Predicting the Energy Efficiency in Wireless Sensor Networks using LSTM and Random Forest Method

Aruna Reddy H.[1], Shivamurthy G.[2], Rajanna M.[3]

[1]Information Science and Engineering, Vemana Institute of Technology (VTU), Bangalore, India
Email: *arunarhvemana[at]gmail.com*

[2]Computer Science and Engineering, VTU CPGSB Muddenahalli, Bangalore, India
Email: *shivamurthygcse[at]gmail.com*

[3]Information Science and Engineering, Vemana Institute of Technology (VTU), Bangalore, India
Email: *rajannam318[at]gmail.com*

Abstract: ***Over wireless sensor networks, consumption of energy due to needless transmitting data is a serious issue (WSNs). By addressing this issue, any station's lifetime can be extended and system practicality for real world applications can be increased. As a result, for WSN's that use low-powered different sensors, energy-efficient information gathering has become a necessity. Data grouping and forecasting approaches based on symmetrical correlation in sensor information can be utilized in order to downplay the complete utilization of energy consumption of the network for sustained collecting data in these situations. We have integrated a group study of Random Forest (RF) Technique, LSTM technique and Particle Filter (PF) which would be used for a efficient method to evaluate and predict the data required by the Sensor nodes to completely minimize the unnecessary data Transmission. Segmentation and data aggregating to each member nodes are used to effectively make data gathering forecasts in WSNs, primarily to reduce the computational overhead cost associated with constructing the prediction model. Simulation trials, comparison, and performance - based assessment in a variety of scenarios reveal that our approach's forecasting accuracy can exceed traditional ARIMA and Kalman filters with Decision tree models, resulting in improved energy consumption due to fewer packet transmissions.***

Keywords: Particle Filter, Random Forest, LSTM, WSN, Clustering

1. Introduction

Self-contained, geographically dispersed sensing devices that govern physical or environmental conditions are known as WSNs. Planning and preparation, smart city traffic surveillance, and environmental sensing are just a few of the applications. Since sensor node cells have a limited capacity and regular maintenance is impracticable in WSNs, power consumption and reliability are significant factors and challenges. The most essential elements influencing energy demand are data collection and propagation of packets of data. This is due to the fact that the nodes must continually and accurately capture all sensed data. During the accurate extraction, consolidation, and transmitting data, such nodes use up a significant total of energy.

Data forecasting [1] is a reasonable technique to cope with these challenges, as it allows one to make predictions based on historical data collected by sensors. There is no need to continuously broadcast the measurements taken by the sensor network while utilizing this method [2]. Some previous research, such as [3,4], has found thatthe information as from full various sensors is transferred to an access point using basic approaches to developing a predictor for the network of sensors. Nevertheless, when measured values vary dramatically and frequently, the prediction approaches used in these studies may not operate properly. This challenge can be solved by local forecasting hierarchical clustering in sensor networks. The local predictive algorithm would've been environmentally friendly since sensing information is exchanged over a smaller transmission route. Local projections hierarchical clustering, from the other side, has a lot of shortcomings. The very first difficulty is the massive price of developing a predictors, which is determined by the informational vs. computation interchange.

Here there is a main principle to save energy, this work has made an efficient use of filtration and clustering techniques can be optimized to enhance the spatial - spectral coherence of sensor information. This study employs a auto self tuned strategy based on the PF [5–8], which has a tremendous opportunity because it allows for non-biased and optimal estimate while reducing covariance errors. By lowering the number of hops, the experts [9] attempted to improve connect energy usage during the data transmission. The energy usage of nodes can be adjusted by installing a new link that allows for faster data transfer. There are various filters which is used for this procedure, such as the Kalman Filter. The Kalman Filter can be used for straight or linearized processes and measurement systems, whereas the PF can be used for nonlinear systems and linear systems as well. Furthermore, the Kalman filter's ambiguity is limited to Gaussian distributions, but the PF can handle non-Gaussian noise distributions.

The data prediction proposed technique for a type of cluster based wireless sensor network's major goal is to reducing the power consumption caused by radio communications by decrease the quantity of transmissions within both the transmitter and the recipient. To that aim, it must efficiently conduct data sampling predictions in WSNs simultaneously aggregating and collecting evidence to each member nodes in order to reduce latency.

The sensor data model must be taken account, and the major components that play a part in the prediction phase must be identified. These elements can be summarised as follows:

The First one would be time connection data has a periodic and is dependent on historical data; second would be spatial correlation—sensory data of wireless sensor nodes is dependent on the data of its neighbors; third would be data integrity regulations as sensitive data is missed or a sensor node fails noised version equated with the novel value, and data quality can be best by data preprocessing.

Machine learning has advanced rapidly in recent years. For its memory storage, the Recurrent Neural Network (RNN) offers a variety of applications in time-series data prediction,speech recognition, and machine translation.

The creation of RNN led to the development of the Long Short-Term Memory (LSTM). It acts well in processing time series data with long-term relationships and forecasting long-interval occurrences [12–15]. The predictions model's efficiency and quality can be improved by utilize LSTM neural network to excerpt and fuse fine sensory input with correlation.

We created a data prediction based on RF, autoregressive moving averages, and PF approaches to eliminate needless information transmission and, as a function, lower energy usage. Based on terms of prediction and data management, this approach uses limited sensors to amass data. In order to generate an amount of trees for collecting sensor data, the suggested system uses the RF to filter information linked for each node. Moreover, estimations are improved while associated defects are reduced using a particle filtration self-tuning method.

The following is a list of the remainder part of this document: The associated works are explained in subdivision 2. The preparation approaches and primary modelling system are represented in Section 3. The performance of the recommended strategy is compared to all different methods in Section 4, and the article is wrapped up in Section 5.

2. Related Works

The production of a psychological events, such as a theory that describes data growth, is what data projections involves. To meet this goal, data projection methodologies could be categorized into three parts: stochastic, time series prediction, and algorithmic strategies. The stochastic classifying technique's biggest disadvantage [10] is its high computational cost, which may be problematic for sensing devices have limited processing and capability. Randomized techniques are more suitable when there are numerous strong sensors [11]. The strategies for time series prediction [12] can provide enough exactitude when using easymethodologies.

To characterise and forecast human motion patterns, Kolodziej and Xhafa [8] suggested an interaction technique Markov chain model. They subsequently utilised the Nonparametric Belief Propagation approach to forecast which regions would've been frequented in the coming and what would not. Liu et al. [9] developed an artificial neural machine-based environmental data forecasting models. The algorithm is designed to increase forecasting velocity while maintaining accuracy. Sinha et al. [10] introduced a traced back prediction-based data aggregating methodology called TDPA. When the forecast meets a predetermined threshold, the model creates an evaluate of future data to examine the forecast error and uses the anticipated value to preserve transmit energy consumption.

In WSN applications, clustering methods are used to organise and cluster nodes, also to indicate which node in each group is responsible for intra-cluster and inter-cluster data substitutes. This technic has the ability to reduce the quantity of data transmissions, hence lowering energy usage and extending the lifespan of the network. A few of these implementations [15] is an approach for local sensor network forecasting in which the central server acts as a sensor network while also storing past data for the sensor nodes of each group.

A few of the decentralized voting systems was created in [16]. The tree architecture of the sensing functions as a small, robust database server at the tree from the root, intended to overcome problems of categorization while applying the algorithms. [3,4,13] presents some techniques for dealing with such failures. References [3] developed a data forecasting methodology for reducing SN load and extending lifetime of the network. In compared to existing data prediction approaches such as the multiple linear regression, experimental findings with the suggested model increased reliability and energy consumption. The authors of [4] presented a data reduction strategy based on the development of a modeling from both the network edge and Smart devices like IOT devices.

Prognostication real - time agglomerative methods [17] is another forecasting approach for arranging a stable clusters of vertices data - based clustering. The given modeling approach has a reasonable performance and is computationally inexpensive. Scientists in [18] set out now to adopt the multi-level full distance segmentation method in decentralized grouping algorithms in order to rescue energy. To shorten the production of unneeded navigation data transmission, the suggested technique creates a cluster and scheduling tree.

A further prediction methodology for constructing suitable long - term of vertex information grouping is forecasting real-time fuzzy clustering approaches [17]. The proposed model structure has a good performance and is computationally cheap. In order to reduce energy consumption, researchers in [18] have decided to use the inter complete range segmentation technique in decentralized clustering techniques. The proposed method establishes a cluster and scheduling tree to reduce the production of unnecessary navigational data transfer. Chen and Du intended a method for stock prediction that incorporates sentiment analysis and online social behavior inspection. It considers the association during transaction volume or price and these traits by generating social activity graphs and determining key features [22]. Teye and Ahelegbey investigated the geographic link among property prices in twelve Dutch regions using Bayesian Graphical Vector Autoregressive model. The findings shows the patterns of housing price diffusion in the Netherlands as well as the patterns [24]. Recurrent convolutional neural networks (RCN) were utilised by Che-Yu Lee to predict stock price. To collect information from the financial news, the suggested predictive algorithm uses convolutions, word embedding, and sequence modelling, followed by technical indicators signals to forecast price of the stock [25].

3. Proposed Methodology

We have demonstrated a full-fledged conceptual approach for energy-efficient data collecting using RF, LSTM, and PF approaches in this research study as presented in the below figure - 1. The methods are described first, followed by the proposed hybrid model in the subsequent subsections.

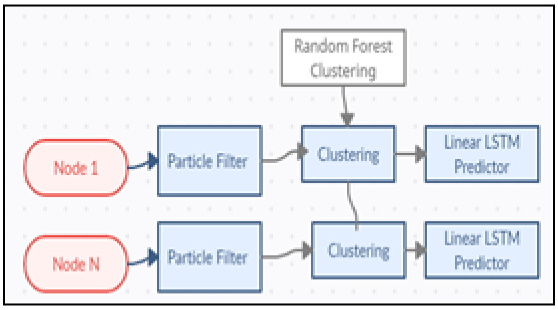

Figure 1: Hybrid Model Architecture

3.1 Techniques Used

The complete techniques used in this process are :

- The particle filter is an algorithm that predicts unknown factors based on the measured taken over period. This filtering is being used to calculate states in state-space based on linear difference equations as well as non-linear formats. It has a basic design, consumes little computational resources, and is more economical and efficient.
- Random forest is an easy-to-understand and interpret classification system. By learning simple decision rules inferred form previous information, The goal of RF is to form a training model that could be used to anticipate the responder variable's category or amount.
- LSTM is a high-level analysis model that predicts future trends. It's a mixture of an autoregressive and a simple movable framework which is highly efficient than ARIMA and other prediction techniques, This LSTM technique can calculate in linear and nonlinear as well.

3.2 The Combined Methodology

An appropriate analyzing needs of a electronic sensor network are predicted by this approach. PF are used to filtering the sensory streams of data connected for each node in the proposed system. This is based on the implementation of auto configuring error correlation with rapid autonomous system when the signals load varies suddenly. RF derives a tree for grouping the detector associated data for each base station using filtered data associated with each node. A good predictive technique based on LSTM aided approach is designed by merging the RF and PF. In comparison to other neural network models such as the Neural convolutional network, the RF is understandable, rapid, and computational efficiency, making it ideal for use in real-time detection and monitoring.

Information inside every cluster can be aggregated to produce certain specific clusters, with it's own member nodes and its heads, thanks to the enormously featured RF typical tree's hierarchy organization. A hierarchical organization is used to group neighboring nodes against with a cluster - head. This graded organizational structure can minimize communication cost and save energy by dynamically through the network node for data collecting in the cluster's range. [21]

As previously said, RF is a categorization technique that generates a tree using data from all vertices. The cluster members gather information about each node, including its ID, remaining energy, and location in the network. The cluster members keep track of this data in a list style. As previously said, RF is a categorization technique that generates a tree using data from all nodes. The Cluster Heads (CHs)collect data about each node, including its ID, remaining power, and location in the network. The cluster members keep track of this data in a list style. Numerical simulations, such as the linear LSTM model [23], can be used to find out the correlation.As a result, the evaluation can be carried out using correct mathematical forms, and the amount of design variables is usually significantly fewer than the sequence' overall length. The sensor network can then choose to transmit its real number first.

PF has a simple design and uses little computing power. To limit the quantity of data transmissions, an efficient information filtration mechanism must be used to remove data redundancy at the sensor as well as at the clusters. This research presents a networked KF to treat data serial noise in order to improve the lifespan of sensor network by minimising data transfer redundancy and preserving power during continuous data gathering. The broad theory on state estimation calculates data filtering to minimise the error of correlation. As a result, the final filter's adaptation to changes is determined by the consensus and its best estimate. The following Equation (1) is considered to define a linearly generative model:

$$y(l + 1) = Gy(l) + W(l) \ldots\ldots\ldots\ldots \text{(i)}$$

Here the l shows the representation of the time dimension and here t is time index; y(l) is state of a system; the Y representation of N * N shows a glimpse of the way the whole architecture is subject to change over the dimension of the time and W(l) is known for evaluating the the uncertainties The Gaussian noise at correlation coefficients was used to simulate as a factor of these weight matrix.

3.3 Auto tuning configuration of Random Forest Clustering Technique

To evaluate local actions, a mechanism for customising grouping is imperative. A total re-clustering is one possibility, and it could be expensive because it requires the creation of a map for aggregating all of the data points for each sensor node. The full change in cluster affiliation recommends that all previous data and models should be constructed from the ground up. RF listed in this section is used to accomplish initial grouping and then dynamic division and combination of clusters with limited communication costs. The cluster centers are selected based on randomized node sets in the early phase. When every sensing node's command data is obtained,

From the whole set of sensor network, RF should be used to identify a strong positive significant CH. Initially, it uses Euclid proximity to generate a clustering tree utilising distances between observations in the data. After that, every sensor node receives a list of CHs. After that, each sensing node connects to its CHs. The updating for adaptable grouping is typically required to monitor the difference in the structures of the localization; nevertheless, most clustering methods can be used in such a study. A comprehensive re-clustering is a possibility, although it could be expensive. This entails creating a map for clustering all of the devices.

From the whole set of sensor nodes RF should be used to identify a strong positive significant CH. Initially, it uses Euclid proximity to generate a clustering tree utilizing similarities between observations in the data. After that, every sensors node receives a list of CH. After that, each sensing node connects to its CH. The updating for adaptable grouping is typically required to monitor the difference in the structures of the localization; nevertheless, most clustering algorithms can be used in such a study. A comprehensive re-clustering is a possibility, although it could be expensive. This entails creating a map for clustering all of the devices.

Utilizing local projections based on an erroneous limitation E> 0, the sensor network can potentially send its numeric value to the CH. In conclusion, the following factors influenced the proposed strategy in this study: I node distance from cluster centre; (ii) motion; (iii) battery technology remaining; and (iv) symptoms of weakening The index value clearly displays an efficient way for calculating node susceptibility in the tree structure of a sensors. To determine each network node vulnerability factor, the following equation can be used as shown below in the Equation ii.

$$L_{index} = \frac{h_i}{h_j} * \frac{B_i}{B_i} \qquad \text{--- eqn (ii)}$$

where hi is the set of nodes well before i-th node is abolished, hj is the number of nodes after the i-th network has been removed, bi the layers even before i-th component has been removed, and aj the number of layers before the i-th node has been removed Each node, as well as the r group centre, is evaluated. A lower number means the node is more likely to be the CH. The larger the power of the batteries, the better. the more likely the node will become a CH. The node's movement has a substantial impact on the network's lifespan. After the begin phase is complete, each sensor node sends its data to a specified CH.

Furthermore, the greater the current battery capacity, more the likely the node will become a stronger leader. The node's movement has a substantial impact on the cable network lifespan. After the setup process is complete, each sensor node sends its information to a designated CH, which then publishes the users' list to other nodes [25]. The clustering procedure is repeated using a specified intervals or by achieving the predefined threshold requirement. The selected transmitted information approximates e-loss as follows: If $| Y_t - \vec{Y}_t | > e_t$ a sensing node send the Y_t value to the cluster due to an error limit E > 0, where indicates a projected data value.Itwillvain to record if the actually gets was close to the awaited value. In evaluating distribution of the data, the deviation of the given value from the predicted value is a critical element.

3.4 LSTM Prediction Method

The methods used here work intended to create a great Long short - term memory algorithm for forecasting sensor network energy consumption during data collecting. The history of the data series is used to explain how the parameters react to a previous unpredictable fluctuation, making LSTM models univariate. Following acquiring historical information for the required parameters, LSTM could be implemented in a multiple approach. The basic stages are as follows: I identify the system; (ii) calculate the attributes; (iii) verify and forecast the prototype; and (iv) verify and estimate the model. [26]. The following is an example of a general LSTM (R, S, M) model for expressing time series:

$$\theta(A) = \Delta^c * Y_t = \varphi(A) e_t \qquad \text{........Eqn (iii)}$$

Yt and et, from the other hand, are employed to indicate energy consumption and error at time t, correspondingly. A is the backward shift operator specified by A * Yt = Yt 1 and roughly 5; d is the degree of interpolation. The abstraction features of sensory data are mined and learned by bidirectional LSTM network. At last, using the merging layer of the neural network, he abstraction characteristics are used in time series forecasting. In several assessment indicators, the suggested MNMF model outperforms the prior techniques, pursuant to the experimental results.

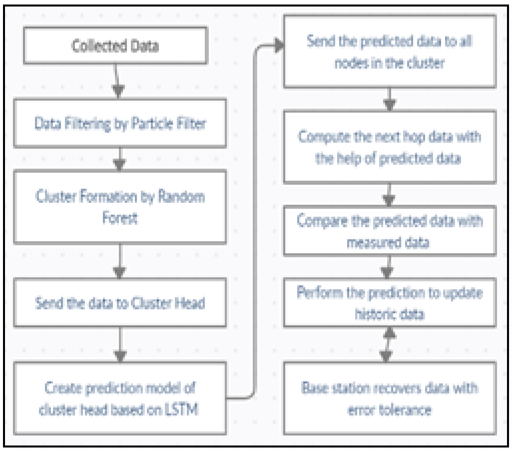

Figure 2: Architecture Flow of the work

4. Result and Discussion

We ran a number of tests to evaluate the suggested computation effectiveness to that of other methodologies. To evaluate the performance of our prediction system, we used sensor datasets from a top Software Research Facility [27]. During such month, data was collected from 54 nodes dispersed across their laboratory. We used the means of the measured at various temporal epochs to fill in the blanks numeric value. For a period of ten days, we picked all of the temperature recordings (17 January to 28 January). In the radio transmission range, each node has an ordinary of five clusters.

To create 2000 numeric value, all vertices were utilised. Alpha = 1 and X0 = 0 were used to initialise each component. The approaches of dispersed densities estimation based on maximum likelihood method [28–30] were chosen for comparison. The technique [31] keeps track of the most recent data sent to the sink and sensors. The sensor network will not be able to broadcast the data value if it has an error bound. A cumulative distribution function (cdf) refers to a set of acquired characteristics in this model's training stage. These include a networked approach to build a Gaussian mixture model (GMM) and a modeling approach [32], as well as an enhanced resilient aggregation technique to pull statistical information from sensor networks reported in [30].

Our method incorporates a Gaussian Model method for group splitting or inter-cluster consolidation in order to provide a valid comparison. Because we plan to employ a mixed Random forest optimised LSTM model in our systems, the communication sizes of the 2 techniques should be comparable to allow for a valid comparison. Each clusters player's history data is saved in a series of concentric by the cluster leader. We estimate the energy usage of member nodes because our model aids cluster individuals rather than cluster chiefs. On average, there are 17 clusters with 35 members. The total energy usage of all 35 cluster members is shown in Figure 16 [16]. Because non-stationary qualities are present in energy consumption data, multiple strategies must be used to alter the non-stationary properties.

The stable mechanism was verified in respect of the LS and MA parameters before adopting LSTM. Table 1 shows the parameters for the LS and MA matrices. The data indicate the LS and MA parameters' ideal values. As measurements increases, the LS and MA indexes will tend to be become zero. The LS and MA indices were all weak, as seen in Table 1. The indexes must add up to one. Furthermore, their good standard deviation of 4.05 to 7.85 were not noteworthy in any of the cases. The heterogeneity of distribution of data resulted in low reliability (or quite high standard deviation). Following the examination of the LS and MA coefficients, the hybrid Language model is determined to be the best option.

The amount of auto-regressive terms is represented by r, the number of lagged prediction errors is represented by s, and the amount of non changes is represented by m in LSTM(r, s, m). Errors (t) are considered to be independent and have an equivalent distribution with a constant variance. Main indices such as average relative error (ARE), (RMSE), and (MAE) were utilized to measure the performance of the model and displayed the models' prediction accuracy. Table 2 lists the fundamental LSTM parameters of the model.

Table 1: Total indices with various quantities of items of Member nodes

	Number of Nodes		
No	500	1000	1300
ARE	0.3	3.8	3.5
MAE	0.4	0.9	0.32
RMSE	0.5	1.1	0.4

Figure 3 depicts how the suggested methodology operates on clusters 1's cluster members and cluster formation node. The limit for inaccuracy is fixed to 0.1. Figure 3a reflects the actual temperatures and predicted temperatures for 2000 observations. Figure 3b is a larger version of Figure 3a that includes 600 observations from 600 to 800 seconds. The matching and predicting results are shown in Figure 3. In all dataset, the optimized hybrid version reduces the prediction error depending on the PF–RF optimized LSTM network. This is a significant advancement in power consumption predictions. The suggested model in comparison to [13], which used LSTM to predict data from sensor networks.

In [13], LSTM was applied to stationarity, which displays that now the serial correlation have had no tendency, have small changes of the average with a consistent magnitude, have the same short-term strange patterns throughout time.

Forecasting accuracy is calculated using conventional performance measures such as ARE, RMSE, and MAE. To test efficiency responsiveness, the number of nodes was adjusted. It's worth noting that each represents the ratio of transmitting power consumption to forecast energy consumption. The data are displayed in Figure 4:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}}$$

$$MAE = \frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{n}$$

$$ARE = \frac{\sum_{i=1}^{n}(|y_i - \hat{y}_i| / y_i)}{n}$$

…eqn (iii)

Fig 4 depicts the performance indices ARE, RMSE, and MAE as a function of the amount of networks (500, 1000, and 1500), with our hybrid version having a lower RMSE, MAE, and ARE for varied node density. This result demonstrated the mixed model's capacity to effectively reduce the predicted value inaccuracy. Appropriate results can be proved with excellent accuracy using best linear and non-linear methods, especially when both methods demonstrated accurate and high projecting strength, especially if both models showed accurate and great predicting ruggedness.

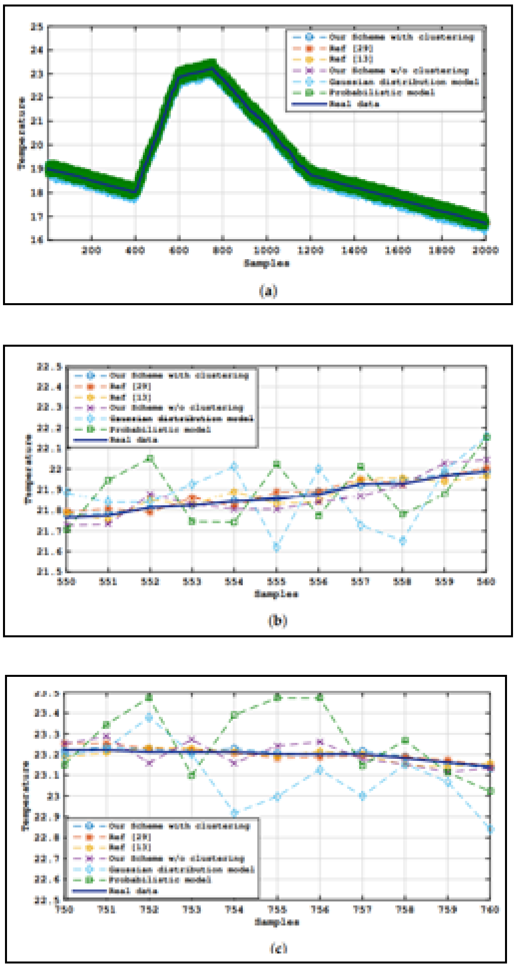

Figure 3: The whole quantitative of packets vs the count of cluster nodes

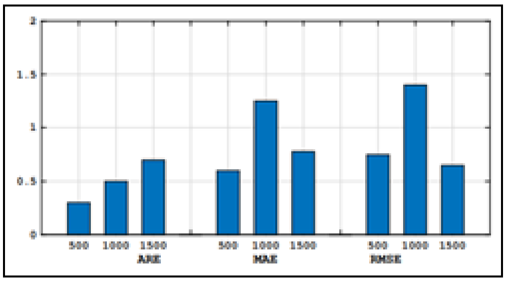

Figure 4: Performance Indices for ARE, MAE, RMSE for number of the nodes

5. Conclusion

An energy-efficient data gathering strategy hierarchical clustering and predictions is designed in this paper. Sensor networks form clusters during the grouping phase, while cluster members gather and store data collected by sensor network. The designed hybrid predictions model was used to consider the communication-prediction tradeoff. The model's efficiency has indeed been assessed using lots of numbers of vertices across various periods. The suggested model greatly outperforms other linkedadvances in prediction accuracy and energy consumption, according to the simulation study. As a result, it can drastically lower the amount of energy consumed for data collection in hierarchical networks while also significantly extending the network lifetime—even when a large number of clusters are allotted.Future research could include integrating the Traffic future producer, such as a SigFox network provider, with the application. To lower the initial implementation expenses, we intend to use the SigFox traffic generator to quickly experimentally test the suggested product's system dispersion over a large geographical area.

References

[1] Arbi, I.B.; Derbel, F.; Strakosch, F. Forecasting methods to reduce energy consumption in WSN. In Proceedings of the 2017 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Turin, Italy, 22–25 May 2017; pp. 1–6.

[2] Jiang, H.; Jin, S.; Wang, C. Prediction or not? An energy-efficient framework for clustering-based data collection in wireless sensor networks. IEEE Trans. Parallel Distrib. Syst. 2010, 22, 1064–1071. [CrossRef]

[3] Agarwal, A.; Dev, A. A Data Prediction Model Based On Extended Cosine Distance For Maximizing Network Lifetime of WSN. Wseas Trans. Comput. Res. 2019, 7, 23–28.

[4] Tayeh, G.B.; Makhoul, A.; Laiymani, D.; Demerjian, J. A distributed real-time data prediction and adaptive sensing approach for wireless sensor networks. Pervasive Mob. Comput. 2018, 49, 62–75. [CrossRef]

[5] Peksa, J. Prediction Framework with Kalman Filter Algorithm. Information 2020, 11, 358. [CrossRef]

[6] Ferrero, R.; Pegoraro, P.A.; Toscani, S. Synchrophasor Estimation for Three Phase Systems Based on Taylor Extended Kalman Filtering. IEEE Trans. Instrum. Meas. 2020, 69, 6723–6730.

[7] Zhou, J.; Gu, G.; Chen, X. Distributed Kalman filtering over wireless sensor networks in the presence of data packet drops. IEEE Trans. Autom. Control. 2018, 64, 1603–1610. [CrossRef]

[8] Wang, J.; Zhu, R.; Liu, S. A differentially private unscented Kalman filter for streaming data in IoT. IEEE Access 2018, 6, 6487–6495. [CrossRef]

[9] Pandey, O.J.; Hegde, R.M. Low-latency and energy-balanced data transmission over cognitive small world WSN. IEEE Trans. Veh. Technol. 2018, 67, 7719–7733. [CrossRef]

[10] Yan, R.; Sun, H.; Qian, Y. Energy-aware sensor node design with its application in wireless sensor networks. IEEE Trans. Instrum. Meas. 2013, 62, 1183–1191. [CrossRef]

[11] Goel, S.; Passarella, A.; Imielinski, T. Using buddies to live longer in a boring world [sensor network protocol]. In Proceedings of the Fourth Annual IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOMW'06), Pisa, Italy, 13–17 March 2006; p. 5.

[12] Jang, W.S.; Healy, W.M.; Skibniewski, M.J. Wireless sensor networks as part of a web-based building environmental monitoring system. Autom. Constr. 2008, 17, 729–736. [CrossRef]

[13] Diwakaran, S.; Perumal, B.; Devi, K.V. A cluster prediction model-based data collection for energy

efficient wireless sensor network. J. Supercomput. 2019, 75, 3302–3316. [CrossRef]

[14] Xiangning, F.; Yulin, S. Improvement on LEACH protocol of wireless sensor network. In Proceedings of the 2007 International Conference on Sensor Technologies and Applications (SENSORCOMM 2007), Valencia, Spain, 14–20 October 2007; pp. 260–264.

[15] Ouni, S.; Gherairi, S.; Kamoun, F. Data aggregation and pipelining scheduling protocols for real–time wireless sensor networks. Int. J. Hoc Ubiquitous Comput. 2013, 12, 56–64. [CrossRef]

[16] Lavric, A.; Petrariu, A.I.; Coca, E.; Popa, V. LoRa Traffic Generator Based on Software Defined Radio Technology for LoRa Modulation Orthogonality Analysis: Empirical and Experimental Evaluation. Sensors 2020, 20, 4123. [CrossRef] [PubMed]

[17] Ashouri, M.; Yousefi, H.; Basiri, J.; Hemmatyar, A.M.A.; Movaghar, A. PDC: Prediction-based data-aware clustering in wireless sensor networks. J. Parallel Distrib. Comput. 2015, 81, 24–35. [CrossRef]

[18] Sabet, M.; Naji, H. An energy efficient multi-level route-aware clustering algorithm for wireless sensor networks: A self-organized approach. Comput. Electr. Eng. 2016, 56, 399–417. [CrossRef]

[19] Yang, X.; Wen, Y.; Yuan, D.; Zhang, M.; Zhao, H.; Meng, Y. 3-D compression-oriented image content correlation model for wireless visual sensor networks. IEEE Sens. J. 2018, 18, 6461–6471. [CrossRef]

[20] Gedik, B.; Liu, L.; Philip, S.Y. ASAP: An adaptive sampling approach to data collection in sensor networks. IEEE Trans. Parallel Distrib. Syst. 2007, 18, 1766–1783. [CrossRef]

[21] Zhang, D.; Li, G.; Zheng, K.; Ming, X.; Pan, Z.H. An energy-balanced routing method based on forward-aware factor for wireless sensor networks. IEEE Trans. Ind. Inform. 2013, 10, 766–773. [CrossRef]

[22] Shokouhifar, M.; Jalali, A. Optimized sugeno fuzzy clustering algorithm for wireless sensor networks. Eng. Appl. Artif. Intell. 2017, 60, 16–25. [CrossRef]

[23] Olofsson, T.; Ahlen, A.; Gidlund, M. Modeling of the fading statistics of wireless sensor network channels in industrial environments. IEEE Trans. Signal Process. 2016, 64, 3021–3034. [CrossRef]

[24] Imani, M.; Dougherty, E.R.; Braga-Neto, U. Boolean Kalman filter and smoother under model uncertainty. Automatica 2020, 111, 108609. [CrossRef]

[25] Yin, Y.; Shi, J.; Li, Y.; Zhang, P. Cluster head selection using analytical hierarchy process for wireless sensor networks. In Proceedings of the 2006 IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications, Helsinki, Finland, 11–14 September 2006; pp. 1–5.

[26] Box, G.E.; Jenkins, G.M.; Reinsel, G.C.; Ljung, G.M. Time Series Analysis: Forecasting and Control; John Wiley & Sons: Hoboken, NJ, USA, 2015.

[27] Madden, S. Intel Lab Data; Intel Research Lab: Berkeley, CA, USA, 2006.

[28] Jiang, H.; Jin, S. Scalable and robust aggregation techniques for extracting statistical information in sensor networks. In Proceedings of the 26th IEEE International Conference on Distributed Computing Systems (ICDCS'06), Lisboa, Portugal, 4–7 July 2006; p. 69.

[29] Wu, M.; Tan, L.; Xiong, N. Data prediction, compression, and recovery in clustered wireless sensor networks for environmental monitoring applications. Inf. Sci. 2016, 329, 800–818. [CrossRef]

[30] Chu, D.; Deshpande, A.; Hellerstein, J.M.; Hong, W. Approximate data collection in sensor networks using probabilistic models. In Proceedings of the 22nd International Conference on Data Engineering (ICDE'06), Atlanta, GA, USA, 3–7 April 2006; p. 48.

[31] Pala, Z. Effects of Mica2-based discrete energy levels on the lifetime of cooperation neighbor sensor networks. Turk. J. Electr. Eng. Comput. Sci. 2016, 24, 2671–2678. [CrossRef]

[32] Barak, S.; Sadegh, S.S. Forecasting energy consumption using ensemble ARIMA–ANFIS hybrid algorithm. Int. J. Electr. Power Energy Syst. 2016, 82, 92–104. [CrossRef]