.

Much of the work in distributed learning differs in the way that the capacity of the links is modeled [1]. The typical assumption is that the topology of these networks is dynamic and perhaps unknown prior to deployment; a fusion center may exist, but the sensors are largely autonomous and may make decisions independently of the fusion center.

V. RESULTS

This paper had surveyed that Data Mining in WSN application from two perspectives, namely the network associated issue and application associated issue. By using the proposed system the simulation results says that the energy consumption is less when compared to existing system. It also says that the packet delivery ratio is more and it also improves the accuracy of data aggregation, performance of overall Wireless Sensor Network.

VI. CONCLUSION

This paper surveys the machine learning techniques applied in WSN from both networking and application perspectives. Data mining techniques have been applied in solving problems such as energy-aware communication, optimal sensor deployment and localization, resource allocation and task scheduling in WSNs. In application domain, data mining methods are mainly used in information processing such as data conditioning, machine inference and etc.

REFERENCES

[1] S. R. K. Joel B. Predd, and H. Vincent Poor, "Distributed Learning in Wireless Sensor Networks - application issues and the problem of distributed inference," IEEE Signal Processing Magazine, 2006.
[2] D. E. D. Culler, and M. Srivastava, "Overview of sensor Networks," IEEE Computer, pp. 41-49, 2004
[3] Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks," presented at ACM MobiCom, Boston, MA, 2000
[4] A. C. W. R. Heinzelman, and H. Balakrishnan,, "Energy-Efficient Communication Protocol for Wireless Microsensor Networks," presented at IEEE Proc. Hawaii Int'l. Conf. Sys. Sci., 2000.
[5] K. Sohrabi et al., "Protocols for Self-Organization of a Wireless Sensor Network," presented at IEEE Pers. Commun., 2000, Oct.
[6] I. E. C. Chien, and C. McCaughey, "Low-Power Direct-Sequence Spread-Spectrum Modem Architecture For Distributed Wireless Sensor Networks," presented at ISLPED Huntington Beach, CA,2001.
[7] A. S. a. A. Chandrakasan, "Dynamic Power Management in Wireless Sensor Networks," presented at IEEE Design Test Comp,2001.
[8] M. M. Yong Wang, and Li-Shiuan Peh, "A Supervised Learning Approach for Routing Optimizations in Wireless Sensor Networks," 2006.
[9] C. P. a. K. J. R. Liu, "Near-Optimal Reinforcement Learning Framework for Energy-Aware Sensor Communications," Communication, 2005.
[10] R. C. a. V. V. T.Srinivasan, "A Fuzzy, Energy-Efficient Scheme for Data Centric Multipath Routing in Wireless Sensor Networks," 2006