Predilection Perspective of Peremptory Evaluation of Wireless Sensor Networks with Machine Learning Approach

Jagadeeswara rao.E[1], Nimmakayala.S.V.Srinivas[2] and Dr.K.V.Ramana,Ph.d[3]

[1] JNTUK,Kakinada, A.P, India.

Abstract

Data mining based information processing in Wireless Sensor Network (WSN) is at its preliminary stage, as compared to traditional machine learning and WSN. Currently researches mainly focus on applying machine learning techniques to solve a particular problem in WSN. Different researchers will have different assumptions, application scenarios and preferences in applying machine learning algorithms. These differences represent 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. 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. This paper will survey Data Mining in WSN application from two perspectives, namely the Network associated issue and Application associated issue. In the Network associated issue, different machine learning algorithms applied in WSNs to enhance network performance will be discussed. In Application associated issue, machine learning methods that have been used for information processing in WSNs will be summarized.

Index terms— Wireless Sensor Network, Machine Learning, Data mining, Fusion Center.

1 Introduction

ireless Sensor Network (WSN) is widely considered as one of the most important technologies for the twenty-first century [1]. 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. 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.

2 II.

3 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 [1]. The design challenges in sensor networks involve the following main aspect: a) Limited Energy Capacity Since sensor nodes are battery powered, they have limited energy capacity. Energy poses a big challenge for network designers in hostile environments, for example, a battlefield, where it is impossible to access the sensors and recharge their batteries. Furthermore, when the energy of a sensor reaches a certain threshold, the sensor

will become faulty and a malfunction can arise, which will have a major impact on the network performance. Thus, routing protocols designed for sensors should be as energy efficient as possible to extend their lifetime, and hence prolong the network lifetime while guaranteeing good performance overall.

4   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 equipped with Global Positioning System (GPS) receivers or use some localization technique to learn about their locations.

5   c) Limited Hardware Resources

In addition to limited energy capacity, sensor nodes have limited processing and also 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 not only consider the energy constraint in sensor nodes, but also the processing and storage capacities of sensor nodes.

6   d) Massive and Random Node Deployment

Sensor node deployment in WSNs is application dependent and can be either manual or random which W 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. If the resultant distribution of nodes is non uniform, optimal clustering becomes necessary to allow connectivity and enable energy efficient network operation.

7   e) Network Characteristics and Unreliable Environment

A sensor network usually operates in a dynamic and unreliable environment. The topology of a network, which is defined by the sensors and 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.

8   f) 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 [2] [3].

9   g) Diverse Sensing Application Requirements

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

10   III.

11   Background

Information processing in WSNs has three major steps [4] namely pre-processing, data aggregation and inference. 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 centre or inference centre 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 to make 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 [5].

IV.

12   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 energyconservation 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 powersaving modal of operation at data link layer to achieve energy-awareness in WSNs.

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 [6]. 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 belongs

to X = IR 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 apriori, then an estimate may be designed within the standard parametric framework. However, if such prior information is unavailable, an alternative approach is necessary. a) 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 =X × Y. For example Si may represent a stationary sensor's measurements of temperature over the course of a day or a mobile sensor's readings at various points in space-time. Suppose further that the sensors form a wireless network, whose topology is specified by a graph. For example, consider the models depicted pictorially in Figure **??**.

13 Figure1: Distribute Learning with Fusion Center

Each node in the graph represents a sensor and its locally observed data; an edge in the graph posits the existence of a wireless link between sensors. Note that the fusion center can be modeled as an additional node in the graph, perhaps with larger capacity links between itself and the sensors, to reflect its larger energy supply and computing power. Apriori, this model makes no assumptions on the topology of the network (e.g., the graph is not necessarily connected); However, the success of distributed learning may in fact depend on such properties.

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 elect the cluster head which is used to collect the data from its members and send the aggregated or summary information to the Fusion Center.

The second approach is in-network processing as shown in the Figure 2.Much of the work in distributed learning differs in a way that the capacity of the links is modeled. 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.

14 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 related to energyaware 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.
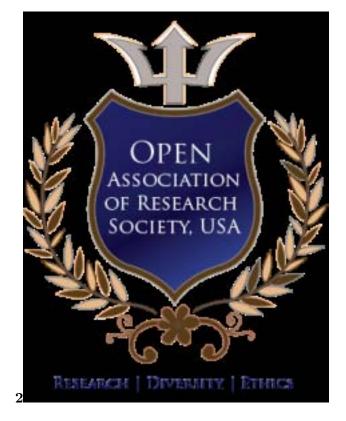
2

Figure 1: Figure 2 :

[126] [Wang and Peh ()] *A Supervised Learning Approach for Routing Optimizations in Wireless Sensor Networks*, MM , Yong Wang , Li-Shiuan Peh . 2006.

[128] [Intanagonwiwat and Estrin ()] *Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks*, R G C Intanagonwiwat , D Estrin . 2000. Boston, MA. (presented at ACM MobiCom)

[130] [Predd and Vincent Poor ()] 'Distributed Learning in Wireless Sensor Networksapplication issues and the problem of distributed inference'. S R K B Predd , H Vincent Poor . *IEEE Signal Processing Magazine* 2006.

[133] [Heinzelman and Balakrishnan] *Energy-Efficient Communication Protocol for*, A C W R Heinzelman , H Balakrishnan .

[135] [Chien and Mcconaghy ()] *Low-Power Direct-Sequence Spread-Spectrum Modem Architecture For Distributed Wireless Sensor Networks*, I E C Chien , C Mcconaghy . 2001. Huntington Beach, CA.

[137] [Culler and Srivastava ()] 'Overview of sensor networks'. D E D Culler , M Srivastava . *IEEE Computer* 2004. p. .

[139] [Wireless Microsensor Networks ()] Wireless Microsensor Networks . *IEEE Proc. Hawaii Int'l. Conf. Sys. Sci*, 2000.