A New Method Based on Machine Learning to Increase Efficiency in Wireless Sensor Networks

Baida'a Abdul Qader Khudor[1], Yousif Abdulwahab Kheerallah[2], Jawad Alkenani[3*]
Email: baidaa.khudur@uobasrah.edu.iq, yousif.abdalwahab@sa-uc.edu.iq, Jawadalkenani@sa-uc.edu.iq
[1]Department of Computer Science, College of Computer Science and Information Technology, University of Basrah, Basrah, IRAQ
[2,3]Department of Computer Science, Shatt Al-Arab University College, Basra, Iraq

Keywords: Data Aggregation, Kalman filter, Support vector machine, wireless sensor networks

Received: September 15, 2022

Wireless sensor networks (WSNs) contain many sensor nodes, and this network is used for many applications such as military, medical, and others. Accurate data aggregation and routing are critical in hostile environments, where sensors' energy consumption must be carefully monitored. There is, nevertheless, a substantial probability of duplicate data due to ambient circumstances and short-distance sensors. Large datasets include a variety of information, some of which is useful, while others are completely superfluous. This redundancy degrades performance in terms of computing cost and redundant transmission. Data aggregation, on the other hand, may eliminate redundant data in a network. In this paper new method called Kalman filter with Support vector machine (KF-SVM) is introduced to classify and data aggregate and get rid of noise in WSNs, which enhances network efficiency and extends its lifetime.

Povzetek: V prispevku je opisana izvirna metoda za agregiranje podatkov v senzorskih omrežjih, ki dela na osnovi Kalmanovega filtra in SVM.

1 Introduction

The creation of low-power sensors and the deployment of large-scale sensor networks are the results of developments in wireless communication and microelectronic devices. With the capacity for pervasive surveillance, sensor networks have received significant interest in a variety of application domains, including habitat monitoring, item tracking, environment monitoring, military, disaster management, and smart environments. In many applications, dependable real-time monitoring is an absolute must[1]. These applications generate an enormous volume of geographically dispersed, dynamic, and heterogeneous data. Data mining can help automated or human-driven tactical/strategic decision-making if this raw data is efficiently evaluated and turned into meaningful knowledge. Therefore, it is crucial to create techniques for mining sensor data for patterns so that intelligent judgments may be made quickly[2].

The fundamental purpose of topology control in WSNs is to guarantee the secure and dependable transfer of data acquired by sensor nodes. ensuring the integrity of the maximum connected graph is the primary strategy for preserving the created network topology. In recent years, the need for WSN application fields has increased steadily. The structure of the topology is produced by a self-organizing network. In many instances, not only perfect communication but also the scalability and universality of the network are necessary. In other words, a topology must accommodate the needs of many sorts of users. This architecture significantly enhances network functionality. The network needs both theoretical operability and real expanding operation, which places increased demands on the topology[3].

The heterogeneous data-transmission network created in this research is a network with a tree topology that exhibits excellent data transmission performance. The tree topology is superior to other network topologies in terms of data transmission and damage resistance. The benefits of the tree structure include efficient data transmission and data aggregation by non-leaf nodes. When designing the minimal tree topology in the references, energy and network latency requirements are thoroughly examined. The advantages and drawbacks of the existing network architecture are measured primarily by the node energy consumption and network throughput of WSNs. There are several network topologies designed for efficient data transmission, including cluster-based topology, tree-based topology, and others[1], [3]. The data mining community has recently paid a significant deal of attention to the extraction of knowledge from sensor data. On-sensor data, many techniques concentrating on clustering, association rules, common patterns, sequential patterns, and classification have proven effective[4]. However, the design and deployment of sensor networks present unique research challenges due to their large scale (up to thousands of sensor nodes), random and risky deployment, loss of communicating environment, limited power supply, and high failure rate. Because traditional data mining techniques are centralized, computationally expensive, and focused on disk-resident transactional data, they are inapplicable[5]. As a result, new algorithms have been