Indian Journal of Computer Science and Technology
Volume 5, Issue2 (May-August 2026), PP: 07-12.
www.indjcst.com
ISSN No: 2583-5300

A Comprehensive Survey of Prediction-Driven Communication Models in WSNs

Nitin Singh[1], Renu[2]
[1]P.G. Student, Department of CSE, Sat Kabir Institute of Technology and Management, Ladrawan, Haryana, India.
[2]Assistant Professor, CSE, Sat Kabir Institute of Technology and Management, Ladrawan, Haryana, India.

To Cite this Article: *Nitin Singh[1], Renu[2], "A Comprehensive Survey of Prediction-Driven Communication Models in WSNs ", Indian Journal of Computer Science and Technology, Volume 05, Issue 02 (May-August 2026), PP: 07-12.*

Abstract: *Wireless Sensor Networks (WSNs) have emerged as a fundamental technology for a wide range of applications, including environmental monitoring, smart agriculture, healthcare, and industrial automation. However, the limited energy resources of sensor nodes and the high cost of data communication pose significant challenges to network efficiency and lifetime. In recent years, prediction-driven communication models have gained considerable attention as an effective approach to reduce unnecessary data transmissions by exploiting temporal and spatial correlations in sensed data. These models enable sensor nodes to locally predict future readings and transmit data only when the prediction error exceeds a predefined threshold, thereby minimizing communication overhead. This survey provides a comprehensive review of prediction-based communication techniques in WSNs, covering model-based, threshold-driven, and machine-learning (ML) based approaches. It also examines how these techniques are integrated with routing strategies, particularly multi-hop communication, to further enhance energy efficiency. The paper analyses existing methods in terms of energy consumption, accuracy, computational complexity, and network scalability. Additionally, key challenges, limitations, and trade-offs associated with prediction-driven communication are discussed. Finally, the survey highlights open research issues and future directions for developing more robust and adaptive prediction models for next-generation WSN applications.*

Key Words: *Prediction-Based Communication, Multi-Hop Routing, Energy Efficiency.*

I. INTRODUCTION

WSNs have emerged as a key enabling technology for a wide range of applications, including environmental monitoring, smart agriculture, healthcare, industrial automation, and smart cities. A typical WSN consists of a large number of sensor nodes deployed over a geographical area to collect and transmit data to a central base station. These sensor nodes are generally resource-constrained, with limited battery power, processing capability, and memory. Among all operations, wireless communication is the most energy-consuming task, making energy efficiency a critical design concern in WSNs [1].

To address energy constraints, numerous routing and communication protocols have been proposed. Early approaches such as LEACH introduced **cluster-based communication** to reduce transmission distance and balance energy consumption among nodes [1]. Similarly, PEGASIS improved energy efficiency by organizing nodes into chains, allowing data to be transmitted through neighboring nodes rather than directly to the base station [2]. While these approaches reduce transmission costs through structural optimization, they do not directly address the issue of redundant data transmission, in which nodes repeatedly transmit similar or unchanged data.

In recent years, attention has shifted toward data-centric and intelligent communication models that aim to reduce unnecessary transmissions. Protocols such as TEEN introduced threshold-based communication, where data is transmitted only when significant changes occur in sensed values [3]. Likewise, negotiation-based protocols like SPIN reduce redundancy by exchanging metadata before actual data transmission [4]. Although these methods improve efficiency, they often lack adaptability and may not fully exploit the temporal correlation present in sensor data. Prediction-driven communication models have been proposed as a promising solution to overcome these limitations. In such models, sensor nodes use local prediction techniques to estimate future data values and transmit information only when the prediction error exceeds a predefined threshold. This approach effectively transforms the communication paradigm from periodic to event-driven, significantly reducing the number of transmissions and conserving energy. Furthermore, when combined with multi-hop routing strategies, prediction-based communication can further minimize transmission distance and more evenly distribute energy consumption across the network [5].

This survey aims to provide a comprehensive overview of prediction-driven communication models in WSNs, with a focus on their integration with energy-efficient routing techniques. The paper reviews existing methods, analyses their performance, and identifies key challenges and research gaps. By highlighting recent advancements and future directions, this survey contributes to the development of more efficient and intelligent communication strategies for next-generation WSNs.

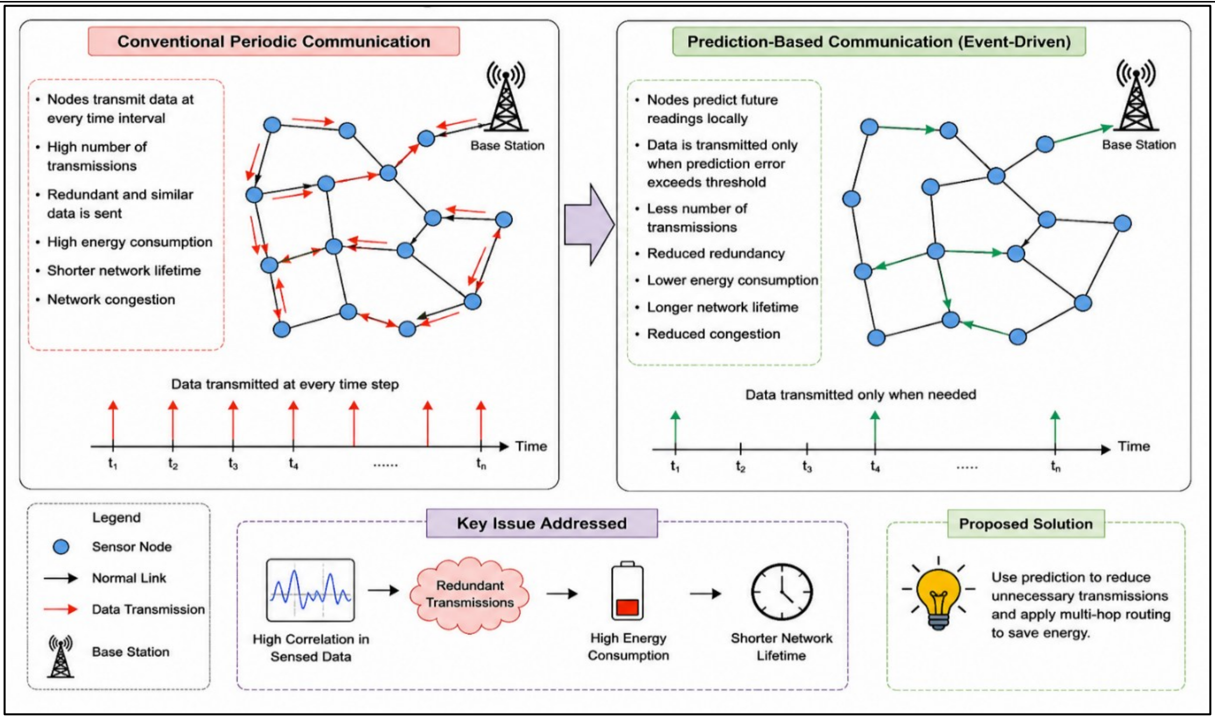


Figure 1: Communication Issues in WSN

II.RESEARCH BACKGROUND

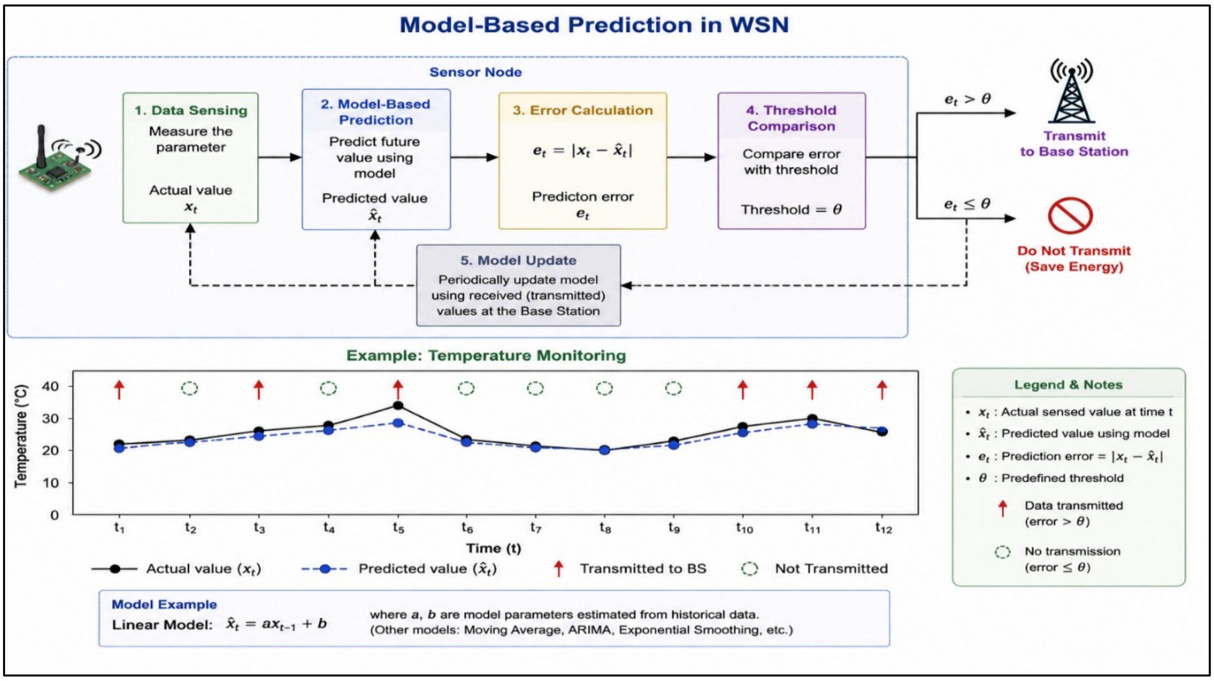
Energy-efficient communication in WSNs has been extensively studied, particularly through clustering-based approaches, which aim to reduce transmission distance and balance energy consumption among sensor nodes. Several researchers have proposed improved clustering strategies to enhance network lifetime and address issues such as heterogeneity, scalability, and load imbalance. Saswati Bandyopadhyay and Edward J. Coyle [6] proposed an energy-efficient hierarchical clustering algorithm that optimizes cluster formation to reduce overall energy consumption. Their approach improves network lifetime by organizing nodes into clusters with efficient communication patterns. However, the method assumes homogeneous node capabilities and does not fully address energy imbalance among cluster heads.

To address heterogeneity in WSNs, Ioannis Matta et al. [7] introduced the Stable Election Protocol (SEP), which considers nodes with different energy levels. SEP improves stability by assigning higher probabilities of becoming cluster heads to nodes with more energy. While this enhances network lifetime, it still relies on periodic data transmission and does not reduce redundant communication. Similarly, Li Qing et al. [8] proposed a distributed energy-efficient clustering algorithm for heterogeneous networks. Their work extends clustering techniques by incorporating energy-aware cluster-head selection, resulting in improved energy distribution. However, like previous approaches, the protocol primarily focuses on routing efficiency rather than minimizing data transmission. Further improvements in clustering strategies were presented by Amin Bari et al. [9], who proposed clustering mechanisms for two-tiered sensor networks. Their approach enhances scalability and lifetime by organizing nodes into hierarchical layers. Despite these improvements, the model still suffers from communication overhead due to continuous data transmission.

Mohamed Younis et al. [10] introduced an energy-aware management framework for cluster-based WSNs, emphasizing efficient cluster formation and maintenance. Their work highlights the importance of balancing energy consumption across nodes, but it does not incorporate mechanisms to reduce redundant sensing data transmission. In [11], J. S. Liu and C. H. R. Lin proposed an energy-efficient clustering protocol that improves communication efficiency through optimized cluster structures. While this approach enhances performance, it still relies on a time-driven communication model, resulting in unnecessary transmissions. Unequal clustering techniques have also been explored to address the hot-spot problem, in which nodes near the base station deplete their energy faster. Sangho Lee et al. [12] proposed a distributed unequal-clustering protocol that assigns smaller clusters near the base station to balance energy consumption. This approach improves network lifetime but does not eliminate redundant communication. The hotspot problem was further analyzed by Michele Perillo et al. [13], who evaluated strategies to mitigate uneven energy depletion in sensor networks. Their study highlights the importance of load balancing and routing optimization, but does not consider data reduction techniques such as prediction-based communication.

III. PREDICTION-BASED COMMUNICATION TECHNIQUES IN WSNS

Model-Based Prediction Approaches: Model-based prediction techniques use mathematical or statistical models to estimate future sensor readings from previously observed data. In these approaches, both the sensor node and the base station maintain a common prediction model, such as linear regression, a moving average, or a time-series model. The sensor node compares the sensed value with the predicted value, and transmits only when the prediction error exceeds a predefined threshold. These methods are computationally lightweight and suitable for resource-constrained sensor nodes, making them widely used in WSNs. They are particularly effective in environments where sensed data exhibits smooth, predictable variations, such as in temperature or humidity monitoring. However, their performance may degrade in highly dynamic or unpredictable conditions [14].



- x_t : Actual sensed value at time t
- $\hat{x}_t$: Predicted value using model
- e_t : Prediction error $= |x_t - \hat{x}_t|$
- θ : Predefined threshold

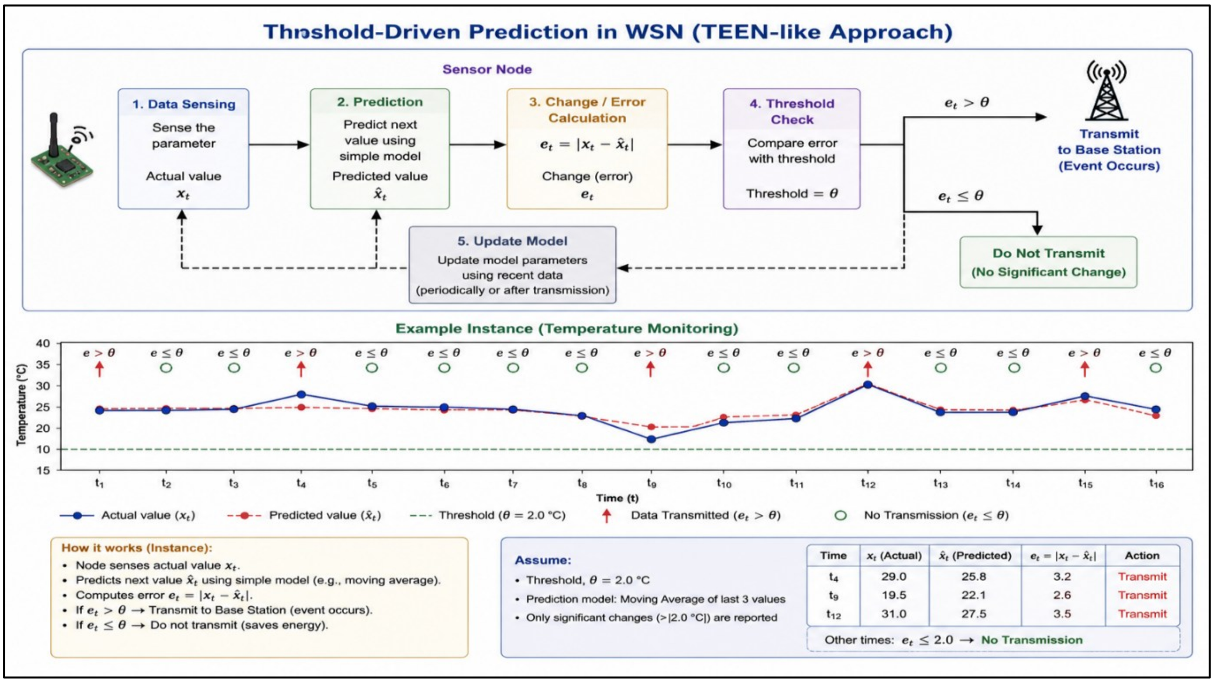
Model Example

Linear Model: $\hat{x}_t = ax_{t-1} + b$ where a, b are model parameters estimated from historical data. (Other models: Moving Average, ARIMA, Exponential Smoothing, etc.)

Figure 2: An instance of Model-Based Prediction in WSN

Threshold-Driven Prediction Approaches: Threshold-driven approaches define one or more threshold values that control data transmission. In this method [15], sensor nodes continuously monitor the sensed parameter but transmit data only when the value crosses a specified hard threshold or when the change in value exceeds a soft threshold. This technique is commonly used in event-driven applications where only significant changes are of interest, such as fire detection or intrusion monitoring. By eliminating the need for continuous transmission, threshold-driven approaches significantly reduce communication overhead and energy consumption. However, selecting appropriate threshold values is critical, as overly strict thresholds may lead to excessive transmissions, while overly relaxed thresholds may result in loss of important information.

Figure 3 illustrates the working of a threshold-driven prediction approach in a WSN, similar to TEEN-based communication. It is divided into two main parts: the sensor node processing flow (top section) and a practical example using temperature data (bottom section).



How it works (Instance):

- Node senses actual value x_t.
- Predicts next value $\hat{x}_t$ using simple model (e.g., moving average).
- Computes error $e_t = |x_t - \hat{x}_t|$.
- If $e_t > \theta$ → Transmit to Base Station (event occurs).
- If $e_t \leq \theta$ → Do not transmit (saves energy).

Assume:

- Threshold, $\theta = 2.0$ °C
- Prediction model: Moving Average of last 3 values
- Only significant changes (>|2.0 °C|) are reported

Time	x_t (Actual)	$\hat{x}_t$ (Predicted)	$e_t = \lvert x_t - \hat{x}_t \rvert$	Action
t_4	29.0	25.8	3.2	Transmit
t_9	19.5	22.1	2.6	Transmit
t_{12}	31.0	27.5	3.5	Transmit

Other times: $e_t \leq 2.0$ → No Transmission

Figure 3: Threshold-Based Prediction Process

In the top section, the diagram shows the step-by-step operation inside a sensor node. First, the node performs data sensing, measuring the actual value of a parameter, such as temperature. Next, a prediction step is applied, where the node estimates the current value using a simple model based on previous readings. After that, the prediction error is calculated as the difference between the actual and predicted values. This error is then compared with a predefined threshold value. If the error exceeds the threshold, it indicates a significant change in the sensed data, and the node transmits the data to the base station. Otherwise, if the error is less than or equal to the threshold, the node does not transmit, thereby saving energy.

Figure 3 also shows that the model can be updated periodically using recent data to improve prediction accuracy. In the bottom section, a time-series example of temperature monitoring is presented to demonstrate how the approach works in practice. The graph shows two curves: the actual sensed values and the predicted values over time. At each time step, the prediction error is evaluated. When the error is greater than the threshold (marked by upward arrows), data is transmitted to the base station. When the error is within the threshold (indicated by circles), no transmission occurs. A horizontal line represents the threshold value, helping visualize when transmission decisions are triggered. The figure clearly demonstrates that the threshold-driven approach reduces unnecessary transmissions by allowing communication only when significant changes occur in the sensed data. This leads to lower energy consumption**,** reduced network congestion**,** and extended network lifetime, making it highly suitable for event-driven WSN applications.

Machine Learning (ML)-Based Prediction Approaches: Machine learning (ML)-based prediction approaches in Wireless Sensor Networks (WSNs) aim to improve prediction accuracy by learning complex temporal and spatial patterns from historical sensor data. Unlike simple model-based techniques, ML methods can handle non-linear and dynamic environments, making them suitable for applications where sensor readings fluctuate unpredictably. These approaches reduce data transmission by predicting future values locally at sensor nodes or at the base station and transmitting only when significant deviations occur.

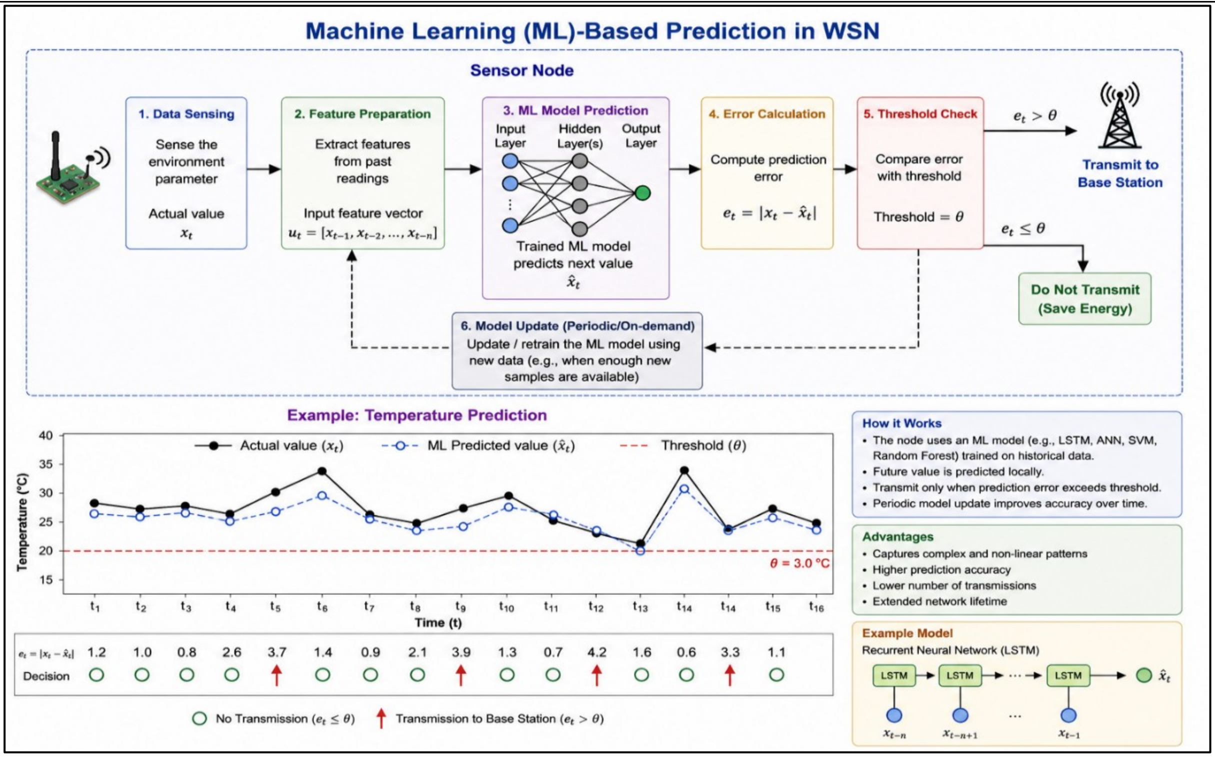
One commonly used ML technique is Artificial Neural Networks (ANNs), which are capable of modeling non-linear relationships in sensor data. ANNs can be trained using historical readings to predict future values with high accuracy, thereby reducing transmission frequency. For example, Haykin [16] provides a comprehensive foundation for neural networks, while Deshpande et al. [17] discuss applications of ANNs in WSN prediction, where neural models are used for efficient data estimation and reduction. Another widely adopted approach is Support Vector Machines (SVMs), which are effective for both regression and classification tasks in WSNs. SVM-based prediction models are particularly useful in handling noisy data and providing accurate predictions with limited training samples.

Vapnik [18] introduced the fundamental theory of SVM, and its application in sensor data prediction has been explored in various studies focusing on energy-efficient communication. In recent years, Deep Learning (DL) techniques, especially Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks, have gained popularity for time-series prediction in WSNs. These models can capture long-term dependencies in sensor data, making them highly effective for forecasting applications. Hochreiter and Schmidhuber [19] introduced LSTM networks, which have since been widely applied in IoT and WSN environments for predictive analytics and data reduction.

Additionally, Decision Tree and Random Forest algorithms are used for lightweight prediction and classification tasks in WSNs. These methods are computationally efficient and can be implemented in edge devices for real-time decision-making. Breiman [20] introduced Random Forests, which are widely used for predictive modeling in distributed systems. Overall, ML-based prediction approaches offer higher accuracy and greater adaptability than traditional methods, enabling a significant reduction in data transmission and energy consumption. However, they require higher computational resources, training data, and memory, which can limit their direct implementation on low-power sensor nodes.

Figure 4 illustrates the working of an ML-based prediction approach in a WSN, showing how intelligent models are used at sensor nodes to reduce unnecessary data transmission. It is divided into two main parts: the sensor node processing flow (top section) and a practical time-series example (bottom section). The node performs **data sensing**, collecting current environmental readings (such as temperature). This is followed by **feature preparation**, where past readings are organized into a feature vector that serves as input to the ML model. Next, a trained ML model (such as ANN, LSTM, or SVM) predicts the current or next value. After prediction, the node calculates the prediction error, which is the difference between the actual and predicted values. This error is then compared with a predefined threshold value. If the error exceeds the threshold, it indicates a significant change in the sensed data, and the node transmits the data to the base station. Otherwise, if the error is within the threshold, the node does not transmit, thereby conserving energy. The diagram also shows a model update step, where the ML model is periodically retrained on new data to improve prediction accuracy over time.

In the bottom section, a time-series example of temperature prediction is presented to demonstrate how the approach works in practice. The graph shows both the actual sensor readings and the ML-predicted values over time. A horizontal line represents the threshold value. At time steps where the difference between the actual and predicted values is large (error exceeding the threshold), data transmission occurs, as indicated by upward arrows. At other time steps, where the prediction is sufficiently accurate, no transmission occurs, as shown by the circular markers. This clearly demonstrates how ML-based prediction reduces the number of transmissions by sending data only when necessary. Additionally, the figure highlights possible ML models, including ANNs, RNNs/LSTMs, Support Vector Regression (SVR), and Random Forests. These models are capable of capturing complex and non-linear patterns in sensor data, leading to higher prediction accuracy compared to simple models. Overall, the figure demonstrates that ML-based prediction enables intelligent, data-driven communication in WSNs by reducing redundant transmissions, lowering energy consumption, and extending network lifetime, although it requires higher computational resources and model training.



How it Works

- The node uses an ML model (e.g., LSTM, ANN, SVM, Random Forest) trained on historical data.
- Future value is predicted locally.
- Transmit only when prediction error exceeds threshold.
- Periodic model update improves accuracy over time.

Advantages

- Captures complex and non-linear patterns
- Higher prediction accuracy
- Lower number of transmissions
- Extended network lifetime

Figure 4: Machine Learning (ML)-Based Prediction

Parameter	**Model-Based Prediction**	**Threshold-Driven Prediction**	**ML-Based Prediction**
Basic Idea	Uses mathematical/statistical models (e.g., linear, AR, moving average)	Uses predefined thresholds to trigger transmission	Uses trained ML models (ANN, SVM, LSTM, RF) for prediction
Decision Criteria	Based on prediction error between actual and predicted values	Based on crossing hard/soft thresholds	Based on the prediction error from the ML model
Complexity	Low	Very Low	High
Computation Requirement	Minimal (suitable for sensor nodes)	Very minimal	High (may require edge/cloud support)
Prediction Accuracy	Moderate	Low (no actual prediction)	High
Adaptability	Limited (fixed models)	Very limited	High (learns patterns dynamically)
Handling Dynamic Data	Poor to moderate	Poor	Very good
Energy Efficiency	Good	Good (event-driven)	Very good
Transmission Reduction	Moderate	Moderate	High
Memory Requirement	Low	Very low	High
Implementation Difficulty	Easy	Very easy	Complex
Example Protocols/Methods	Linear prediction, AR models	TEEN, APTEEN	ANN, LSTM, SVM, Random Forest
Best Suitable Applications	Slowly varying environments (temperature, humidity)	Event-driven applications (fire detection, intrusion)	Complex and dynamic environments (smart cities, IoT analytics)
Advantage	Simple and efficient	Extremely lightweight and fast	Highly accurate and intelligent
Limitation	Cannot handle complex patterns	May miss important data if threshold poorly set	High computational and training cost

Table 1: Comparison of Prediction-Based Communication Approaches in WSNs

IV.CONCLUSION

This survey presented a comprehensive review of **prediction-driven communication models in Wireless Sensor Networks (WSNs)**, focusing on their role in reducing redundant data transmission and improving energy efficiency. Traditional WSN communication approaches rely heavily on periodic data transmission, leading to excessive energy consumption and a shortened network lifetime. To address this limitation, prediction-based techniques have emerged as an effective solution by enabling sensor nodes to transmit data only when significant changes occur in sensed values.

The survey analyzed three major categories of prediction-based approaches: model-based, threshold-driven, and machine learning (ML)-based techniques. Model-based methods offer a simple, computationally efficient solution suitable for resource-constrained nodes, while threshold-driven approaches provide lightweight, event-based communication mechanisms for applications that require quick responses to critical events. On the other hand, ML-based approaches demonstrate superior predictive accuracy and adaptability in dynamic environments, though they entail higher computational and memory requirements.

A comparative analysis of these approaches shows that no single method is universally optimal; instead, each technique involves trade-offs among accuracy, complexity, and energy efficiency. While model-based and threshold-driven techniques are suitable for simpler, energy-constrained scenarios, ML-based methods are more appropriate for complex, data-intensive applications. Furthermore, the integration of prediction techniques with multi-hop routing strategies has the potential to further enhance energy savings and network performance.

References

[1] Heinzelman, W.R., Chandrakasan, A. and Balakrishnan, H., 2000, January. Energy-efficient communication protocol for wireless microsensor networks. In Proceedings of the 33rd annual Hawaii International Conference on System Sciences (pp. 10-pp). IEEE.

[2] Lindsey, S. and Raghavendra, C.S., 2002, March. PEGASIS: Power-efficient gathering in sensor information systems. In Proceedings, IEEE aerospace conference (Vol. 3, pp. 3-3). IEEE.

[3] Manjeshwar, A. and Agrawal, D.P., 2001, April. TEEN: ARouting Protocol for Enhanced Efficiency in Wireless Sensor Networks. In ipdps (Vol. 1, No. 2001, p. 189).

[4] Kulik, J., Heinzelman, W. and Balakrishnan, H., 2002. Negotiation-based protocols for disseminating information in wireless sensor networks. Wireless networks, 8(2), pp.169-185.

[5] Intanagonwiwat, C., Govindan, R. and Estrin, D., 2000, August. Directed diffusion: A scalable and robust communication paradigm for sensor networks. In Proceedings of the 6th annual international conference on Mobile computing and networking (pp. 56-67).

[6] S. Bandyopadhyay and E. J. Coyle, “An energy efficient hierarchical clustering algorithm for wireless sensor networks ”, in Proc. INFOCOM, Mar. 2003, pp. 1713–1723.

[7] I. Matta, G. Smaragdakis, and A. Bestavros, “SEP: a stable election protocol for clustered heterogeneous wireless sensor networks ”, in SANPA, Aug. 2004.

[8] L. Qing, Q. Zhu, and M. Wang, “Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks ”, Computer Commun., vol. 29, pp. 2230–2237, 2006.

[9] A. Bari, A. Jaekel, and S. Bandyopadhyay, “Clustering strategies for improving the lifetime of two-tiered sensor networks ”, Computer Commun., vol. 31, pp. 3451–3459, 2008.

[10] M. Younis, M. Youssef, and K. Arisha, “Energy-aware management for cluster-based sensor networks ”, Computer Networks, vol. 43, pp. 649–668, 2003.

[11] J. S. Liu and C. H. R. Lin, “Energy-efficiency clustering protocol in wireless sensor networks ”, Ad Hoc Networks, vol. 3, no. 3, pp. 371–388, May 2005.

[12] S. Lee, J. Lee, H. Sin, S. Yoo, S. Lee, J. Lee, Y. Lee, and S. Kim, “An energy-efficient distributed unequal clustering protocol for wireless sensor networks ”, in Proc. PWASET, Dec. 2008, pp. 1274–1278.

[13] M. Perillo, Z. Cheng, and W. Heinzelman, “An analysis of strategies for mitigating the sensor network hot spot problem ”, in Proc. Second Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services, 2005, pp. 474–478.

[14] Jain, A., Chang, E.Y. and Wang, Y.F., 2004, June. Adaptive stream resource management using kalman filters. In Proceedings of the 2004 ACM SIGMOD international conference on Management of data (pp. 11-22).

[15] Manjeshwar, A. and Agrawal, D.P., 2002, April. APTEEN: A hybrid protocol for efficient routing and comprehensive information retrieval in wireless sensor networks. In Parallel and distributed processing symposium, international (Vol. 3, pp. 0195b-0195b). IEEE Computer Society.

[16] Haykin, S., 2009. Neural networks and learning machines, 3/E. Pearson education india.

[17] Deshpande, A., Guestrin, C., Madden, S.R., Hellerstein, J.M. and Hong, W., 2004, August. Model-driven data acquisition in sensor networks. In Proceedings of the Thirtieth international conference on Very large data bases-Volume 30 (pp. 588-599).

[18] Vapnik, V.N., 1998. Statistical learning theory wiley-interscience. New York.

[19] Hochreiter, S. and Schmidhuber, J., 1997. Long short-term memory. Neural computation, 9(8), pp.1735-1780.

[20] Breiman, L., 2001. Random forests. Machine learning, 45(1), pp.5-32.