Muthanna Journal of Engineering and Technology, Vol. (11), Issue (1), Year (2023)

Muthanna Journal of Engineering and Technology

Website: https://muthjet.mu.edu.iq/

Submitted 14 April 2023, Accepted 22 May 2023, Published online 22 May 2023

A new method for predicting and controlling security risk in sensor networks using big data analysis

[a]Wisam Mohammed Abed and [b]Mohanad Hameed Rashid

[a]Computer Science Preparation and Training Department, Ministry of Education, Anbar Education Directorate, Ramadi, Iraq
[b]Computer Science Quality Department, Ministry of Education, Anbar Education Directorate, Ramadi, Iraq
**Corresponding author E-mail:* rashid_mohanad@yahoo.com

DOI:10.52113/3/eng/mjet/2023-11-01/26-32

Abstract

Sensor network programs that predict and control security have been studied using a huge amount of data. This work views The inadequacy of current security measures is a problem with security standards. To make it simpler to choose security measures that would meet needs, have an impact on genuine security, and maintain track of the network's counting process, the concept of a "network security line" has been proposed. Network security was altered both before and after the event, according to these changes. The security of the topology is then assessed as a counterattack defense utilizing a number of stochastic process-based models. Lastly, by creating a distributed data model, a risk-receiving data model, data feature removal, reconstruction, and an evaluation approach are used to create the experimental risk model. According to the testing findings, when the distance is between 10 and 40 meters, this method is more than 90% successful. The operating power, which is less than the standard requirement, is 196~461 Hz. They perform better in real-time than conventional models. The date timing discrepancy is no more than 0.3 seconds. Despite the improved precision, the relative error is reduced by 3%.

Keywords*: sensor networks, controls security, network security, real time.*

1. Introduction

Microelectronics, computers, wireless communications, and sensors are all examples of technological devices. Each of these technologies has developed swiftly in recent years, and they are all based on the wireless sensor network (WSN), which went from idea to reality quickly. By combining many inexpensive micro sensor nodes with wireless connections, we build a multichip self-organizing network with processing and communication capabilities. It may put a stop to the procedure of collecting, transferring, and integrating various analytical data sets in the joint venture region. Safety is not a key issue for the great majority of non-commercial usages, including preserving the environment, putting out forest fires, and monitoring bird migration. It may also be utilized in other fields, such as data sampling and transmission programs, sensor networks, commercial community wireless security networks, and other specialties. Security-wise, the second application is very weak [1].

A group of identical nodes is referred to as a header. In order to save power, the sensor node's wireless communication module may be switched off and only turned on for brief periods of time. Both the transmitting and receiving states share the same state. Data that is transferred during the data collection process is received by the node's board [2]. A node in a wireless sensor network has the ability to function both as a host and a router (Figure1). In addition to identifying, maintaining, and building network topological structures, they also perform their own unique, specialized tasks for applications (such as data sensing and transmission). As a result, topology control and the routing protocol are the two primary technologies used in wireless sensor networks. As a result of extensive research into Recently, topology control methods, multiple routing protocols, and wireless sensor networks have been presented [3].

They include "location-based routing," which is similar to geographically efficient routing; "query-based routing," which is similar to directed diffusion routing; "energy-aware routing," which is similar to multipath routing; and dependable routing, which is similar to multipath routing (MPR). The cluster-based routing system combines a unique security technique with a traditional topology management method. However, enabling data transmission in wireless sensor networks relies on choosing a reliable control method that meets the specific requirements of It is vital to provide a variety of security assessment models for routing protocols and architectures in wireless sensor networks among various topology control and routing approaches [4]. This article evaluates the protocol, compares it to alternatives, and considers the unique characteristics and quirks of these networks. It discusses its advantages and disadvantages as well as suggestions for creating a secure environment. a certain application environment's routes. This study achieves the overall goal of precise sensor network information transmission risk prediction by accurately predicting risk factors such as data integrity, confidentiality, and availability during network transmission using big data technologies.

Hussain et al [5]. claim that WSN distributes data among technologies and merges electronics, computers, wireless communication, and existing networking technologies into one system. It ushers in a revolution in information understanding and technology purchasing by enabling data gathered and evaluated on sensor nodes allows individuals to receive precise and trustworthy information at any time and in any place. WSN acts as a heat engine in the long term. Like the Internet, this too will have a tremendous influence on people's lives. Communication has changed as a result of WSN and will continue to have an impact on how people interact with events since it integrates historical data from around the world with the present. The potential impact of wireless sensor networks is comparable to that of the Internet. Due to wireless sensor network characteristics, which have a substantial impact on the accuracy and integrity of data transmission, Cheng and others believe that the network is susceptible to several threats. Targeted network routing and architectural patterns prevent precise or efficient data flow towards the last node. Hare and others [6] recommended the use of optical fiber sensors for the security and avoidance of all-optical networks. Using an all-optical network with high transmission bandwidth and processing capacity, they evaluated the network transmission threat value and concealed hazard likelihood. Using a three-scale analytical hierarchy technique, they developed the significance weight of data transmission and achieved network data transmission risk prediction .

The Bayesian network developed by Javeed and colleagues uses the fuzzy significance theory to quantify risk factors in network data transmission. The approach for predicting transmission risk estimates its probability. The network system's total diagnostic reliability is doubled, according to Babaeer and others, and its fault tolerance is boosted since it may include more sensor nodes. When it is difficult to pinpoint the precise monitoring location, the combination of many distinct kinds of sensors increases the efficacy of detection. When the sensors are situated close to the monitoring site, scattered detection monitoring is more efficient than using only one sensor [7]. Zhang and others [8] claim that nodes need to be controlled. As opposed to being application-centric, Due to the data-centric nature of wireless sensor networks, data aggregation, fusion, caching, and compression are essential. Data fusion and compression techniques may result in fewer processing needs, the elimination of duplicate data, and more efficient communication.

According to Dong and others [9] a single sensor has more restricted capabilities than the typical large-scale sensing device, so data fusion may boost measurement accuracy by merging numerous sensors. Additionally, data transmission is minimized even at long distances since the sensor nodes employ scattered data processing. Collaborative computing only transmits user-requested data that has been partially processed when utilized locally.

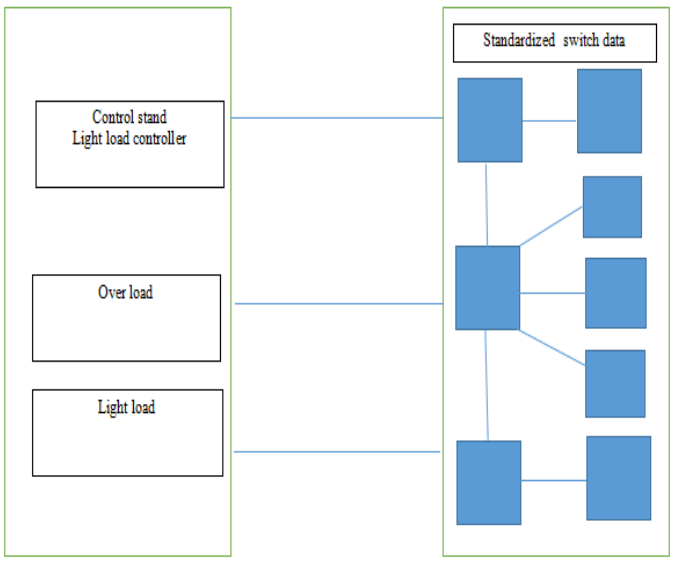

Fig. 1: Network of wireless sensors

2. Research Methodology

The sensor network is considered a large-scale network consisting of two nodes, a routing node and a peripheral node, which is known as peer-to-peer. When the shared mail passes between the two nodes, then it passes through many jumps, which gives the other party the attack more opportunities to destroy the normal transmission of data packets, and attacks are carried out in what is known as attacks. The network layer, and the attacker has a full understanding of the deployment of the wireless network. Therefore, wireless sensor networks in open spaces are common, and the security status of the network will change with the change of the type of attack [10].

2.1. WSN Security Requirements and the Limitations of Current Evaluation Models

Limitations many scientists have long been concerned about the security of conventional transmission. Wireless sensor networks, however, have distinct needs from regular wireless networks [11]. Service management and routing in traditional wireless networks are mainly concerned with optimizing service effectiveness quality of service and high bandwidth utilization. The battery can be recharged for as long as is feasible, and power consumption is merely a small concern. Wireless sensor networks, however, may include thousands of nodes. The majority of nodes stay the same after setup, although some are removable. The characteristics of traffic are similar to those of traditional connections. Furthermore, most sensor networks have unanticipated topologies prior to deployment. Additionally, the network's sensor nodes' functionality and overall architecture have changed repeatedly since the data was transmitted [12]. Based on its characteristics, the security of WSN may be analyzed from the two angles described below:

(a) **Data integrity, authentication, and freshness:** These are only a few examples of communications that may be protected while being sent between nodes. Even if the adversary has the encrypted content, they won't be able to decrypt the original text if a proper symmetric key encryption scheme is utilized. Determine if data is being sent by an authorized node as part of the data identification procedure. "Data integrity" relates to the correctness and absence of data manipulation. This might be accomplished using a similar summary function or data identification [13].

(b) **Node security:** If a sensor node enters the network and is then corrupted into a malicious node, the network may be infiltrated. To stop more hazards, quickly identify various nodes. Nodes in WSN networks are less vulnerable to damage and failure than nodes in traditional networks. Therefore, just offering an encryption method that can secure communications is insufficient. The crucial choice is how to keep things secure. Therefore, if a bomb detonates and private information, such as keys, is stolen, it's crucial to focus on the primary advantages of distribution. An intelligently designed security system should be capable of stopping the hostile behavior brought on by such malevolent acts [14]. Furthermore, different applications call for varying levels of to enhance the efficiency of key component distribution, there must be a minimum amount of variability, which is also necessary in other real-world applications.

2.2. Building a Model for Evaluation Based on Attack Effect

Route security is essential for wireless sensor networks to have secure data transfer. In order to create a secure route and offer networking advice, the security of network routers is assessed using the routing security assessment methodology. Establishing a routing security assessment model based on attack effect [15] entails selecting an evaluation index, outlining a routing attack, calculating node security, and assigning pertinent weights.

2.2.1. Evaluation, Indicator Selection

The creation and application of WSN technology has led to a constant expansion of the meaning of information security. Currently, a variety of performance indicators are routinely used to gauge information security [16]. To assess the impact of route assaults, it is crucial to identify the metrics that can truly and objectively indicate security reliable measurement, no arbitrary standard; simple to gather, It should appear quantitatively rather than qualitatively with labels like "high," "middle," and "low" and utilize at least one measuring unit to describe [17]. It's more suited for active acquisition because of these qualities. There are three stages of security measurement: attack effect layer, security index layer, and protocol layer. Figure (2) illustrates how to assess the direct security difference by comparing the relevance and the impact of each attack on the routing protocol is proportionate [18]. The safety of WSN is closely analyzed using the statistical quantitative approach suggested in this study in order to demonstrate the network's data transmission procedures' security. The individual indices of the security index layer are affected differently by various routing attacks. Assessing them gives the effects of the attack more weight. In this study, a systematic analytical technique is used to continuously strengthen the relationship.

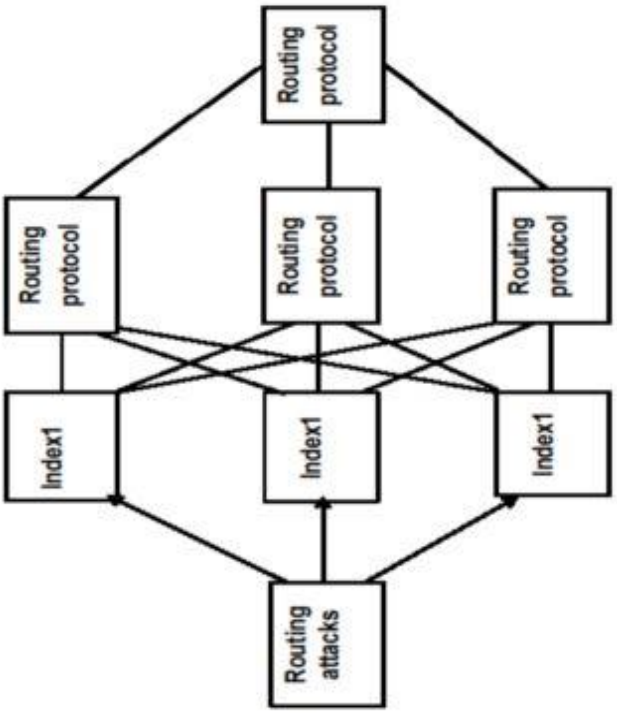

Fig. 2: The hierarchy of safety measurement

2.2.2. The Routing Attack Effect's description

The formula $p(x_i)\ (0 \leq\ p(x_i)\ \leq 1)$ shows how well xi performs the security feature. The safety-related size study of the amount of protection to correctly depict the ambiguity, uncertainty, and disorder of the safety factor, it must have a safety degree comparable to the safety degree. There is a decrease that is directly related to the degree of safety. in ambiguity, uncertainty, and disorder, and vice versa. A WSN's security after and before an attack can be compared to determine how effective the attack was. Based on the information theory idea of "information entropy," which is a good way to represent variations in security performance, the concept of "network security" is offered [19]. This is because only changes in the assessment includes security effectiveness both during and after an assault on the system. attack consequences. A system's state probability may be calculated using the formula $S\ =\ k\ in\ p,$, where $k\ is\ an$ unknown coefficient with a fixed parameter value. $K\ =\ 1$ is used in the calculations shown in this research. Network security Direct is a metric used to quantify information security, the safer the network, the lower the direct number. The direct value of a certain security index xi of WSN is defined as

$$S(x_i) = -\ log_2 P(x_i) \tag{1}$$

According to formula (1), the corresponding safety degree increases as the safety degree $P(xi)$ increases.$S\ (x_i)$ is less than It seems to reason that following a WSN network assault, security would drop and direct value would rise. Consequently, "direct difference" (2) may be used.

$$\Delta s = -\ log_2\ p\left(\frac{(x_i)}{p(x_i)}\right) \tag{2}$$

2.2.3. Determining Node Security

In create a probability model that includes the security evaluation indexes for the attack effect of the WSN routing protocol on the data flow process. The Monte Carlo method is used to develop a process model, and the degree of security index is then used to calculate use the weighted entropy of the index to assess the routing protocol's security in the context of variations [20] in terms of soil moisture The ostensibly Monte Carlo technique entails creating an appropriate random process or probability model based on the statistical rule of randomly occurring issue that has to be addressed, and then doing a sizable number of statistical trials using the example A set of random numbers is used to simulate this process, and the parameters are then determined through an experiment in which the model or process is monitored or sampled. The approximate answer is then provided [21]. A complicated probability computation is all that is involved in its study, which may It may be applied to mimic assaults, normalize security indicators, and assess security. All security-related indications should be recast as data flow and a tried-and-true process paradigm to avoid compromising network security. The failure rate for each index is represented by a random integer between (0 and 1) which represents the normal condition and the failure state, respectively.

$$F(Z) = \begin{cases} 0, z \geq \beta, \\ 1, z < \beta, \end{cases} \quad (3)$$

Indicator $x\,i$ must have a predetermined minimum number of nodes with a single user-defined indication in order to be deemed normal (22). The Monte Carlo technique's guiding principle asserts that the following is true for the probability that index $x\,i$ is normal:

$$P(x_i) \approx \frac{1}{s} \sum_{j=1}^{s} f(X_j), \quad (4)$$

where j stands for the test of the j- th assault.

3. Results And Discussion

The efficiency of this method in high reliability fiber Bragg grating sensor networks for risk prediction in information transmission was evaluated through simulated. A MATLAB simulation environment is used to set up the experiment. The monitoring data acquisition number is launched between $3\ and\ 12\ kHz$, and the clustering center's transmission risk data collection frequency range is set to 2000. (0.3 to 0.6). For the FBG sensor network, the time breadth of the information transmission risk sampling carrier frequency is $1.25\ ms$. The normalized start frequency for risk prediction is $f11 = 0.05$, $f12 = 0.15$, $N = 120$ sample points, and $-12 \sim 12\ dB$ for the fluctuating interference signal-to-noise ratio range. In light of the aforementioned, the danger information from wireless sensor networks is used in this method as well as other algorithms, which are selected for the simulation environment and parameter settings. high-performance fiber the dynamic prediction of information transmission risk was modelled using Bragg grating sensor networks [23,24] (Figure3). High-reliability fib grating sensor networks' information transmission time and risk assessment are also assessed. According to Table (1) this method's prediction accuracy is consistently high and much greater than that of other approaches the table presents the values in Hz, and the results show that the proposed method in this research has a lower load overhead compared to other methods at all distances. As an illustration, at a distance of 15m, the load overhead for this paper's method is 0.14 Hz, while for other methods it is 0.26 Hz.

Table (2) demonstrates how the load overhead increases with predicted distance. This method's load overhead is smaller than earlier ones, coming in at $196 \sim 461\ Hz$ at a distance of $10 \sim 40$ m as opposed to $246 \sim 797\ Hz$. This indicates that the need for a priori sample information can be met at a lower cost. The analysis in Table 2 demonstrates that, at a distance of $10 \sim 40m$, Using this method with a temporal delay of no more than 0.3s, the high-reliability fiber Bragg grating sensor network can forecast information transmission risk. Real-time performance is much superior to $0.25–0.52\ Hz$ for other approaches, and the burden is reduced.

Table 2 presents a comparison of delay and relative error for different distances (m) between the proposed method in this paper and other methods, with values given in seconds (s). The results indicate that the proposed method has a lower delay for all distances compared to the other methods. For instance, at a distance of 24m, the delay for the proposed method in this paper is only 0.26 s, whereas, for other methods, it's significantly higher at 5.04s. The term "196~461 Hz" refers to a frequency range in hertz that is being compared to standard requirements with regard to power consumption, timing accuracy, and real-time performance. It is possible that this frequency range represents the upper limit of frequencies that results in lower power consumption, making it less strict than the standard power requirements, while potentially resulting in improved real-time performance. The meaning of this term can vary depending on the context in which it is being used, and it may be specified in other contexts as well.

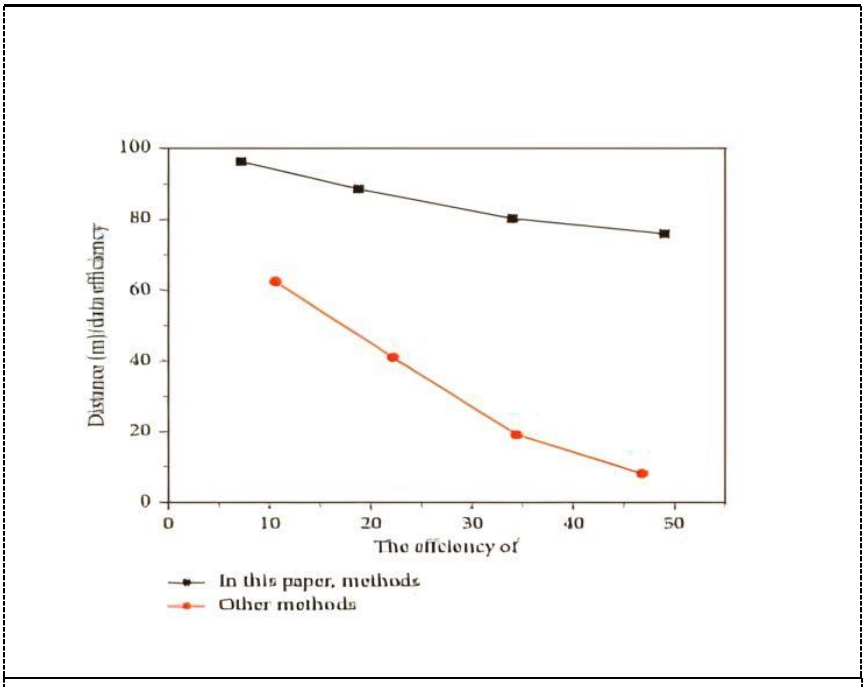

Fig. 3: Prediction efficiency comparison

Table 1. Overhead load comparison (unit:Hz)

Distance (s) /time delay	The method of this paper	Other method
16	0.14	0.26
25	0.18	0.25
41	0.26	0.35
31	0.80	0.42

Table 2. Overhead load comparison (unit:Hz)

Distance (m) /relative error	The method of this paper	Other method
16	1.70	6.73
25	0.26	5.04
41	0.60	6.49
31	0.29	0.84

Table 3. Errors in risk prediction compared (unit:%)

Distance (m) /relative error	The method of this paper	Other method
16	1.70	6.73
25	0.26	5.04
41	0.60	6.49
31	0.29	0.84

Accordingly,

$$error = \frac{x_i - x_j}{x_i} \times 100\% \quad (5)$$

The prediction error of risk information transmission can be calculated using a variety x_i denotes the forecast data and x_j denotes the actual data, where x_i is used throughout the methods. Table (3) displays the results the percentage error of the risk prediction is shown in to increase as the prediction distance approaches $10\sim40$ m. is far less than the relative inaccuracy of the two conventional risk prediction methodologies, which in this study is not more than 3%.

Table 3 presents the comparison of prediction errors for risk with relative error and at different distances (m) between the proposed method in this paper and other methods, with values given in percentage (%). The results suggest that the proposed method has lower prediction errors for risk compared to the other methods at all distances. For example, at a distance of 15m, the prediction error for the proposed method in this paper is only 1.70%, while for other methods, it's as high as 6.73%. The method proposed in this paper also demonstrates lower prediction errors for all other distances when compared to other methods.

4. Conclusion

In this study, the models of WS security evaluation currently in use are analyzed. In order to address After assessing the advantages and disadvantages of various security measures, a model for security assessment based on attack effects is presented. A security assessment index that can accurately and fairly depict security is chosen and made easy. The idea of a "network security line" is introduced and a method for calculating network security moisture is examined. These modifications in network security performance both prior to and during the attack caused these modifications. It employs the correlation feature rearrangement method. By building a large-scale data structure model, high-performance fiber for dispersed data, Bragg grating sensor networks are utilized to lower the risk of information transfer. In order to establish the attribute of dynamic distribution, large data sets are combined using the big data feature clustering technique and the big data fusion clustering analysis. Based on the results of data clustering and information fusion, the dynamic information transmission risk prediction of a high-reliability fiber Bragg grating sensor network is carried out. According to the study, the dynamic prediction of information transmission risk approach maintains a high level of over 90% effectiveness for distances of 10~40 m and has a lower overhead than the traditional method at frequencies of 196~461 Hz. Nowadays, real-time performance is improved, and latency is under 0.3 s. Because the relative inaccuracy is less than 3%, the precision is better and a wide range of applications may be made of it. Security is a major problem with sensor networks. This article only looks at a handful of the security-related concerns that still require in-depth research. No extensive networking testing or other tests are carried out because of the environment's restrictions.

References

[1] R. Geetha, A. K. Suntheya, and G. U. Srikanth, “Cloud integrated iot enabled sensor network security: research issues and solutions,” Wireless Personal Communications, vol. 113, no. 2, pp. 747–771, 2020.

[2] W. Wang, Z. Deng and J. Wang, “Enhancing sensor network security with improved internal hardware design,” Sensors, vol. 19, no. 8, p.1752, 2019.

[3] M. H. Rashid and W. M. Abed, "IoT sensor network data processing using the TWLGA Scheduling Algorithm and the Hadoop Cloud Platform," Wasit Journal of Computer and Mathematics Sciences, vol. 2, no. 1, 2023.

[4] Y. C. Hsu, Y. Zhao, K. H. Huang, Y. T. Wu, and K. L. Tsui, “A novel approach for fall risk prediction using the inertial sensor data from the timed-up-and-go test in a community setting,” IEEE Sensors Journal, vol. 20, no. 99, p. 1, 2020.

[5] M. Hussain, J. Ren, and A. Akram, “Classification of dos attacks in wireless sensor network with artificial neural network,” International Journal on Network Security, vol. 22, no. 3, pp. 542–549, 2020.

[6] J. Hare, S. Gupta, and T. Wettergren, “Pose.3c: prediction based opportunistic sensing using distributed classifification,clustering and control in heterogeneous sensor networks,” IEEE Transactions on Control of Network Systems, vol. 6, no. 99, p. 1, 2019.

[7] H.A. Babaeer and S.A.Al-Ahmadi,“Efficient and securedata transmission and sinkhole detection in a multi-clustering wireless sensor network based on homomorphic encryption and watermarking,” IEEE Access, vol. 8, no. 99, p. 1, 2020.

[8] C. Zhang and Y. Wang, “Sensor network event localization via non-convex non-smooth admm and augmented Lagrangian methods,” IEEE Transactions on Control of Network Systems, vol. 6, no. 99, p. 1, 2019.

[9] R. H. Dong, H. H. Yan, and Q. Y. Zhang, “An intrusion detection model for wireless sensor network based on information gain ratio and bagging algorithm,” International Journal on Network Security, vol. 22, no. 2, pp. 218–230, 2020.

[10] H. Zhang, S. Xing, and J. Wang, “Security and application of wireless sensor network,” Procedia Computer Science, vol. 183, no. 3, pp. 486–492, 2021.

[11] Rashid, M. H. "Examining the Impact of Cyber security on Wireless Sensor Networks." Eurasian Research Bulletin, vol. 19, pp. 58-68, 2023.

[12] Y. Wang, J. Ma, A. Sharma et al., “An exhaustive research on the application of intrusion detection technology in computer network security in sensor networks,” Journal of Sensors, vol. 2021, no. 3, 11 pages, Article ID 5558860, 2021.

[13] C. Chunka, S. Banerjee, and R. S. Goswami, “An efficient user authentication and session key agreement in wireless sensor network using smart card,” Wireless Personal Communications, vol. 117, no. 2, pp. 1361–1385, 2021.

[14] M. F. Moghadam, M. Nikooghadam, M. Jabban, M. Alishahi, and A. Mohajerzadeh, “An efficient authentication and key agreement scheme based on ecdh for wireless sensor network,” IEEE Access, vol. 8, no. 99, p. 1, 2020.

[15] Z. Ma, L. Wang, and W. Zhao, “Block chain-driven trusted data sharing with privacy-protection in iot sensor network,” IEEE Sensors Journal, vol. 21, no. 99, p. 1, 2020.

[16] W. M. Abed, "Deep learning-based Internet of Things intrusion detection," Eurasian Research Bulletin, vol. 19, pp. 47-57, 2023.

[17] B. Lza, A. Jl, and B. Jc, “Security control for t-s fuzzy systems with multi-sensor saturations and distributed event-triggered mechanism—ScienceDirect,” Journal of the Franklin Institute, vol. 357, no. 5, pp. 2851–2867, 2020.

[18] S. Diwakaran, B. Perumal, and K. Vimala Devi, “A cluster prediction model-based data collection for energy efficient wireless sensor network,” The Journal of Supercomputing, vol. 75, no. 6, pp. 3302–3316, 2019.

[19] S. Y. Cui, C. Li, Z. Chen, J. J. Wang, and J. X. Yuan, “Research on risk prediction of dyslipidemia in steel workers based on recurrent neural network and lstm neural network,” IEEE Access, vol. 8, no. 99, p. 1, 2020.

[20] K. Zhu, P. Xun, W. Li, Z. Li, and R. Zhou, “Prediction of passenger flow in urban rail transit based on big data analysis and deep learning,” IEEE Access, vol. 7, no. 99, p. 1, 2019.

[21] M.-w. Fan, C.-c. Ao, and X.-r. Wang, “Comprehensive method of natural gas pipeline efficiency evaluation based on energy and big data analysis,” Energy, vol. 188, no. Dec.1, pp. 116069.1–116069.12, 2019.

[22] Y. Qiu, X. Zhu, and J. Lu, “Fitness monitoring system based on internet of things and big data analysis,” IEEE Access, vol. 9, no. 99, p. 1, 2021.

[23] Y. Yang, “Medical multimedia big data analysis modeling based on dbn algorithm,”IEEE Access,vol.8,no.99,p.1,2020.

[24] Y. Zheng, “Key technologies of media data in-depth analysis system based on artificial intelligence-based big data,” Mobile Information Systems, vol.2021, no.3, Article ID 7191567, 10 pages, 2021., 2021.