Context-Aware Query for High-Voltage Transmission Line Fault Detection using Wireless Sensor Network

Paulo Régis C. de Araújo
Department of Telematics
Instituto Federal do Ceará
Fortaleza, Ceará, Brasil
Email: pregis@ifce.edu.br

Raimir Holanda
Department of Computer Science (PPGIA)
University of Fortaleza
Fortaleza, Ceará, Brasil
Email:raimir@unifor.br

Antonio Wendell de Oliveira Rodrigues
and André Luiz Carneiro de Araújo
and João Paolo C. M. Oliveira
Department of Telematics
Instituto Federal do Ceará
Fortaleza, Ceará, Brasil
Telephone: 55 085 33073607

José de Aguiar Moraes Filho
and Angelo Brayner
Department of Computer Science (PPGIA)
University of Fortaleza
Fortaleza, Ceará, Brasil
Telephone: 55 085 3477.3268

***Abstract*—Due to vandalism or weather conditions, disruption or damage in high voltage transmission line insulators may occur, causing a risk to the electricity supply for many cities. Studies show that the insulators′ states can be known by the value of the leakage current which passes through them and by other variables. In this paper, we present a context-aware query strategy in a Wireless Sensor Network (WSN), installed on the electricity supply system, in order to detect faults in insulators on the high voltage transmission line. The approach presented is based on the strategy of continuous and real-time query of some variables. We have created two new modules, the Data Prediction Module (DatPredMod) and the Data Aggregation Module (DatAggrMod) to support this query strategy.**

I. INTRODUCTION

Due to vandalism or weather conditions, disruption or damage in high voltage transmission line insulators may occur, causing a risk to the electricity supply for many cities. The damage to the insulator generates a high leakage current which is sufficient to open switchgear in the electricity substation. Such a situation interrupts the power supply. One of the ways to investigate the health (state) of insulators is to know the leakage current which passes through them. Besides the physical and structural damage to the insulator, the leakage current can vary with temperature, humidity and the environmental pollution. Other important characteristic of the leakage current is that when it passes through the insulator, it generates a radio frequency signal. As the leakage current varies with environmental conditions, it requires an analysis of the environment context to know whether the increase in leakage current is due to the insulator state or if there are environmental conditions causing the increase.

As we have used the WSN to monitor the insulators, we have to observe the most critical point in WSNs, which is energy consumption [5]. The largest part of energy consumption in a sensor node occurs during data transmission or reception. For this reason, the main goal of most algorithms designed for WSN applications is communication cost reduction in terms of energy consumption [2], [3], [4]. For example, the use of in-network aggregation operators [2], [3], [4] is an efficient strategy to reduce the volume of data transmitted in a given WSN, and consequently, the energy consumption. Even the WSN installed on an electricity transmission line, the possibility of harnessing this energy is by magnetic induction which presents a very poor energy conversion.

Thus, our work addresses two issues. The concept of data aggregation and data compression, whose coordination is realized by DatAggrMod (embedded in the cluster head and base station), is used due to the large amount of information that should be passed to the base station. The other issue is related to the concept of data prediction, whose coordination is realized by DatPredMod (embedded in the base station). This concept is also important to minimize the cost of packet traffic in the WSN and to improve the quality of information provided.

This work is structured in five sections. In section II, the basic concepts of this work are depicted. Section III demonstrates all the components of the approach and its functionality. In section IV, the evaluation stage is presented, with graphics that demonstrate the efficiency of our approach. And, finally, section V presents the conclusion and future works.

September 11, 2014

II. BASIC CONCEPTS

In this model, a WSN is composed of three sensor node types: (i) end node, which is responsible for sensing and transmitting sensed data to a cluster head. Each insulator chain has only one end node; (ii) cluster head node, which compresses and aggregates sensed data, forwards the (sub)query for the end nodes, and returns the query results to the base station (sink node). Each transmission tower has only one cluster head node, and; (iii) sink node (base station), which contains the Query Processor, Data Aggregation Module (DatAggrMod) and the Data Prediction Module (DatPredMod). The sink node is installed in the electricity substation. Figure 1 illustrates the topology model depicting the modules that compose the sink node.

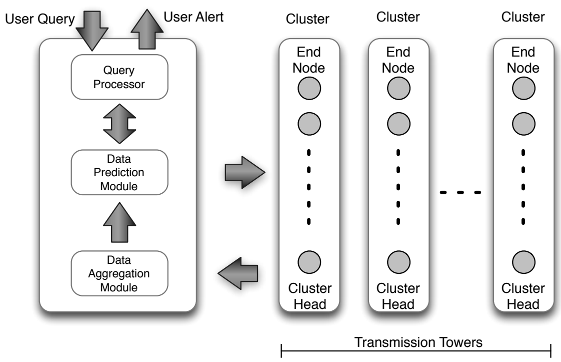


Fig. 1. Sink Node Architecture and Clusters.

In our work, an application in the sink node submits, automatically and periodically, queries to the query processor, which is the interface between the application and the DatPredMod module. Tasks such as data prediction are performed by DatPredMod. Then the query is passed to the DatPredMod which is responsible for selecting a model, converting it to query parameters and, in turn, creating sub-queries. The sub-queries are forwarded to the cluster heads.

When the sub-query arrives at the cluster head, it is forwarded to the end nodes which answer the sub-query, sending back the sensed data. Then the cluster head waits for the sensed data of the end nodes of its cluster, and then aggregates and compresses these sensed data. The aggregated and compressed sensed data are sent back directly to the sink node, if it is closer, or are passed from cluster head to cluster head until the sensed data arrive at the sink node.

When the packets (aggregated and compressed sensed data) arrive at the sink node, a data reverse aggregation mechanism, which is represented by the Data Aggregation Module, applies a reverse aggregation and compression operation to convert incoming aggregated and compressed packets into a single packets as the output. The data aggregation and compression steps aim to reduce network energy consumption, and to increase WSN lifetime.

III. DATA AGGREGATION, COMPRESSION AND PREDICTION IN CONTEXT-AWARE QUERY

In this section, the DatPredMod and DatAggrMod modules are described, as well as their strategies to aggregate, compress and predict data.

A. DatAggrMod - Data Aggregation Module

The Data Aggregation Module is implemented in each cluster head and in the sink node. In the cluster head, its goal is to aggregate and compress sensed data received from the end nodes and from the predecessor cluster head. In the sink node, it aims to realize a data aggregation and compression reverse operation.

To explain the strategy, we suppose a sub-query sent by the cluster head to the end nodes of its cluster. The language used to produce such sub-query is SNQL [2]. To answer the sub-query, the end nodes return the query results to the cluster head. The Data Aggregation Module, in the cluster head, produces an average value of the received measurements. The equation 1 represents the average value of the sensed variables.

$$Aq_{sd} = \frac{\sum_{j=1}^{6} \frac{\sum_{n=1}^{nm} sd(n,j)}{nm}}{6} \qquad (1)$$

In equation 1, the Aq_{sd} is the average value of certain amount of variable measurements (nm) in the time interval of 1 minute for the six insulator chains. In this way, the average value equation of the leakage current is produced by using index lc instead of index sd. The same analysis is used for the average value of the temperature, humidity, and noise signal variables.

The aggregation strategy is also used when incoming packets are arriving at a cluster head. The temperature and humidity variables have a tightly spatial and temporal correlation. These variables can be aggregated based on their similarity. As an example, if Cluster Head 2 sends to Cluster Head 1 its non aggregated packet. Cluster Head 1 extracts the temperature and humidity fields from the packet and compares them with its temperature and humidity average values. If the difference between the values is less than the factor De (De_t for temperature and De_h for humidity), the local value is suppressed and an aggregated packet is created. Based on the database collected in [1], there is no substantial variation of leakage current when temperature variation is less than 1^oC. Then the temperature factor is $De_t = 1$. The same approach is used for humidity whose factor is $De_h = 5$.

After the action of data aggregation in the cluster head, the data compression is implemented. In the next subsection, this strategy is presented.

1) Data Compression in the DatAggrMod: After sensed data aggregation takes place, a data compression strategy is further used to reduce the sensed data packet size, which will be sent back to the sink node. In our experiments, based on [1], we have observed that the leakage current variation sensibility due to temperature variation and humidity variation presented important values, with temperature variation between 20^oC and 50^oC and humidity variation between 50% and 89%. In this way, we can represent the measurement's average value with less number of bits, instead of using one byte to represent each value. With these temperature and humidity variations, we can use two bits to represent the dozen (ex. ten - **2**5) of the average values of the temperature and humidity, and four bits to represent the unity (ex. unity - 2**5**) of the average values of the temperature and humidity. Then we can use only six bits instead of eight bits for each variable. The noise signal ns, which is related to the leakage current, presented a variation between 0 dB and 15 dB, and can be represented by four bits. And, finally, the leakage current has a variation between 1560μA and 2960μA, with the variation of its dozen (ten) representing the sensibility due to other variables. So we have 140 (296 - 156) units to represent the leakage current variation, which can use 8 bits to report it.

It is important to observe that the initial packet of six bytes used to represent the four variables and the cluster identifier was compressed to four bytes. This strategy represents a reduction of 33.3% in the packet size, and, consequently, a reduction in the network energy consumption. After the compression action, the aggregated and compressed packet is forwarded to the next cluster head or to the sink node, if it is closer.

B. DatPredMod - Data Prediction Module

When aggregated and compressed data arrive at the sink node, the DatAggrMod receives these data and realizes a reverse operation of data aggregation and compression. In this way, incoming aggregated and compressed packets are converted to single packets. Then they are forwarded to the DatPredMod.

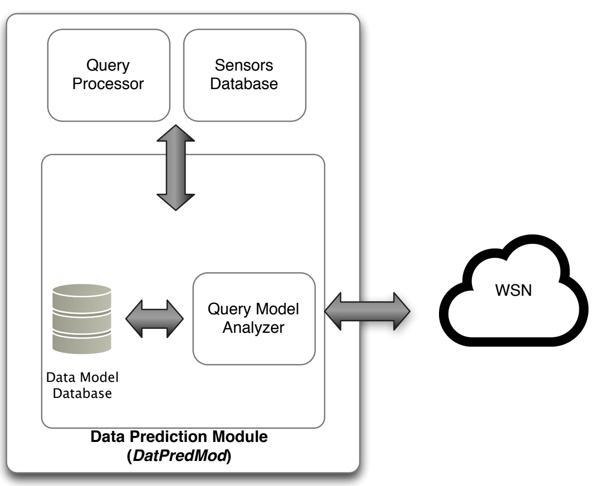


Fig. 2. Data Prediction Module.

As shown in Figure 2, the DatPredMod is constituted by a Query Model Analyzer, which is responsible to: (i) receive the query of the Query Processor; (ii) make sub-queries to the WSN based on the predefined models; and (iii) receive sensed data of the DatAggrMod to build a linear regression-based sensor-field model and to store it in the sensors database. The DatPredMod has a Data Model Database, which is responsible for storing predefined data models. These data models are based on environmental and electrical systems studies realized at the Electrical Engineering Department of UFCG (Federal University of Campina Grande) [1].

In the proposed prediction strategy, the linear regression model is used to select the set of parameters into the Data Model Database. With the prediction equation, equipped with the last ten queries, the Data Model Database is searched to produce the parameters to the next sub-query.

The in-network prediction implemented by Query Model Analyzer is based on the following equations:

$$lc'(t) = a_t + b_t t \tag{2}$$

$$lc'(h) = a_h + b_h h \tag{3}$$

$$ns'(lc) = a_{lc} + b_{lc} lc \tag{4}$$

In equation (2), the temperature t is an independent variable. $lc'(t)$ represents the estimated value of $lc(t)$ (leakage current) and it varies with t (temperature). The same explanation is used for equations (3) and (4). The parameter a_t is the interceptor-t (value of $lc'(t)$ for temperature $t = 25^0C$) and b_t is the stretch slope, and they are computed as follows in equations (5) and (6).

$$a = \frac{1}{N}\left(\sum lc_i - b\sum t_i\right), \tag{5}$$

$$b = \frac{\sum(t_i - \bar{t})(lc_i - \bar{lc})}{\sum(t_i - \bar{t})^2} \tag{6}$$

IV. EVALUATION

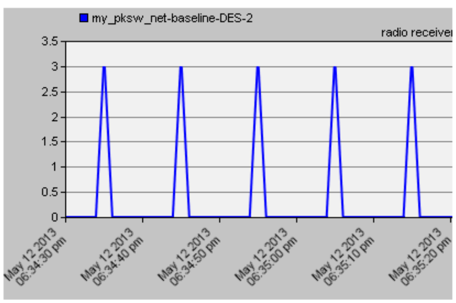
We have used the OPNET Modeler 14.5 as a simulator to validate this work. Our experiments are based on a Campus type scenario. We have analyzed two scenarios as depicted below.

i. A scenario with 24 sensor nodes (installed on the transmission towers) and 1 sink node (installed into the electricity substation). It is equivalent to 8 transmission towers (1 cluster head and 2 end nodes in each tower). In this scenario, we have fixed 25°C for the temperature and 70% for the humidity for clusters 2, 3 and 4; 27°C for the temperature and 60% for the humidity for clusters 5 and 6; and 29°C for the temperature and 50% for the humidity for clusters 7 and 8.
ii. A scenario with 99 sensor nodes and 1 sink node. It is equivalent to 33 transmission towers. In this scenario, we have fixed 25°C for the temperature and 70% for the humidity for clusters of 2 to 25; 27°C for the temperature and 60% for the humidity for clusters of 26 to 60; and

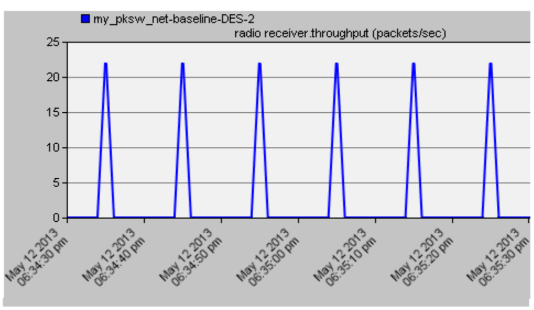
29°C for the temperature and 50% for the humidity for clusters of 61 to 99.

For each scenario, we have generated ten queries (every 10 seconds) for the temperature and the humidity for all sensor nodes in the WSN. Two situations have been implemented for each scenario: in one situation, we have used a data aggregation technique to reduce the number of packets traveling on the network, and in the other situation, we have not used the data aggregation technique.

To validate the Data Prediction Module, we have generated an anomaly in the last cluster of the WSN, cluster 8. We have fixed 2 mA for leakage current, 25°C for temperature and 60% for humidity. This anomaly, which needs a more specific query, is a typical case of flashover risk (insulation breakdown) at the transmission line of the electrical system, and it was represented by a data model in the module DatPredMod.



3.a



3.b

Fig. 3. Number of packets received in the cluster head 1 due to query results

Figure 3 shows the data flow in the radio receiver of Cluster Head 1 for two situations: Figure 3.a represents scenario (ii) (with data aggregation technique) and Figure 3.b represents scenario (i) (without data aggregation technique). This cluster head was chosen due to its closeness to the sink node, and relays all queries and results to the WSN. So it is the node with the largest number of retransmissions of the network, and with the greatest criticality as to the energy consumption. In Figure 3.a, we have noted that, due to scenario (ii), with 3 groups of parameters (temperature and humidity) and using data aggregation technique, only 3 data packets have been received by the radio receiver of this node as a result of each query in the network.

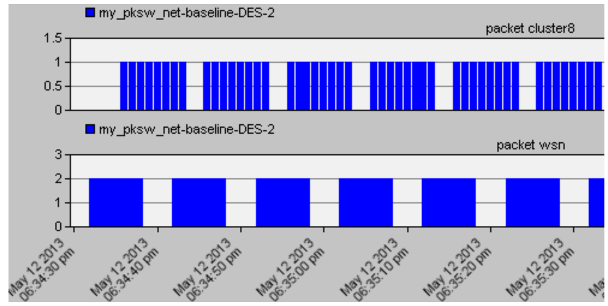


Fig. 4. Results of queries with anomaly in the cluster 8

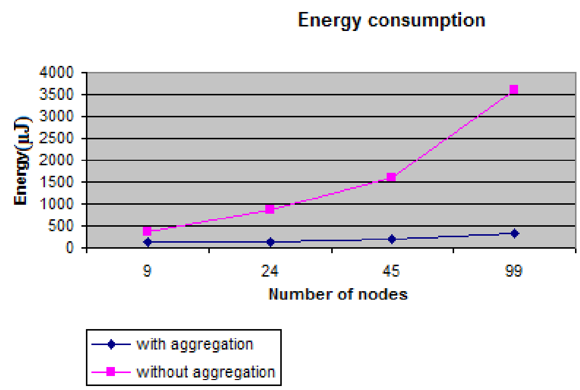


Fig. 5. Energy expended by the receiver of the cluster head 1

Figure 3.b represents the data flow in the radio receiver of Cluster Head 1, but now the scenario (i) is addressed, with 3 groups of parameters (temperature and humidity) configured for all nodes, as shown in scenario (i). In this configuration, we have disabled the data aggregation module for all cluster heads. We have observed that 22 data packets have been received for each query in the WSN. In this case, there was packet loss, because 23 data packets were expected. When we have simulated with data aggregation technique enabled for this scenario, we have observed a reception of only 3 data packets. For this scenario, a comparison between the application of the data aggregation technique (reception of 3 packets) and its absence (reception of 22 packets) shows a reduction of 86.36% in the energy consumption of the CH1 (Cluster Head 1) radio receiver.

In Figure 4, we have generated a query to the WSN every 10 seconds, and we have observed a specific query to cluster 8 every 1 second. This situation has occurred due to the Data Prediction Module (DatPredMod), in the sink node, in that it has compared the linear regression model of the first ten query results (with anomalous result for cluster 8) with one of the data models. As the anomalous result was a case of risk, the DatPredMod has initiated a specific query to the cluster that generated the anomalous result. This simulation has been implemented to prove the efficiency of the data prediction.

Figure 5 presents a graph that relates the energy consumption of the CH1 radio receiver with the number of network sensor nodes. We have used, for energy consumption of the radio receiver, the value of 0.5μJ/bit [6] and 10 Bytes for the packet size. The blue and red lines represent "respectively" the energy consumption of the CH1 radio receiver for scenarios (i) and (ii)(24 and 99 nodes) "with" and "without" data aggregation technique enabled in all cluster heads. We have also used scenarios with 9 and 45 nodes to validate this test. The best result of the simulations was achieved with scenario (ii), with the reduction of energy consumption of 3,600μJ (without aggregation) to 320μJ (with aggregation), representing a reduction greater than 10 times.

V. CONCLUSION

In this work, we have presented data prediction, aggregation and compression mechanisms to provide queries on a WSN with energy efficiency and efficient fault detection. The approach aims at efficiently detecting transmission lines faults using a WSN with a long lifetime. The data prediction mechanism is embedded in the sink node which has a large energy supply and processing capability. That is, this mechanism does not consume the network energy, and the data aggregation and compression mechanism allows a gain in the reduction of the network energy consumption.

The simulation experiments have demonstrated that the combination of data prediction, aggregation and compression mechanisms is effective, and increases significantly the WSN lifetime.

As future work, we should implement context-aware query, based on event prediction, using Artificial Intelligence or Pattern Recognition techniques. We also intend to implement a voting (election) of cluster head based on communication reliability due to radio frequency and corona interference in the transmission line.

ACKNOWLEDGMENT

The project that resulted in this paper is supported by Energy Company of Pará (CELPA - Pará, Brasil) and National Agency of Electric Energy (ANEEL - Brasil).

REFERENCES

[1] P. R. Araujo and R. Ramos. Indoor tests of insulator chain electric characteristics using high voltage - p&d chesf. Campina Grande, PB, BRA, 2003. CHESF.

[2] A. Brayner, A. Lopes, D. Meira, R. Vasconcelos, and R. Menezes. Toward adaptive query processing in wireless sensor networks. *Signal Processing*, 87(12):2911 – 2933, 2007. Special Section: Information Processing and Data Management in Wireless Sensor Networks.

[3] A. Brayner, A. Lopes, D. Meira, R. Vasconcelos, and R. Menezes. An adaptive in-network aggregation operator for query processing in wireless sensor networks. *J. Syst. Softw.*, 81(3):328–342, Mar. 2008.

[4] A. Deligiannakis, Y. Kotidis, and N. Roussopoulos. Processing approximate aggregate queries in wireless sensor networks. *Inf. Syst.*, 31(8):770–792, Dec. 2006.

[5] I. Dietrich and F. Dressler. On the lifetime of wireless sensor networks. *ACM Trans. Sen. Netw.*, 5(1):5:1–5:39, Feb. 2009.

[6] A. W. S. H. D. E. C. J. Hill, R. Szewczyk and K. S. J. Pister. System architecture directions for networked sensors. In *Proceedings of the 9th International Conference on Architectural Support for Programming Languages and Operating Systems*, pages 93–104, 2000.