ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

A Survey about Data Prediction in Wireless Sensor Networks with Improved Energy Efficiency.

Thamizhselvi S*, and Prince Mary S.

CSE Department, Sathyabama University, Chennai, India

ABSTRACT

Wireless sensor networks (WSN) infrastructure with data mining is one of the more powerful and research topic in most of the applications. This paper explains the survey on WSN virtualization with data prediction from sensed data of Wireless Sensor networks. It describes technique followed to achieve data prediction and way of organized WSN. This paper addressed all issues and drawbacks occurred in the existing system and how it could be tackled. Survey has been done with many existing papers to understand better about the WSN infrastructure, to implement data prediction aggregated in the base station of WSN with compressed, signature-based and secured data. Survey is also done to understand more about various query processing strategies in data mining of sensor networks.

Keywords: WSN, Aggregation, Compression, Data prediction.

Corresponding author

ISSN: 0975-8585

INTRODUCTION

Wireless sensor network consists of many sensor nodes. These sensor nodes are clustered together based on different categories (based on nature of data, ranges defined for sensors). These sensor nodes are used to sense the environment conditions such as temperature, pressure, humidity, sound etc., with specified periodic time. Each sensor node will sense the environment conditions and send it to the cluster head by means of shortest path which is the only node having direct communication with base station. When there is direct communication between each sensor nodes and base station, then congestion and data traffic occurs which will reduce the performance of network system. Data clustering, data aggregation, data security and compression of data is also important in wireless sensor networks since it improves energy consumption of sensor nodes, reduces communication overheads and to achieve the data integration.

SURVEY OF EXISTING RESEARCH

T. Palpan as [1] reviews Data Analytics in Sensor Networks in real time, effort of the research by describing two important problems in the concept of WSN, (i.e.) real time data collection from sensors and it's processing. Also it describes about data management techniques and network protocols.

D. Tulone and S. Madden [2] proposed an energy-efficient framework, called SAF designed for approx. query and clustering of all nodes in a sensor network. SAF runs with time sequence forecasting models to predict sensor readings. An energy efficient approximate query framework which dramatically reduces the amount of communication in sensor networks.
This scheme works by detecting data similarities among sensor nodes by comparing their local models rather than their raw data.

This definition of similarity, coupled with the linearity of our models, allowed us to derive an efficient algorithm for clustering that is optimal in the number of clusters formed by the network.

M. Hassani[3] describes organizing of energy -based clustering of sensor nodes in wireless sensor networks using related attributes by self. Physical clustering in sensor networks allows sensor nodes to group together according to some similar criterion like neighbourhood. Out of each cluster group one sensor node will be chosen as the group representative and it is responsible for forwarding the data which has been collected. This considerably reduces the consumption of energy. Physical clusters are developed according to similar measurements of sensor nodes. Many Existing data mining approach for physical clustering concentrated on the similarity overall dimensions of measurements. Hassani proposed ECLU, an energy aware method for clustering of sensor nodes physically based on both spatial and measurements similarities.

N. Q. V. Hung [4] describes about the volumes of sensor data accumulation, data compression in a broad- range of sensor data applications. Model based approaches have been highlighted due to their significant compression improved performance. These methods have never been compared and hence compression technique for a particular application seems to be difficult. To address this problem, Hung presents a benchmark that offers a comprehensive study on the comparative performance of the model-based compression techniques. Hung re-implemented several state-of-the-art methods with comparative analysis and various performance factors with benchmark, compression ratio, calculation time, maintenance cost, quality approximation and robustness to noisy data. Hassani provided deep analysis of the benchmark results, obtained using 11 real data set consisting 346 heterogeneous sensor data signals. Findings considered after the benchmark will observed as a real time guideline for sensor applications which are in need to compress sensor data.

N.Q.V.Hung [4] compares representative techniques from each category according to data reduction and prediction accuracy. The study concludes that constant and linear models outperform the others in the presence of small variations in the data.

D. Chu [5] describes the approximation of data with data collection of sensor networks using probabilistic models, a user-specified confidence, but special data characteristics, such as periodic drifts, should been coded explicitly by domain experts.

D. Chu proposed a robust technique called Ken which used replicated dynamic probabilistic models for minimizing interaction and congestion from sensor nodes to the network's base station. In addition to data collection, Ken is well suited for anomaly and event detection applications.

C. Guestrin [6] proposed an efficient framework used for modeling sensor network data using distributed regression parameter approximation technique, nodes collaborate to fit a global function converted to local measurements, but this also requires an assumption about the number of estimation required to fit the data.

FUTURE WORK

In our future work, we proposed to define the cluster the sensors based on sensor range. Three types of clusters defined are homogeneous low cluster, homogeneous high cluster and heterogeneous cluster. We proposed to give high authenticity of each sensing data and integrity of the same in a recoverable environment for concealed Data Aggregation (CDA) by privacy Homomorphism Encryption Scheme using Ecc-Elgamal Signature in a Binary transmission for three completely different Network Clusters. Various query processing such as Top-K Based Query Processing, Necessary Set Based Query Processing, Sufficient Set Based Query Processing and Boundary based Query Processing is also achieved.

CONCLUSION

In this paper, we studied some well-known methodologies applied in WSN infrastructure with data mining. Various framework designs for clustering of the sensors are studied. Studied results are compared and analysis is done in accordance to their efficiencies. This made us to overcome the drawbacks of existing approaches by proposing the new technique to cluster the nodes in sensor networks and to achieve the data security and different query processing.

REFERENCES

[1] Palpanas T., Real-time data analytics in sensor networks in Managing and Mining Sensor Data C. Aggarwal Ed. Springer 2012.
[2] Tulone D, Madden, An energy-efficient querying framework in sensor networks for detecting node similarities in Proc. of the Int. Conf. on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM) 2006.
[3] Hassani M, Mller E, Spaus P, Faqolli A, Palpanas T, Seidl T, Self-organizing energy aware clustering of nodes in sensor networks using relevant attributes on Knowledge Discovery from Sensor Data 2010.
[4] Hung N. Q. V, Jeung H, Aberer K, An evaluation of model based approaches to sensor data compression on Knowledge and Data Engineering IEEE Transactions 2013; 25: 2434–2447.
[5] Chu D, Deshpande A, Hellerstein J.M, Hong W, Approximate data collection in sensor networks using probabilistic models in ICDE 2006.
[6] Guestrin C, Bodik P, Thibaux R, Paskin M, Madden S, Distributed regression: an efficient framework for modeling sensor network data in Proc. of the Int. Symp. on Information Processing in Sensor Networks (IPSN) 2004.
[7] Wu, Lim, Tan, Query-driven data collection and data forwarding in intermittently connected mobile sensor networks in Proc. of Int. Conf. on Data Mgmt. for Sensor Networks(DMSN) 2010.
[8] Gruenwald L, Sadik M.S, Shukla R, Yang. H, DEMS: a data mining based technique to handle missing data in mobile sensor network applications in Proc. of the Int. Conf. on Data Mgmt. for Sensor Networks (DMSN) 2010.
[9] Silberstein A, Filpus G, Munagala K, Yang G, Data-driven processing in sensor networks in Proc. of the Conf. on Innovative Data Systems Research (CIDR) 2007.