Saziya Tabbassum1* , Chandra Kumar Jha² , Sneha Asopa³

¹ Research Scholar, Department of Computer Science, Banasthali Vidyapith, Rajasthan, India

Email: saziyatabassum@gmail.com

² Professor & Head, Department of Computer Science, Banasthali Vidyapith, Rajasthan, India

³ Assistant Professor, Department of Computer Science, Banasthali Vidyapith, Rajasthan, India

Abstract: The research focuses on hierarchical clustering-based intrusion detection using artificial neural networks (ANNs) for secure data transmission in wireless sensor networks (WSNs). WSNs consist of numerous tiny sensor nodes deployed to monitor environmental phenomena. These nodes face significant challenges, including restricted energy, memory space, communication range, and limited capacity for managing energy, storing, transmitting, and processing data. To address these limitations, a machine learning-based approach is proposed to detect intrusions and efficiently utilize energy by properly selecting cluster heads using a secure clustering protocol. The proposed method was implemented and tested using MATLAB software, employing the NSLKDD and UNSW-NB15 datasets for intrusion detection. The results demonstrated promising outcomes in detecting intruders and enhancing network efficiency, achieving a 92% packet delivery ratio (PDR) and 1.82 Mbps throughput. The study concludes that while WSNs are gaining popularity due to their simplicity, flexibility, and scalability, innovative solutions are necessary for efficient energy management and security. Future research should focus on advanced machine learning models, energy harvesting techniques, scalable protocols, real-time data processing, and integration with IoT platforms for broader applications and enhanced functionality.

Keywords: Wireless Sensor Networks, Intrusion Detection, Artificial Neural Networks, Energy Efficiency, Secure Clustering Protocols

- X -

INTRODUCTION

Wireless Sensor Networks (WSNs) are composed of a large number of sensor nodes deployed to monitor physical or environmental conditions, such as temperature, humidity, and motion, and to cooperatively pass their data through the network to a main location (Akyildiz et al., 2002). These networks are characterized by their ability to self-organize and self-heal, making them highly suitable for deployment in dynamic and remote environments where human intervention is impractical. WSNs are used in various applications ranging from military surveillance to environmental monitoring and smart agriculture.

Sensor nodes are the fundamental units of WSNs, each equipped with a microcontroller, transceiver, power source, and one or more sensors (Heinzelman et al., 2002). These nodes are responsible for sensing data, processing it, and communicating with other nodes or a central base station. The key components of a sensor node include the sensing unit, processing unit, transceiver unit, and power unit. Due to their limited power supply, typically from batteries, energy efficiency is a critical consideration in the design and operation of sensor nodes. WSNs have diverse applications across various domains some of them are illustrated in figure 1. In environmental monitoring, WSNs are used to track climate changes, detect natural disasters, and monitor wildlife habitats (Hart and Martinez, 2006). In healthcare, they facilitate remote patient monitoring, tracking of vital signs, and management of chronic diseases (Lo et al., 2005). Industrial applications include monitoring manufacturing processes, detecting equipment failures, and ensuring workplace safety (Yang, 2014). The flexibility and scalability of WSNs make them

valuable for real-time data collection and analysis in these and other areas.

Figure 1: Applications of WSNs along with its Subcategories

The protocol stack of a WSN is designed to support the unique communication and energy efficiency requirements of sensor networks. It includes physical, data link, network, transport, and application layers, along with power, connection, and task management planes (Al-Karaki and Kamal, 2004). The physical layer deals with frequency selection, carrier frequency generation, signal detection, modulation, and data encryption. The data link layer is responsible for multiplexing data streams, data frame detection, medium access, and error control. The network layer handles routing, while the transport layer ensures reliable data transfer. The application layer provides the necessary software for specific applications such as sensor management protocols (SMP) and query processing. Designing WSNs involves addressing several challenges, including energy efficiency, scalability, fault tolerance, and latency. The limited battery life of sensor nodes necessitates the development of energy-efficient protocols and algorithms (Pantazis et al., 2013). Scalability is crucial to accommodate a large number of nodes and ensure robust communication. Fault tolerance is needed to maintain network functionality despite node failures. Routing protocols in WSNs are typically categorized into flat, hierarchical, and location-based protocols. Flat routing protocols treat all nodes as equals, hierarchical protocols organize nodes into clusters, and locationbased protocols use the physical location of nodes to guide routing decisions (Chen and Zhang, 2005). Security is a paramount concern in WSNs due to their deployment in potentially hostile environments and the critical nature of the data they collect. Common security threats include eavesdropping, node capture, denial of service (DoS) attacks, and data tampering (Perrig and 2004). Ensuring the confidentiality, integrity, and availability of data is essential for the reliable operation of WSNs.

Various mechanisms have been developed to secure WSNs against these threats. Cryptographic techniques are used to protect data confidentiality and integrity.

Key management schemes ensure secure communication between nodes (Eschenauer and Gligor, 2002). Intrusion detection systems (IDS) are employed to detect and respond to malicious activities (Onat and Miri, 2005). Additionally, secure routing protocols are designed to safeguard data transmission from attacks (Karlof and Wagner, 2003). These mechanisms collectively enhance the resilience of WSNs against security breaches.

REVIEW OF LITERATURE

Clustering is a crucial technique in wireless sensor networks (WSNs) aimed at enhancing the scalability and energy efficiency of the network. Research has extensively explored various clustering algorithms, such as LEACH (Low-Energy Adaptive Clustering Hierarchy), HEED (Hybrid Energy-Efficient Distributed clustering), and TEEN (Threshold-sensitive Energy Efficient Sensor Network protocol) (Heinzelman et al., 2001). These algorithms primarily focus on reducing energy consumption by minimizing the distance that data needs to travel, thereby prolonging the lifetime of the sensor nodes. Clustering helps in load balancing and improves the network's robustness by organizing the sensor nodes into clusters, each with a designated cluster head responsible for data aggregation and communication with the base station. Studies have shown that effective clustering can significantly impact network performance, particularly in terms of energy conservation and latency reduction (Abbasi and Younis, 2007).

Energy efficiency is a paramount concern in WSNs due to the limited power resources of sensor nodes. Various methods have been proposed to optimize energy consumption, such as duty cycling, data aggregation, and energy-efficient routing protocols. Duty cycling involves alternating nodes between active and sleep states to conserve energy (Kuliket al., 2002). Data aggregation techniques reduce the volume of data transmitted by combining multiple data packets into one, thus saving energy (Krishnamachar et al., 2002). Energy-efficient routing protocols, such as PEGASIS (Power-Efficient Gathering in Sensor Information Systems) and SPIN (Sensor Protocols for Information via Negotiation), focus on selecting optimal paths for data transmission to minimize energy consumption (Lindsey and Raghavendra, 2002; Kulik et al., 2002). These methods collectively enhance the longevity and efficiency of WSNs, making them more sustainable for long-term applications.

Security in WSNs is critical due to the susceptibility of these networks to various attacks, including eavesdropping, spoofing, and denial of service (DoS). Several security protocols have been developed to counter these threats, such as TinySec, SPINS, and SNEP (Karlof et al., 2002). TinySec is a lightweight security protocol designed for sensor networks, providing confidentiality, authentication, and integrity at the link layer (Karlof, Sastry, & Wagner, 2004). SPINS (Security Protocols

for Sensor Networks) include two secure building blocks: SNEP (Secure Network Encryption Protocol) and μTESLA (Micro Timed, Efficient, Streaming, Losstolerant Authentication), which ensure data confidentiality, two-party data authentication, and broadcast authentication, respectively (Perrig et al., 2002). These protocols have been effective in mitigating various security threats, thus ensuring the secure operation of WSNs.

Secure data transmission is essential to protect sensitive information in WSNs. Protocols such as LEAP (Localized Encryption and Authentication Protocol) and ZigBee are widely used for secure data communication (Zhu et al., 2003). LEAP provides a set of key management protocols that support in-network processing and restrict the impact of node compromise to the immediate vicinity of the compromised node (Zhu et al., 2003). ZigBee is a specification for a suite of high-level communication protocols using low-power digital radios, widely used for secure and reliable data transmission in WSNs (Kinney, 2003). These protocols employ various encryption and authentication mechanisms to ensure the integrity and confidentiality of the transmitted data, making them resilient to attacks.

Intrusion Detection Systems (IDS) in WSNs are designed to identify and respond to unauthorized activities within the network. Various IDS have been proposed, including anomaly-based, signature-based, and hybrid systems (Sun et al., 2004 and Rajasegarar et al., 2006). Anomaly-based IDS detects intrusions by identifying deviations from normal behavior patterns, while signature-based IDS uses predefined patterns to recognize known attacks. Hybrid IDS combines both approaches to enhance detection accuracy and reduce false positives (Rajasegarar et al., 2006). These systems play a crucial role in maintaining the security and reliability of WSNs by providing real-time monitoring and alerting mechanisms.

The literature review highlights significant advancements in clustering techniques, energy efficiency methods, security protocols, secure data transmission, and intrusion detection systems in WSNs. However, several research gaps remain. For instance, there is a need for more efficient and scalable clustering algorithms that can adapt to dynamic network conditions. Additionally, while existing security protocols address many threats, new and evolving attack vectors necessitate continuous improvement and innovation in security mechanisms. Furthermore, integrating energy-efficient techniques with robust security measures without compromising network performance remains a critical challenge. Addressing these gaps will be essential for advancing the state-ofthe-art in WSNs.

RESEARCH METHODOLOGY

The methodology integrates hierarchical clustering and artificial neural networks (ANNs) to enhance intrusion detection and secure data transmission in wireless

sensor networks (WSNs). This approach aims to develop a robust, energy-efficient, and secure WSN.

Wireless sensor networks (WSNs) face significant challenges, including limited energy resources, vulnerability to various security attacks, and the need for effective intrusion detection mechanisms. Traditional intrusion detection systems (IDS) often struggle with high false positive rates and energy inefficiency, compromising the network's performance and longevity. There is a critical need for a solution that not only detects intrusions accurately but also manages energy consumption effectively to prolong the network's operational life. The primary research problem addressed in this study is the development of a hierarchical clustering-based IDS that leverages artificial neural networks to detect and mitigate security threats while ensuring energy efficiency and secure data transmission in WSNs.

The main objectives of this study are as follows:

1. Design a Hierarchical Clustering Algorithm: To create a clustering algorithm that enhances energy efficiency and extends the network's lifetime by optimizing the distribution of sensor nodes and balancing energy consumption.

2. Develop an ANN-Based Intrusion Detection System: To build an artificial neural network model capable of accurately identifying and responding to security threats within the WSN by analysing network traffic patterns and distinguishing between normal and malicious activities.

3. Ensure Secure Data Transmission: To integrate robust encryption and authentication protocols that secure data transmission between sensor nodes and the base station, maintaining data integrity and confidentiality.

4. Evaluate System Performance: To assess the proposed system's performance in terms of energy efficiency, detection accuracy, false positive rate, network lifetime, and data transmission delay, and to compare these results with existing techniques to demonstrate the system's effectiveness and advantages.

Variables under Study

This study involves several key variables critical to evaluating the performance and effectiveness of the proposed intrusion detection system in wireless sensor networks (WSNs). These variables include:

 Energy Consumption: This variable measures the amount of energy used by sensor nodes during data transmission and processing. Energy consumption is crucial as WSN nodes typically operate on limited battery power. The goal is to minimize energy consumption to prolong the network's operational lifetime.

- **Detection Accuracy**: This variable evaluates the ability of the intrusion detection system (IDS) to correctly identify security threats. It is typically measured as the ratio of correctly detected intrusions to the total number of intrusions.
- **False Positive Rate**: This variable measures the rate at which benign activities are incorrectly identified as threats. A lower false positive rate indicates a more reliable IDS.
- **Network Lifetime**: This variable represents the duration for which the WSN remains operational before the nodes' energy is depleted. It is influenced by the energy consumption patterns and the efficiency of the energy management strategies employed.
- **Data Transmission Delay**: This variable measures the time taken for data to be transmitted from sensor nodes to the base station. Minimizing transmission delay is essential for real-time applications.

Dataset Description

1. **NSL-KDD Dataset**: The NSL-KDD dataset is a refined version of the KDD Cup 1999 dataset, widely used for evaluating intrusion detection systems. It provides labeled data on various types of network intrusions and normal traffic, making it suitable for training and testing the ANN-based IDS. Every record of NSL-KDD dataset has 42 attributes among which 41 attributes are as given in table whereas the 42nd attribute of dataset has five different classes categorized as a normal and four attack class.

UNSW-NB15 Dataset: The UNSW-NB15 dataset is another benchmark dataset for network intrusion detection. It contains a diverse range of attack scenarios and normal network traffic data, providing a comprehensive evaluation environment for the IDS. The dataset includes features such as source IP, destination IP, source port, destination port, and protocol type, which are essential for intrusion detection.

Table 2: Categorization of UNSW-NB 15 dataset features

Proposed Methodology

Figure 2: Proposed Methodology Workflow

The proposed methodology involves the following steps:

- 1. **Hierarchical Clustering Algorithm**:
	- o Develop a clustering algorithm that organizes sensor nodes into clusters based on their energy levels and communication costs.
	- o Rotate cluster heads periodically to balance energy consumption across the network.

- o Evaluate the algorithm's performance in terms of energy efficiency and network lifetime.
- 2. **Artificial Neural Network-Based Intrusion Detection System**:
	- o Design an artificial neural network (ANN) model to detect intrusions based on network traffic data.
	- o Train the ANN model using labelled datasets (NSL-KDD and UNSW-NB15) to distinguish between normal and malicious activities.
	- \circ Implement the trained model within the WSN to monitor and analyze network traffic in real-time.

3. **Secure Data Transmission Protocols**:

- Integrate robust encryption and authentication mechanisms to secure data transmission between sensor nodes and the base station.
- o Ensure the integrity and confidentiality of data using protocols like LEAP and TinySec.
- o Evaluate the security performance of the proposed protocols against various attack scenarios.

4. **Performance Evaluation**:

- o Conduct extensive simulations to evaluate the proposed system's performance in terms of energy efficiency, detection accuracy, false positive rate, network lifetime, and data transmission delay.
- o Compare the results with existing techniques to demonstrate the effectiveness and advantages of the proposed system.

FINDINGS AND ANALYSIS

Wireless sensor networks (WSNs) are becoming popular due to their simplicity, flexibility, and scalability, making them suitable for various applications. These networks consist of tiny sensors scattered in an area to monitor environmental phenomena. The sensors communicate with the Internet of Things (IoT) to screen and document environmental data. However, they have significant limitations, including restricted energy (battery), memory space, communication range, and limited capacity for managing energy, storage, transmission, and data processing. To address these limitations, a machine learning-based approach for detecting intrusions and efficiently utilizing energy has been proposed. This involves selecting cluster heads using a secure clustering protocol to optimize energy usage. The proposed method was implemented and

tested using MATLAB software, employing the NSLKDD and UNSW-NB15 datasets for intrusion detection.

Table 3: Simulation Variables for Analysis

Preparing the Network

Initially, an area of 500 \times 500 m² is chosen for deploying the sensor nodes. Around 100 sensor nodes are randomly placed in this area without any specific pattern, and all sensor nodes are static. The figure 3 depicts the node deployment in the monitoring location.

Figure 3: Deployment of Nodes

Once the deployment of sensor nodes is complete, cluster formation begins. A threshold is defined using the consumed energy ratio, and an enhanced random number is generated. Based on these, cluster heads are chosen. These cluster heads then advertise to other sensors that they have been chosen. Depending on the signal strength from the cluster head node, other nodes decide to join the cluster head, forming six different clusters as illustrated in the figure 4.

Figure 4: Cluster Formation

Efficient Energy Utilization

The proposed method's energy efficiency is evaluated using various metrics, and the results are summarized in table 4 and illustrated in the figures below.

Figure 5: Average Residual Energy

Figure 6: Packet Delivery Ratio

Figure 6: Delay of Proposed Work

Figure 7: Throughput of Proposed Work

Securing Network from Intruders

The proposed method's performance in detecting malicious nodes and securing the network from intruders is compared with existing techniques. The performance metrics considered include detection accuracy, false positive rate, and energy consumption.

Figure 8: Comparison of Average Residual Energy

Figure 9: Comparison of Packet Delivery Ratio

Figure 10: Comparison of Packet Loss

Figure 11: Comparison of Throughput

Figure 12: Comparison of Delay

Performance Comparison

The performance of the proposed method in detecting intrusions is compared using the NSLKDD and UNSW-NB15 datasets. The results are summarized in Tables 4 and 5.

Table 4: Comparing Performance Metrics of Proposed Work for Intrusion Detection Using NSLKDD Dataset

Techniques				
Metrics	ANN	RF	KNN	DT
Accuracy	98%	91%	84%	88%
Specificity	97%	85%	71%	94%
Sensitivity	97%	96%	97%	83%
Error	2%	9%	16%	12%
FPR	3%	15%	29%	6%
FNR	3%	4%	3%	17%
FDR	3%	14%	23%	7%
Precision	97%	86%	77%	93%
F1 Score	97%	91%	85%	88%

Table 5: Comparing Performance Metrics of Proposed Work for Intrusion Detection Using UNSW-NB15 Dataset

CONCLUSION AND FUTURE SCOPE

The research on hierarchical clustering-based intrusion detection using artificial neural networks for secured data transmission in wireless sensor networks (WSNs) highlighted significant findings, such as the primary limitations of sensor nodes, including restricted energy, memory space, communication range, and limited capacity for managing energy, storing, transmitting, and data processing. The proposed machine learning-based approach effectively detected intrusions and optimized energy utilization by properly selecting cluster heads using a secure clustering protocol. Tested on MATLAB software with NSLKDD and UNSW-NB15 datasets, the proposed method showed promising results in detecting intruders and enhancing network efficiency, achieving a 92% packet delivery ratio (PDR) and 1.82 Mbps throughput. The study concluded that while WSNs are gaining popularity due to their simplicity, flexibility, and scalability, their inherent limitations necessitate innovative solutions for efficient energy management and security. Future research should focus on enhancing security and energy efficiency, exploring advanced machine learning models, investigating energy harvesting techniques, developing scalable protocols, implementing realtime data processing, and integrating with IoT platforms for broader application and enhanced functionality.

REFERENCES

- 1. Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., & Cayirci, E. (2002). A survey on sensor networks. *IEEE Communications Magazine*, 40(8), 102-114.
- 2. Al-Karaki, J. N., & Kamal, A. E. (2004). Routing techniques in wireless sensor networks: a survey. *IEEE Wireless Communications*, 11(6), 6-28.
- 3. Chen, D., & Zhang, P. K. (2005). Dataintensive applications, challenges, techniques and technologies: A survey on routing in wireless sensor networks. *International Journal of Sensor Networks*, 1(1-2), 104-112.
- 4. Eschenauer, L., & Gligor, V. D. (2002). A keymanagement scheme for distributed sensor networks. In *Proceedings of the 9th ACM conference on Computer and communications security* (pp. 41-47).
- 5. Hart, J. K., & Martinez, K. (2006). Environmental sensor networks: A revolution in the earth system science? *Earth-Science Reviews*, 78(3-4), 177-191.
- 6. Heinzelman, W. B., Chandrakasan, A., & Balakrishnan, H. (2002). An applicationspecific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications*, 1(4), 660-670.
- 7. Karlof, C., & Wagner, D. (2003). Secure routing in wireless sensor networks: Attacks and countermeasures. *Ad Hoc Networks*, 1(2-3), 293-315.
- 8. Lo, B., Wang, Q., & Yang, G. Z. (2005). From research to reality: Wireless body sensor networks for healthcare. *Proceedings of the 2005 IEEE EMBS International Conference on Information Technology Applications in Biomedicine* (pp. 162-165).
- 9. Onat, I., & Miri, A. (2005). An intrusion detection system for wireless sensor networks. In *Proceedings of the 2005 IEEE International Conference on Wireless And Mobile Computing, Networking And Communications* (Vol. 3, pp. 253-259).
- 10. Pantazis, N. A., Nikolidakis, S. A., & Vergados, D. D. (2013). Energy-efficient routing protocols in wireless sensor networks: A survey. *IEEE Communications Surveys & Tutorials*, 15(2), 551-591.
- 11. Perrig, A., Stankovic, J., & Wagner, D. (2004). Security in wireless sensor networks. *Communications of the ACM*, 47(6), 53-57.
- 12. Yang, X. (2014). Wireless sensor networks principles and applications. In *Handbook of Networks in Power Systems I* (pp. 41-77). Springer, Berlin, Heidelberg.
- 13. Abbasi, A. A., & Younis, M. (2007). A survey on clustering algorithms for wireless sensor networks. *Computer Communications*, 30(14- 15), 2826-2841.
- 14. Karlof, C., Sastry, N., & Wagner, D. (2004). TinySec: A link layer security architecture for wireless sensor networks. In *Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems* (pp. 162-175).
- 15. Kinney, P. (2003). ZigBee technology: Wireless control that simply works. *Communications Design Conference*, 2(1), 1-7.
- 16. Kulik, J., Heinzelman, W. R., & Balakrishnan, H. (2002). Negotiation-based protocols for disseminating information in wireless sensor networks. *Wireless Networks*, 8(2-3), 169- 185.
- 17. Lindsey, S., & Raghavendra, C. S. (2002). PEGASIS: Power-efficient gathering in sensor information systems. In *Proceedings, IEEE Aerospace Conference* (Vol. 3, pp. 3-1125).
- 18. Manjeshwar, A., & Agrawal, D. P. (2001). TEEN: A routing protocol for enhanced efficiency in wireless sensor networks. In *Parallel and Distributed Processing Symposium., Proceedings International* (pp. 2009-2015).
- 19. Perrig, A., Szewczyk, R., Tygar, J. D., Wen, V., & Culler, D. E. (2002). SPINS: Security protocols for sensor networks. *Wireless Networks*, 8(5), 521-534.
- 20. Rajasegarar, S., Leckie, C., & Palaniswami, M. (2006). Anomaly detection in wireless sensor networks. *IEEE Wireless Communications*, 15(4), 34-40.
- 21. Sun, B., Osborn, S., & Xu, Y. (2004). Intrusion detection techniques in wireless ad hoc networks. *IEEE Wireless Communications*, 11(5), 56-63.
- 22. Younis, O., & Fahmy, S. (2004). HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. *IEEE Transactions on Mobile Computing*, 3(4), 366-379.
- 23. Zhu, S., Setia, S., & Jajodia, S. (2003). LEAP: Efficient security mechanisms for

large-scale distributed sensor networks. In *Proceedings of the 10th ACM Conference on Computer and Communications Security* (pp. 62-72).

Corresponding Author

Saziya Tabbassum*

Research Scholar, Department of Computer Science, Banasthali Vidyapith, Rajasthan, India

Email: saziyatabassum@gmail.com