A Data Prediction in Wireless Sensor Networks using Deep Learning-based RSA Algorithm

Anand Dohare, Tulika, Sweta Sachan, B. Mallikarjuna

Abstract. In wireless sensor networks (WSN) data collection and gathering data from surroundings and removing the redundancy, process the appropriate data is a challenging task, in spite of low battery, less memory space, low computational speed, reduce the energy consumption are major research areas in WSN. Temperature sensor, humidity sensors are majorly used to climate monitoring, agricultural, humidity observation, due to energy consumption of sensors, low battery power, computational speed of sensors are used to maintain a long time is a crucial issue. To overcome such type of problems data prediction techniques are required, several data prediction, aggregation techniques are proposed in this issue and several research has been done, but not solved all challenging issues. In this paper proposed deep learning-based RSA algorithm to provide security and efficiently handle the data by using a feed-forward filter to remove the aggregated data, Least Mean Square (LMS) variable step-size method to remove error rate that will improve the energy consumption and size of the memory space for data collection, the experimental results proved that 98% predicted data and minimum error rate on cluster network as per considered (Intel Lab) data set.

Keywords: wireless sensor networks, prediction, aggregation, redundant, LMS

I. INTRODUCTION

In the present-day scenario, sensors are used in traffic monitoring, healthcare, burglar alarm and various domestic appliances for safety purposes [1]. Sensor is the most wanted device for accessing the data and process to the corresponding application, the most common device used in the military for border security enforcement, in irrigation for crop purpose to rectify an attack monitoring, in health care applications, to check heart pulse and heart beep reading. The temperature and humidity sensors are most useful sensors in many applications, earthquake, to check the vibrations and measure the sound and estimate the pollution in the air, air pressure and vibrations in the air, etc... The following Table 1 provides the challenging issues and major areas of research used in WSN [2].

| Wireless Sensor Networks (WSN) | Challenging Issues on WSN | Major Area used WSN |
|-------------------------------|--------------------------|----------------------|
| Energy consumption            | Health care Applications |
| Energy conservation           | Military Applications    |
| Hardware and Software related | Transportation system    |
| Synchronization              | Agricultural Applications |
| Quality of Service (QoS)      | Irrigations and crop     |
| Data Computation              | Environmental monitoring |
| Security                      | Animal habitats          |
| Data collection               | Forest fires             |
| Minimum communication         | Natural disasters        |
| High sensing                  | Inventory Management     |
| Flexibility                   | Smart home               |

Combination of two or more trees it becomes the cluster tree network architecture, it combinations of several sink nodes, the root node can act as head node[1], all remaining nodes can gather the data from the respective head node and transmit the data to the base station, it is responsible for preserving the data for a long time. Energy consumption, improve the performance of battery life time was a challenging task for researchers, sometimes huge amount of data stored in the WSN database and occupy huge memory space and increase transmission rate [2]. Data prediction can be done three stages i) time series forecasting ii) Algorithm approaches iii) Stochastic approaches. In the time series forecasting used filters and calculates the error and adjusts the weights and it provides acceptable accurate and it is easy to implement. Algorithmic approach works with the as per the algorithm defined. In stochastic approach works with the probability density function, it provides the large computational speed. Aggregate prediction techniques are used to transfer the data based probabilistic values, aggressive prediction techniques are easy to implement. The data compression technique used to reduce the data, in this technique the data can be moved to the sink node by applying code strategy and semantic information. Network aggregation strategy, in this process data transfer between the source nodes to sink node this approach is depend upon the minimum and maximum average result of data [3]. Deep learning-based models [23-25] can be implemented in mobile ad hoc networks [18], sentimental analysis for recommender system on the cloud [21-22], deep learning models consist of the feed forward network [19-20] model and it consist of the recursive approach as shown in the Figure 1, feed ford network consists of the one input layer, one output layer many hidden layers.
A Data Prediction in Wireless Sensor Networks using Deep Learning-based RSA Algorithm

Deep learning-based models can be implemented in Facebook and obtain the effective results of the performance of data prediction 97% while transferring the pictures, messages, etc. [21]. Google used deep learning models on its data centers and minimize energy consumption upto 40% [22].

Figure 1: Deep Learning Models with Feed Forward Network [23]

In this paper, we proposed a deep learning-based RSA algorithm [18-25] with feed-forward filter computes the error rate (it named as feed forward filter it works with the same principal of Kalman filter [15]). LMS based variable step-size [13] method can compute the data rate in wireless sensor networks, which helps to reduce the energy consumption and increase battery performance as well as memory size by decreasing the transmission rate. Feed-forward [19] filter find the error rate and deep learning-based RSA method is used for reducing the redundant data and provide the security at source as well as a sink node. The proposed approach contains two techniques used in data prediction first one is feed-forward filter and the second one is LMS( Least mean square) based variable step size technique, in feed forward filter based prediction algorithm, adjust coefficient of predicted value and compared with threshold value if the prediction error is greater than the threshold error then predicted data will be removed if prediction error is less then threshold error than data will be considerable. While LMS used initial prediction at the actual value and based on the current time data error for reduced the mean square error. It compared step size parameter with previous values by using a recurrence relation. The best things regarding LMS algorithms require previous knowledge of collected data. In this paper we used LMS algorithms for data reduction with variable step size and apply at three different network topology [4] like star, tree, and cluster-based on real-time Intel data set [17] as shown in the Figure 2, it shows the three different architectures.

Figure 2: Tree, Star and Cluster topologies [4]

II. RELATED WORK

Temperature and humidity sensors best devices from collecting the data and monitoring the pollution, temperature, levels of humidity, sensors are used in most sensitive places like health care and military enforcement, border security [5]. Energy consumption [5], battery life is the most challenging issue for researchers in WSN [5]. Data prediction means to reduce the redundancy data and remove the useless data [6-13], huge amount of research has been done on data prediction but no literature available on a deep learning-based RSA algorithm. The following Table 2 shows the major research gaps are identified by the prediction methods are used by the WSN.

Table 2: The major research gaps identified in data prediction in WSN

| Author                | Research Gaps                                                                 |
|-----------------------|-------------------------------------------------------------------------------|
| Ramsey Faragher       | In this study provides the basic principle of working ‘Kalman filter’, error occurrence has been done after adjustment of waits, energy consumption is not discussed. |
| Guivi Wei             | Combining Gray model based data aggregation (GMDA) in Kalman filter but it improves the communication only.                        |
| Olston                | Proposed TRAPP (Tradeoff in replication precision and performance) query based optimization approach but it is not suitable for all challenges in WSN. |
| Rajagopalan           | This study gives a survey on various data prediction techniques on wireless sensor networks                                    |
| Anastasi              | The study techniques receive data redundancy technique, but not solve Energy consumption, memory space, etc.                   |
| Deshpande             | This study gives classifies the data aggregation approach in three ways stochastic approaches, algorithmic and time series forecasting approaches, it solves energy consumption only. |
| Chu                   | Data collection in sensor networks solved by using probabilistic models, but it reduces the anomalous data, it not solved by the energy consumption, memory space etc. |
| Stojkoska             | Data prediction in wireless sensor networks solved by using variable step size LMS algorithm, but not used updated descriptions. |

As per Table 2, data prediction is a technique to remove the useless data and reduce the redundant data, the various research has been done on data prediction, data compression in WSN, Kalman Filters [6] was proposed prediction based data models are used in wireless networks, Guiyi Wei[7] proposed a novel approach for data prediction by the combination of (GMDA)with Kalman-Filter named as Kalman filter data aggregation (KFDA), it improves accuracy, but suffering from energy consumption. Olston [8] proposed TRAPP based data reduction used query-based optimization and deliberated different models of the sensor node, it performs data transmission between synchronous intervals, it is mainly focused on energy savings but not solve the major challenging issues. Rajagopalan [9] survey on various data normal operational techniques and stochastic models to reduce the communication overhead, remove data redundancy and increase the battery lifetime the data need not be reported on every sink node where the sink node does not give the information to the root node, this model fails on aggregated data in this model lack communication and also cannot improve memory space.
Despande [11] the probabilistic models only used to solve this approach, this is model not suitable for improvement of energy, and lifetime of the battery. Chu [12] this prediction model useful to reduce the...

III. PROPOSED APPROACH

Sensor is a small low battery power device which has the ability to collecting information, gathering data and transmitting to the base station, due to low battery power and the low computational speed we cannot utilize all features of a sensor node to overcome its drawback [1]. The level of data aggregation decides by the deep learning approach [23], it improves the depth of the network by using deep learning approach [24]. The more number of layers are used while RSA encryption and decryption algorithm provides end-to-end confidentiality for collecting and aggregating the data. The entire data collection it performs the individual data operation as follows [25].

\[
\begin{align*}
\sum E(D_1+ D_2) &= \sum E(D_1) + \sum E(D_2) \\
\sum E(D_1 \times D_2) &= \sum E(D_1) \times \sum E(D_2)
\end{align*}
\]

It performs addition and multiplication of encrypt the data without the need of individual decrypted data, the pseudo code for key generation as shown in Algorithm 1.

Algorithm 1: Pseudo code for Key generation procedure for RSA algorithm:

Step 1: Consider two different random prime numbers P1 and P2 where P1 ≠ P2
Step 2: Calculate the Value M=P1×P2. N value will be calculate
Step 3: operation of both P1 and P2
Step 4: Find an integers values 1<i<(M), and calculate i, (M), no divisors until
Step 5: Finally compute the value of private key based on
Step 6: Repeat step 1 to 5 until private key is obtained

After generating the key it is useful to find encryption and decryption as follows, Public key is obtainable for entire cluster data, it can be represented as (\(\delta, M\)). Private key is existing to moreover sink nodes the, the pseudo code for encryption and decryption as shown in Algorithms 2 & 3.

Algorithm 2: Pseudo code for RSA Encryption Algorithm

Step 1: Sensor node contains public key Pu(i,M)
Step 2: En=(Ni) mod (M)
Step 3: The data is encrypted

Algorithm 2: Pseudo code for Decryption Algorithm

Step 1: sink nodes or base stations have private key such as Pr (Pr1,M)
Step 2: N= (Enp) mod (M)
Step 3: The data decrypted.

Kalman Filter [15] is ancient approach (1960), later the model is updated with Gray model and Dual Kalman filter etc, while data processing, Neel Arm Strong used in to predict earth surface, moon surface, satellite navigation system[16] LMS method is also proposed Widrow-Hoff rule (1960) used in adaptive circuit switching method for high data transmission in AT&T-bell laboratories, in this proposed methodology the same LMS algorithms has implemented with the feed-forward filter as shown in Figure 3.
Each node can process the data to the next node, each and every sensor node have equal time period, the prediction technique is applicable to each and every node in the network. To compute the energy consumption in wireless sensor networks $E_c$, it receive the data $D$ bits per second per unit of time, the energy consumed to transmit the data $T$ bits per second per unit of time, $E_{fs}$, the amount of energy required for feed forward filter, $E_{fs}$ is the relay activation energy then it follows

$$E_c = D \times T \times E_{fs}$$

The distance between one sensor node to another sensor node $d_i$ then it calculate the distance function as follows.

$$d_i = \sqrt{\frac{E_{fs}}{E_{elc}}}$$

The LMS algorithms works with the three major techniques such as i) Filter output ii) Estimation error iii) weight or input adaption as shown in Algorithm 4.

Algorithm 4: Least Mean Square (LMS) Algorithm

Step 1: Compute the filter output where $F[n]$ be the output signal, $a'[n]$ be the adjective weight and input coefficient, the current input signal $\mu[n]$ used as follows.

$$F[n] = a'[n] \mu[n].$$

Step 2: Then it calculated the estimated error, where $e[n]$ is the estimated error occurrence has been given to the feed forward filter, $C[n]$ is the expected output signal that subtracted from the given input signal $F[n]$ as follows.

$$e[n] = C[n] - F[n].$$

Step 3: Adjust the weights as per the input adaption from equattion 2, $\mu$ is the step size coefficient which can help to calculate the filter length, step size coefficient, it can adjust the weights as follows.

$$a[n + 1] = a[n] + \mu e[n].$$

The feed forward filter length $M$ where $\mu$ is the input signal of feed forward filter, $u[n]$ is the current input signal $d[n]$ be the desired out signal, $n$ be the iterative signals to adding the iterative data, the feedback can be obtain the adjustment of the desired input signal, it give the error it can adjust signal then obtain accurate feedback of feed forward filter, the auto correlation $\eta$ can be obtained from $n$ number of layers, the predicted signal obtained from desired output signal which gives error as feedback to adjust input signal. The total input signal $IP_{total}$ is devisable $n$ layers matrix that is obtained by learning rate coefficient.

$$\eta = \frac{n}{IP_{total}}$$

The total input power $IP_{total}$ obtained by the largest eigenvalue of autocorrelation matrix $\eta$. The step size variable $\mu$ lies in between the 0 and autocorrelation matrix .

$$0 < \mu < \frac{n}{IP_{total}}$$

Where $IP_{total}$ it can be compute through mean value as follows

$$IP_{total} = \sum_{n=1}^{n} \frac{\mu_n}{n}$$

In the WSN the data collection by using deep learning-based RSA at the application layer, first sensor node collects data from the physical layer and sends data to the sink node, the sink node sends data to the sensor node, the sensor node checks the redundant data and remove the useless data and it sends to the data center [20] and it reduces the memory space as well as increase the data transmission rate, and it improves battery life. The overall proposed methodology has been described in Figure 4.

![Image](image_url)

In Below Figure 4, it performs three different stages, first stage decides the topology tree, star and cluster based widely used in WSN, while data transmission stage feed-forward filter used to collect data for each topology and compare the value of step size variable $\mu$ with the threshold value, it contains variable length $M$ and calculate the error rate and it reduced to $n$ number of hidden layers. The third stage LMS value used to collect the data from each sensor and transmit to respective sink node at that time same algorithm is synchronously work and effective transmission of data to
the sensor nodes, the aforementioned approach continually work until whole data transmission in the entire network.

IV. EXPERIMENTAL EVALUATION

The proposed architecture, data collection using deep learning RSA prediction is implemented ns2 tool kit, the iterate approach after obtaining the LMS value with variable step size given to the network for transmitting to the data than given to the network such as tree, star and cluster. The feed-forward filter provides the prediction value μ, the IntelBerkeley Research Lab [17] provides the data set as follows.

Table 3: Intel Berkeley Research Lab resultant data set [17]

| Topology | Sink node | Child node | Filter length (M) | Step Size |
|----------|-----------|------------|-------------------|-----------|
| Tree     | 19        | 20, 21, 22, 23, 26, 27 | 6                | 4         |
| Star     | 4         | 2, 3, 5, 6, 7 | 6                | 5         |
| Cluster  | 14, 19    | 17, 18, 20, 21, 12, 13, 15, 16 | 2.4              | 6         |

In tree topology where sink node 19 the appropriate child nodes are 20, 21, 22, 23, 26, the feed forward filter length value M=6 (1.2 × 10^5) and step size 4, it provides the data prediction rate 95%. In star topology the sink node becomes 4, the intermediate nodes are 2, 3, 5, 6, 7 the M value 6, the step size and filter length are same as tree topology. It considered sink nodes are 14 and 19 in cluster topology, the intermediate nodes are 17, 18 and 19, the filter length M = 2.4 (1.2×10^8). The aforementioned data set values with mathematical models are simulated with Matlab tool kit evaluated the results, it obtains the error rate results on the humidity sensor and temperature sensor, the effective results are carried out by the cluster topology compared to a tree and star topology, the highest error rate is measured on tree topology because it contains the maximum number of sink nodes, the following Figure 5 for error rate is measured on temperature sensor and Figure 6 error rate is measured on humidity sensor.
A Data Prediction in Wireless Sensor Networks using Deep Learning-based RSA Algorithm

Figure 8: Performance of Battery life time.

To evaluate the data prediction approaches by using LMS algorithms for variable step-size method in networks such as tree, star, and cluster, the MATLAB simulator evaluated the results, the average data rate of humidity and temperature sensor assumed as per topology as a tree, star and cluster on as per the data set [17] as shown in the Table 3.

Table 3: Obtained average data rate as per the topology

| S.No | Topology | No. of nodes for topology | Average temperature | Average humidity | Average data transferring rate | Step size |
|------|----------|---------------------------|---------------------|------------------|------------------------------|-----------|
| 1    | Star     | 2.3, 4.5, 6              | 18.95               | 39.97            | 43.24                        | 4         |
| 2    | Tree     | 2.02, 22, 3.6, 27        | 19.08               | 39.05            | 41.18                        | 5         |
| 3    | Cluster  | 17.18, 20.2, 1.15, 16, 17, 19 | 18.76             | 39.82            | 42.94                        | 6         |

Average data rate of humidity and temperature sensors, cluster network provides the higher performance rate compared to tree and star topology that it gives approximately 98% of data reduction and transmitting rate because data transmit between a minimum numbers of sink nodes as shown in Table 4.

Table 4: Average data rate as per topology

| Topology | Sink node | Child node | Filter length (M) | Step size | Data prediction |
|----------|-----------|------------|------------------|-----------|-----------------|
| Tree     | 19        | 20.21, 22, 32, 26, 27 | 6               | 7         | 95%             |
| Star     | 4         | 2.3, 5.6   | 4                | 4         | 92%             |
| Cluster  | 14, 1     | 17, 18, 20.2, 12, 13, 15, 19 | 4 | 4 | 98% |

V. CONCLUSION

The experimental evaluation as per the Intel lab measured temperature humidity sensors, the valuations of results by using the proposed architecture by using the deep learning RSA prediction algorithm using the LMS method and variable step size approach implemented in ns2 tool kit, it obtains the result as implanted in MATLAB R2016b shown in Table 5.

Table 5: Star, Tree, Cluster corresponding type of sensor

| Type of Sensor | Amount of Transfer (MBPS) | Prediction data rate |
|---------------|---------------------------|----------------------|
| Temperature   | 18.59                     | 94.9%                |
| Humidity      | 38.64                     | 95.01%               |
| Temperature   | 43.24                     | 92.01%               |
| Humidity      | 44.12                     | 91.96%               |

It provides the 95% prediction rate in start topology on step size 4, the data transfer rated 18.59 MBPS on the temperature sensor and 95 MBPS in humidity sensor. In the tree, topology obtains 92%, and cluster topology obtain 98% higher prediction data rate when compared to the tree and star topologies. Tree and star topology needed less prediction rate due to the various sink and intermediate nodes and moreover it contains parent nodes, but the cluster network shows better performance when compared to tree and star topology.

REFERENCES

1. Akylidiz, T. Melodia, & K.R. Chowdhury. (2007). A Survey on wireless multimedia sensor networks. Computer Networks (Elsevier), 51, 921–960.
2. Mohammad Salahuddin, & Al-Fuqaha, A. (Eds.). (2015). Introduction to wireless sensor networks (pp. 3–32). New York: Springer.
3. Santini, S., & Römer, K. (2006). An adaptive strategy for quality-based data reduction in wireless sensor networks. In Proceedings of the 3rd International Conference on Networked Sensing Systems (INSS 2006) (pp. 29–36). June 2006. Chicago, IL, USA.
4. Park, J., & Moon, K. (2012). Optimal stop points for data gathering in sensor networks with mobile sinks. Wireless Sensor Network.
5. Slizielman, W. R., Chandrakasan, A., & Balakrishnan, H. Energy-efficient communication protocol for wireless micro-sensor networks. In Proceedings of the IEEE Transactions on 33rd Hawaii International Conference on System Sciences.
6. Faragher, R. (2012). Understanding the basis of the Kalman filter via a simple and intuitive derivation. IEEE Signal Processing Magazine.
7. Weia, G., & Ling, Y. (2011). Prediction-based data aggregation in wireless sensor networks: Combining grey model and Kalman filter. Computer Communications, 34, 793–802.
8. Oston, C., & Widom, J. (2005). Efficient monitoring and querying of distributed, dynamic data via approximate replication. IEEE Data Engineering Bulletin.
9. Rajagopalan, R., & Varshney, P. K. (2006). Data aggregation techniques in sensor networks: A survey. Electrical Engineering and Computer Science, 22.
10. Anastasi, G., Conti, M., & Di Francesco, M. (July 2008). Data collection in sensor networks with data mules: An integrated simulation analysis. In 2008 IEEE Symposium on Computers and Communications (pp. 1096–1102).
11. Deshpande, A., G Guestrin, C.,Madden, S., Hellerstein, J.M., & Hong,W. (2004). Model-driven data acquisition in sensor networks. In Proceedings of the 30th International Conference on Very Large Data Bases (pp. 588–599).
12. Chu, D., Deshpande, A., Hellerstein, J. M., & Hong, W. (2006). Approximate data collee295 ion in sensor networks using probabilistic models. In Proceedings of the 22nd International Conference on Data Engineering (pp. 48–59).
13. Stojejkoa, B. R. (2012) Variable step size LMS algorithm for data prediction in wireless sensor networks. Sensors and Transducers Journal, 14(2), 111–124.
14. Carmalatta J. (2017). A comparative analysis among certain data prediction techniques for wireless sensor network” in the academic science IEEECS ISSN 2348-117X Volume 6, Issue 9 September 2017.
15. Kalman, R. E. (1960). A new approach to linear filtering and prediction problem. Journal of Basic Engineering, 82(1), 35–45.
16. El-Telbany, M. E. (2017).An energy-efficient wireless sensor networks utilizing LMS filter and matrix completion. International Journal of Applied Engineering Research, 12(5), 591–597, ISSN 0973-4562.
17. Intel lab data, Web Page, http://db.ics.mtu.edu/labdata/labdata.html.

Retrieved Number: IJ135079920/2020/BEIESP
DOI: 10.35940/ijitee.IJ135.079920
Published By: Blue Eyes Intelligence Engineering & Sciences Publication

403
18. Saritha, S. B. M. S. V., & Puspalatha, S. C. (2012). Fault tolerant resource management with mutual exclusion algorithm for mobile adhoc networks. International Journal of Scientific and Research Publications, 94.
19. Malikarjuna, B, Anand Dohare, Mhd.Shahajad, Tulika (2019). Feed Forward Approach for Data Processing in IoT over Cloud. International Journal of Innovative Technology and Exploring Engineering (IJITEE), ISSN, 2278-3079, Volume-8 Issue-5 March, 2019.
20. Malikarjuna, B, Anand Dohare, Mhd.Shahajad, Tulika (2019). Master Slave Scheduling Architecture for Data processing on Internet of things. International Journal of Innovative Technology and Exploring Engineering (IJITEE), ISSN, 2278-3079, Volume-8 Issue-5 March, 2019.
21. Malikarjuna, B., & Krishna, P. V (2018), A Nature Inspired Bee Colony Optimization Model for Improving Load Balancing in Cloud Computing. International Journal of Innovative Technology and Exploring Engineering (IJITEE), ISSN, 2278-3079. Volume-8 Issue-2S2 December, 2018.
22. Malikarjuna, B., & Krishna, P. V. (2018). A Nature Inspired Approach for Load Balancing of Tasks in Cloud Computing using Equal Time Allocation, International Journal of Innovative Technology and Exploring Engineering (IJITEE), ISSN, 2278-3079. Volume-8 Issue-2S2 December, 2018.
23. Goodfellow, Yoshua Bengio, Aaron Courville, (2017) Deep Learning, Cambridge, MA:MIT Press.
24. Preethi, G & Krishna, P & S. Obaidat, Mohammad & Saritha, V & Yenduri, Sumanth. (2017). Application of Deep Learning to Sentiment Analysis for recommender system on cloud. 93-97. 10.1109/CITS.2017.8035341
25. Zhang, Shuai et al.(2017) Deep Learning based Recommender System: A Survey and New Perspectives. CoRR abs/1707.07435.

AUTHORS PROFILE

Anand Dohare, Research scholar in SHAUTS, University, Allahabad, India, Published various research paper in top journal, research area is WSN and cloud computing.

Dr. (Mrs) Tulika Assistant professor in the department of Computer Science and Information technology in Sam Higginbottom Institute of Agriculture Technology and Sciences, Allahabad Since November 2006. She received her Ph.D degree in Computer Science from Thapar University Patiala in Vehicular Network field. Her areas of interests are Vehicular Ad hoc network, Ad hoc network and Peer-to-peer network, she Published various book and research paper in renowned journals.

Dr. Basetty Mallikarjuna currently works at the Scool of Computing Science & Engineering, Galgottas University, Greater Noida. Dr. Basetty does research in Cloud computing, Internet of Things, Green Computing, Blockchain and 'Health care Applications'

Mrs. Sweta Sachan pursuing M.Tech from MDU, Rohtak in Computer Science & Engineering. Research area is AI & ML.