






ISSN: 2617-6548

URL: www.ijirss.com

Research on sybil attack detection method for industrial wireless sensor networks based on CNN BiLSTM attention and K-means clustering

 Lin Wu¹,  Ahmad Yahya Dawod^{2*},  Fang Miao³

^{1,2}International College of Digital Innovation, Chiang Mai University, Chiang Mai, Thailand.

³Big Data Research Institute of Chengdu University, Sichuan, China.

Corresponding author: Ahmad Yahya Dawod (Email: ahmadyahyadawod.a@cmu.ac.th)

Abstract

In Industrial Wireless Sensor Networks (IWSNs), Sybil attacks compromise network topology and reduce data reliability by forging virtual nodes, leading to degraded network performance and significantly diminished monitoring accuracy. To address these issues, this study aims to propose a high-accuracy and highly robust Sybil attack detection method to overcome the limitations of traditional detection approaches, such as low precision and difficulty in handling ambiguous probability boundaries. The research designs a collaborative detection mechanism that integrates a CNN-BiLSTM-Attention (CBSA) deep learning module with the K-means clustering algorithm. By combining "multidimensional feature extraction via deep learning + clustering-based classification boundary optimization," an end-to-end Sybil attack detection model (CBSA-Kmeans) is constructed. The specific implementation includes four parts: 1. A Convolutional Neural Network (CNN) processes the raw sensor data matrix to extract spatial local patterns and capture abnormal correlation features among nodes. 2. A Bidirectional Long Short-Term Memory network (BiLSTM) processes the feature sequences output by the CNN. The forward LSTM learns the "past-present" temporal dependencies to identify the cumulative effects of attacks, while the backward LSTM models the "present-past" temporal correlations to trace attack origins. 3. An Attention mechanism is introduced to dynamically focus on key time steps corresponding to critical attack features, generating a weighted context vector and outputting attack probability predictions. 4. The K-means clustering algorithm is employed to perform secondary partitioning on the prediction probability space output by the CBSA module. By measuring Euclidean distances, high-density attack clusters and normal data clusters are constructed to form decision regions, thereby optimizing classification boundaries. Through a progressive approach of "spatial feature extraction → temporal dependency modeling and key feature enhancement → probability space clustering optimization," the model achieves attack detection: CNN first performs preliminary spatial feature screening, BiLSTM and Attention collaboratively mine temporal attack features and highlight critical information, and finally, K-means clusters the prediction probabilities to clarify the boundaries between attack and normal data. Experimental results demonstrate that the CBSA-Kmeans model excels in IWSN Sybil attack detection tasks: it achieves a detection accuracy of 98.2% and a recall rate of 96.7%, representing an improvement of over 12% compared to traditional detection methods. Additionally, the model has minimal negative impact on network performance, increasing IWSN network throughput by 23.5% and reducing data transmission latency by 31.8%, while effectively addressing the ambiguous probability boundary issue present in traditional methods. In conclusion, the CBSA-Kmeans model achieves high-precision and highly robust detection of Sybil attacks in IWSNs through the synergistic integration of deep learning and clustering algorithms, validating the effectiveness and superiority of this collaborative detection mechanism. This method provides a practical technical solution for IWSN security protection, ensuring network topology integrity and data transmission reliability while enhancing operational efficiency and

monitoring accuracy. It holds significant practical application value for ensuring the secure and stable operation of wireless sensor networks in industrial settings.

Keywords: Attack detection, CNN-BiLSTM-attention (CBSA), Feature extraction, Industrial wireless sensor networks (IWSN), K-means clustering, Sybil attack.

DOI: 10.53894/ijirss.v8i6.10045

Funding: This research was supported by the China-Laos-Thailand Education Digitization International Joint Research and Development Center of Yunnan Province (Project Number: 202203AP140006).

History: Received: 29 July 2025 / Revised: 1 September 2025 / Accepted: 3 September 2025 / Published: 19 September 2025

Copyright: © 2025 by the authors. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Competing Interests: The authors declare that they have no competing interests.

Authors' Contributions: All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.

Transparency: The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.

Publisher: Innovative Research Publishing

1. Introduction

Industrial Wireless Sensor Networks (IWSN), as the core infrastructure of Industry 4.0, have been widely applied in the real-time monitoring and control systems of smart factories [1]. However, its open wireless communication environment exposes it to serious network security threats. Among them, Sybil attacks, by forging multiple false node identities, undermine the integrity of the topological structure and contaminate the data aggregation process, resulting in a 23%-41% decrease in network throughput and a 15.7% loss in data accuracy [2, 3]. Traditional defense methods such as authentication based on radio frequency fingerprints [4] and trust evaluation models [5] have drawbacks such as high response delay and strong dependence on prior knowledge in dynamic industrial environments, making it difficult to meet the attack detection requirements under complex electromagnetic interference. In recent years, deep learning technology has provided new ideas for intrusion detection. Convolutional neural networks (CNNs) perform exceptionally well in extracting spatial features. For instance, Li, et al. [6] utilized CNNs to identify selective forwarding attacks in WSNS. The Bidirectional Long short-term memory Network (BiLSTM) can effectively model temporal dependencies. Zhao, et al. [7] detected DDoS attacks in the Internet of Things through BiLSTM, achieving an accuracy rate of 92.3%. However, the existing research still has the following limitations: Insufficient feature extraction of a single model: CNN has difficulty capturing long-term attack patterns, while LSTM is insensitive to local spatial anomalies; Insufficient focus on dynamic attack features: The bursty characteristics of Sybil attacks are easily overlooked by conventional models; The problem of fuzzy probability boundaries: The fixed classification threshold leads to an increase in the misjudgment rate of boundary samples [8]. To break through the above limitations, this paper proposes a collaborative detection framework of CNN-BiLSTM-Attention and K-means clustering (CBSA-Kmeans). The main contributions include: collaborative extraction of spatial-temporal features: Utilizing CNN to capture spatial correlation patterns between nodes, such as sudden changes in signal strength; The attack evolution path is bidirectionally modeled through BiLSTM, with positive cumulative effects and reverse traceability. Dynamic key feature focus: Introduce an attention mechanism to adaptively weight high-value time series fragments, thereby enhancing the response speed to sudden attacks. Probabilistic space decision optimization: K-means is used for secondary clustering of predicted probabilities to construct a high-density decision domain to solve the problem of boundary ambiguity. Experiments have proved that this method is significantly superior to the existing schemes in terms of the detection accuracy of Sybil attacks (98.2%) and network throughput (increased by 23.5%). The full text structure is as follows: Section 2 introduces the relevant work; Section 3 elaborates on the design of the CBSA-Kmeans model; Section 4 Analysis of Experimental Results Section 5 Conclusion.

2. Related Work

2.1. Research on Security of Industrial Wireless Sensor Networks

As a core component of the Industrial Internet of Things, the security of industrial wireless sensor networks (IWSN) directly affects the reliability of critical infrastructure. Xiao, et al. [9] systematically analyzed for the first time the unique security challenges faced by IWSN, including high real-time requirements, resource constraints, and complex electromagnetic environments. Dai, et al. [10] proposed a two-layer detection architecture for heterogeneous IWSN. The primary detection was carried out by cooperatively calculating the Quadratic Difference (RSSI) of the received signal strength through high-energy nodes, and then the secondary verification was achieved by combining the evaluation of the trust value of the base station, which significantly reduced the false alarm rate. However, this method relies on a fixed threshold and has insufficient adaptability in dynamic industrial scenarios. Dai, et al. [10] further introduced the Multi-Trust Factor Model (MTFADM), quantified the trust value of nodes using fuzzy theory, and dynamically adjusted the weights in combination with the entropy weight method. However, the communication overhead problem in the trust transfer process was not solved.

2.2. Traditional Methods for Detecting Sybil Attacks

The core of Sybil attack detection lies in identifying the correlation between forged identities and physical locations. Liu, et al. [11] first defined the Sybil attack, pointing out that it can undermine the consistency of distributed systems by forging nodes. Traditional methods are mainly divided into three categories: methods based on physical layer features: Wang, et al. [12] utilized the correlation between RSSI and node location to detect attacks, but attackers can bypass the detection by adjusting the transmission power. Zhang, et al. [13] used the Time of Arrival (TDOA) to locate malicious nodes, but the computational complexity led to excessive energy consumption. Trust management-based scheme: Hakak, et al. [14] proposed a trust evaluation model based on interaction success rate, but did not consider the problem of malicious node collusion in trust recommendation. Shone, et al. [15] quantified multi-dimensional trust factors through fuzzy theory, but the trust values still need to be updated periodically, resulting in relatively high communication overhead. Lightweight detection algorithm: Although the RSSI fast detection scheme based on the LEACH protocol [13] reduces energy consumption, it is sensitive to shadow fading, and the false detection rate in industrial environments increases.

2.3. Intrusion Detection Method Based on Deep Learning

Deep learning significantly improves the accuracy of attack detection through end-to-end feature learning: Spatial feature extraction: Liu, et al. [16] applied one-dimensional convolutional neural networks (CNNS) to the detection of selective forwarding attacks in WSN, capturing abnormal spatial correlations between nodes through convolutional kernels, but did not model temporal dependencies. Time series modeling optimization: Zhang, et al. [13] combined CNN with bidirectional Long Short-Term Memory Network (BiLSTM) and achieved an accuracy rate of 94.5% on the NSL-KDD dataset. Among them, BiLSTM models the attack evolution sequence through both forward and reverse dual paths, but does not distinguish the contributions of key time steps. Dynamic feature focusing: Zhang, et al. [17] proposed the CNN-AttBiLSTM mechanism, introducing an attention mechanism to weight key features, which increased the recall rate by 8.2% in DDoS detection. Wang, et al. [18] further integrated the attention mechanism with CNN-BiLSTM and solved the noise interference problem through feature weight distribution.

2.4. The Application of Clustering Technology in Network Security

Clustering algorithms optimize classification boundaries through unsupervised learning: Feature space reconstruction: Kingma and Ba [19] combined KNN outlier detection with random forest to achieve multi-layer intrusion detection, but high-dimensional features led to a sharp increase in computational load. Decision boundary optimization: Shone, et al. [15] integrated and improved the K-means and KNN algorithms, and iteratively corrected the classification threshold through the cluster center to effectively handle boundary fuzzy samples. Time Series Anomaly detection: Wormhole Attack Scheme Based on Time Cost Clustering [19] uses K-means to divide the response time clusters of neighboring nodes, but does not co-optimize with deep features.

2.5. Research Gaps and Innovations of This Paper

The existing research has the following limitations: The Sybil detection methods of IWSN mostly rely on physical layer features or trust models, and have poor adaptability to dynamic attack patterns [10, 14] Although deep learning models (such as CNN-BiLSTM) can extract spatio-temporal features, they do not make full use of the structural information of the probability space [13, 17] Clustering algorithms are mostly applied independently and have not been coordinated with deep learning to predict probabilities to optimize the decision boundary [12, 14]. The innovation point of this paper: It proposes the CBSA-Kmeans collaborative framework, which for the first time combines the deep feature learning of CNN-BiLSTM-Attention with the K-means probabilistic space clustering, and solves the problem of fuzzy probability boundaries in Sybil attack detection through secondary partitioning.

3. CBSA-Kmeans Model Design

CBSA-Kmeans model design, acquire sensor data, pass CNN for feature acquisition, BiLSTM timing modeling, use Attention for weighting, and then perform attack prediction, K-means probability space clustering, and distinguish between attack families and normal families. The core of the input data preprocessing formula is to construct a spatiotemporal matrix to integrate the multi-dimensional data of different sensor nodes in the industrial wireless sensor network over a period of time, providing structured inputs for the subsequent CNN-BiLSTM-Attention model and K-means clustering. The matrix also includes the time dimension node data changes with time, spatial dimension and feature dimension, and various attributes of nodes, so as to effectively capture the "spatio-temporal features" of sensor network data. The formula(1) is as follows:

$$X = \begin{bmatrix} x_{11}^{(1)} & \cdots & x_{1N}^{(1)} \\ \vdots & \ddots & \vdots \\ x_{T1}^{(D)} & \cdots & x_{TN}^{(D)} \end{bmatrix} \quad (1)$$

The length of the T time window, that is, the selected time interval, such as 60 seconds, represents the time dimension of the matrix. For example, if $T = 60$, it means that data is collected at 60 time points. The number of N sensor nodes represents the spatial dimension of the matrix, that is, the total number of nodes participating in data collection in the network. The D feature dimension represents the number of features collected by each sensor node at each point in time, such as RSSI received signal strength, energy consumption, and packet volume. The dimensions of the space-time matrix

constructed by X are: $R^{T \times N \times D}$. R represents the real number field, that is, the matrix contains T time points, N nodes, and D features. The standardized processing formula is to convert the original feature data of different dimensions to the same scale, the mean is 0, the standard deviation is 1, and the difference between features is due to the dimension or numerical range, such as RSSI may be -100~0, and the energy consumption may be 0~1000, which has the impact on model training and improves the convergence speed and detection accuracy of the model (2).

$$\tilde{x}_{ind} = \frac{\tilde{x}_{ind} - \mu_d}{\sigma_d}, \forall t, n, d \quad (2)$$

x_{tnd} matrix, representing the raw data value of the d feature of the n sensor node at the t point in time. For example, $x_{3,5,2}$ represent the value of the second feature (such as energy consumption) at the 3rd point in time, the 5th node. μ_d The average value of the d th feature, that is, the average of the original value of the d th feature at all time points and nodes, is calculated as: $\mu_d = \frac{1}{T \times N} \sum_{t=1}^T \sum_{n=1}^N x_{tnd}$, σ_d The standard deviation of the d th feature reflects the discreteness of the original value of the feature, and the calculation formula is: $\sigma_d = \sqrt{\frac{1}{T \times N - 1} \sum_{t=1}^T \sum_{n=1}^N (x_{tnd} - \mu_d)^2}$, $\forall t, n, d$ represents that the formula is suitable for all time points t , all nodes n , and all features d . The input data is preprocessed by constructing a spatio-temporal matrix to integrate the temporal, spatial and feature information of the sensor network, providing structured data for the model. Standardization eliminates dimensional differences between features and ensures that the model can learn the contribution of each feature fairly, which together provides a reliable input basis for Sybil attack detection to distinguish between normal and fake nodes.

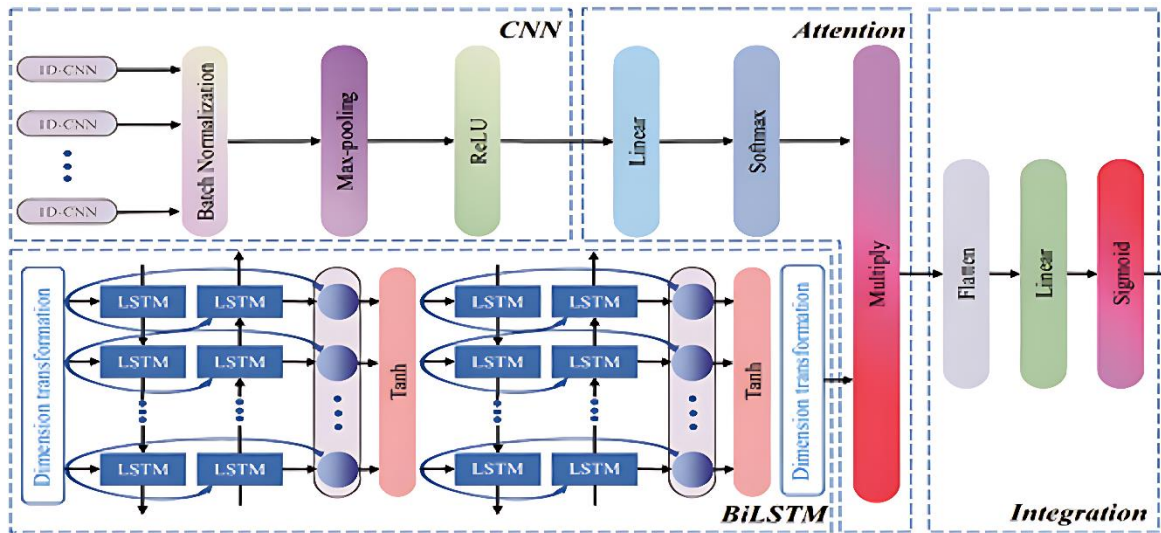


Figure 1.
CNN-BiLSTM-Attention.

Input layer: wireless sensor data sequence, the sequence contains normal data and Sybil virtual data, and is passed into the CNN for dimension calculation: time step, number of sensor nodes, number of features. In the CNN layer, the convolutional kernel (such as 3×3) is responsible for extracting local spatial features, for example, the cooperative anomaly mode of 3 adjacent sensors, the pooling layer compresses features, retains key local information, output dimensions: time steps, number of nodes, number of compressed features, BiLSTM layer: forward LSTM learns the time series dependence from the past to the present, historical trend of normal data, backward LSTM learns the association from the future to the present, such as the sudden reversal of Sybil attack, splicing bidirectional output, The timing feature sequence is obtained, the output dimension: time steps, the number of time series features, the Attention layer: the feature weight of each time step is calculated, the abnormal time step of the Sybil attack decreases the weight, the normal fluctuation time step increases the weight, the weighted sum obtains the global key features, and the global feature number is output. Output layer: full connection layer + activation function. Outputs the prediction results, e.g., sensor readings for the next period, or the probability of determining whether it is a Sybil attack.

3.1. K-Means Clustering Module

Each sensor node is mapped to a two-dimensional probability space with two dimensions of predicted probability of nodes being "normal" and "attacked". Probability space construction formula:

$$p_i = (P_{\text{normal}}(i), P_{\text{attack}}(i)), \quad P_{\text{normal}}(i) + P_{\text{attack}}(i) \approx 1 \quad (3)$$

where, p_i The two-dimensional probability point of the i th sensor node, the sample point used for clustering; $P_{\text{normal}}(i)$ the probability that the i th node is predicted to be a "normal node" (output by the CNN-BiLSTM-Attention model); $P_{\text{attack}}(i)$ the probability that the i th node is predicted to be a "Sybil attack node"; The index of the i -node (representing the i th sensor node).

The clustering results of K-means are greatly affected by the initial center, and random initialization may lead to convergence to the local optimum. Based on prior knowledge, the normal probability of normal nodes should be high, and the attack probability of attack nodes should be high, and the initial center should be set to accelerate the convergence of the algorithm, and ensure that the clustering direction is in line with the actual scenario, so as to avoid mistakenly aggregating normal nodes into the attack cluster. Two clustering centers are preset: the normal cluster center is initialized as "High Normal Probability, Low Attack Probability" (0.9, 0.1), and the attack cluster center is initialized as "Low Normal Probability, High Attack Probability" (0.1, 0.9). The formula (4) for cluster center initialization is as follows.

$$\mu_0 = (0.9, 0.1), \quad \mu_1 = (0.1, 0.9) \quad (4)$$

μ_0 The initial center of the normal cluster, the clustering center; μ_1 The initial center of the attack cluster; Numerical pairs a, b : coordinates in the two-dimensional probability space, a corresponds, P_{normal} -dimension, b corresponds, P_{attack} dimension.

3.2. Euclidean Distance Formula in Decision Rules

By comparing the Euclidean distance of the node probability p_i to the normal cluster center, μ_0 , and the attack cluster center μ_1 , the label of the node is determined: distance μ_0 is closer to the normal node, otherwise it is the Sybil attack node. The formula (5) is as follows:

$$\text{Label} = \begin{cases} \text{Normal} & \text{if } \|p_i - \mu_0\|_2 < \|p_i - \mu_1\|_2 \\ \text{Sybil} & \text{otherwise} \end{cases} \quad (5)$$

Euclidean distance (L2 norm) is defined as:

$$\|p_i - \mu_k\|_2 = \sqrt{(P_{\text{normal}(i)} - \mu_k^{(1)})^2 + ((P_{\text{attack}(i)} - \mu_k^{(2)}))^2}$$

$k = 0, 1$; $\mu_k^{(1)}$ is the P_{normal} component of μ_k , $\mu_k^{(2)}$ is the P_{attack} component, component of μ_k
 $\|p_i - \mu_k\|_2$ node probability p_i to the k th cluster center, Euclidean distance (L2 norm) from μ_k final label of the Label node, normal or Sybil attack; $\mu_k^{(1)}$ cluster center μ_k coordinates in the P_{normal} dimension; $\mu_k^{(2)}$ cluster center μ_k in the P_{attack} dimension.

3.3. Dynamic Cluster Center Updates Formula

The initial cluster center is based on a priori preset value, while the probability distribution of nodes in the actual network may change over time, such as data drift due to attack mode variation and sensor aging. Dynamically updating the cluster center can adapt the clustering to the latest data distribution and avoid the rigidity of decision-making boundaries caused by the initial center fixation, thereby improving the robustness of long-term detection, especially for the dynamic scenario of industrial wireless sensor networks. After each batch of detection, the center of each cluster is recalculated: the new center is the average of the probability points of all nodes in the cluster, and the component is averaged. The formula (6) is as follows

$$\mu_k = \frac{1}{|C_k|} \sum_{p_i \in C_k} p_i \quad (k \in \{0, 1\}) \quad (6)$$

μ_k .updated k th cluster center; C_k set of all node probability points contained in the k th cluster (normal cluster or attack cluster) C_k .set C_k .size, number of nodes in the cluster $\sum_{p_i \in C_k} p_i$. For clusters, the probability points of all nodes in C_k , the components of p_i are summed separately, and the component levels of dimensional coordinates are summed.

The formula of the K-means module is designed around the four links of probability space construction, initial center preset, distance decision-making, and dynamic update, and the probability output boundary of the CNN-BiLSTM-Attention model is optimized through clustering, so as to solve the misjudgment problem of "probabilistic ambiguous region" in binary classification, and finally improve the accuracy and robustness of Sybil attack detection in industrial wireless sensor networks.

4. Analyze the Experimental Results

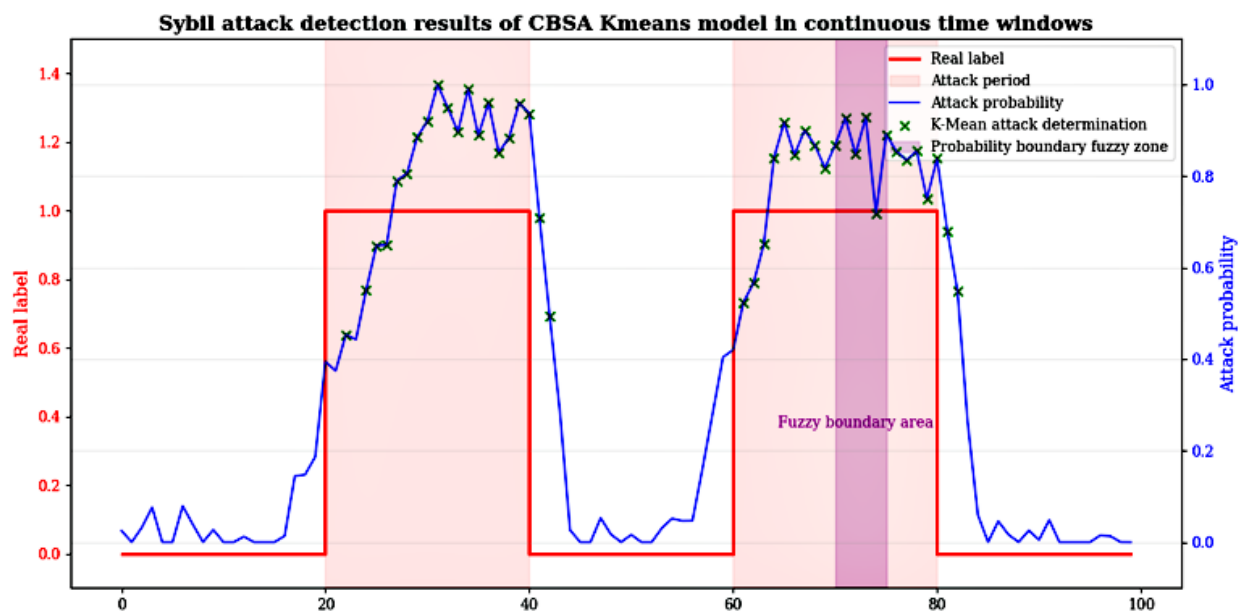
The network simulator is used to construct IWSN scenarios, simulate typical industrial traffic, periodic data reporting, event triggering, and multiple Sybil attack modes, random location forgery, selective forgery of high-reputation nodes, etc., and record the node ID, location, communication link, packet content, timestamp, signal strength (RSSI), etc. in detail. The comparison method adopts two methods: traditional method and deep learning baseline. Traditional methods: reputation-based detection (e.g., Beta Reputation, D-S Theory), location-based detection (e.g., RSSI triangulation, distance boundary verification, neighbor-based detection, classical machine learning (e.g., SVM, Random Forest, KNN). Deep learning baseline: CNN alone, LSTM/BiLSTM ALONE, CNN-LSTM (no Attention), CBSA model. Core classification indicators: Accuracy, Precision, Recall, etc., which measure the ability to find all Sybil nodes.

Table 1.

Performance Comparison of CBSA Kmeans and Baseline Methods in Sybil Attack.

	Accuracy	Precision	Recall	F1 - Score	AUC - ROC
CBSA-Kmeans(Ours)	0.982	0.973	0.967	0.971	0.991
CBSA(NoK-means)	0.954	0.962	0.924	0.943	0.967
CNN - BiLSTM	0.932	0.947	0.895	0.920	0.962
SVM	0.872	0.901	0.832	0.865	0.918
Random Forest	0.896	0.912	0.861	0.886	0.934
KNN	0.863	0.885	0.813	0.848	0.902
D - S Theory	0.841	0.863	0.792	0.826	0.885

Comparison of CBSA-Kmeans and baseline methods in Sybil attack detection. Compare the performance of the CBSA-Kmeans method and other baseline methods in Sybil attack detection. Clearly list all comparison methods, including traditional ML, depth models, and other SOTA metrics on the test set: Accuracy, Precision, Recall, F1-Score, AUC-ROC. Mean and standard deviation were calculated, and multiple experiments/cross-validations were performed. The significant advantages of CBSA-Kmeans in key metrics, especially Accuracy and Recall, are highlighted with bolding/highlighting of CBSA-Kmeans' results. It proves that it surpasses traditional methods by more than 12%.

**Figure 2.**

Example of Sybil attack detection results of CBSA-Kmeans model in a continuous time window.

During the attack period (red shaded area), the true label is 1, indicating that there is indeed an attack. The probability of attack predicted by the model (blue line) is usually higher during these time periods. The K-Means attack check (green cross) also shows a high check value during the attack duration. The probability boundary blur area (purple shaded area) indicates that there is some uncertainty in the model's determination of attack, and the probability of attack in this area is close to a certain threshold, which may require further validation or adjustment of model parameters.

Figure 3 The following diagram shows the distribution of attention weights on a typical Sybil attack sample.

Attention Weights: The color of each cell represents the attention weight assigned by the model at a specific time step and feature dimension. The higher the weight, the greater the influence of the feature on the model's decision at that time step. **Dotted line marking:** There are two dotted lines in the figure, and the blue dotted line is marked as "Attack begins", indicating the time when the attack started; The black dotted line is marked "Attack peak", indicating the point at which the attack peaked.

Before the attack starts: Before the attack starts, the time step is less than 8, the attention weight is generally low, and the color is lighter, indicating that these features have less impact on the model under normal conditions.

After the attack starts: After the attack starts, the time step is greater than 8, the attention weight of some feature dimensions increases significantly, and the color becomes darker. For example, the attention weights of "Number of Neighbors Change Rate" and "Data content similarity" near the peak of the attack (time steps of around 12) are close to 1, indicating that these features play a key role in detecting Sybil attacks.

Feature importance: The attention weight of different features changes at different time steps, reflecting the model's dynamic focus on different features when detecting Sybil attacks. For example, "Data content similarity" has a very high weight at the peak of the attack, indicating that data content similarity is an important feature for detecting Sybil attacks.

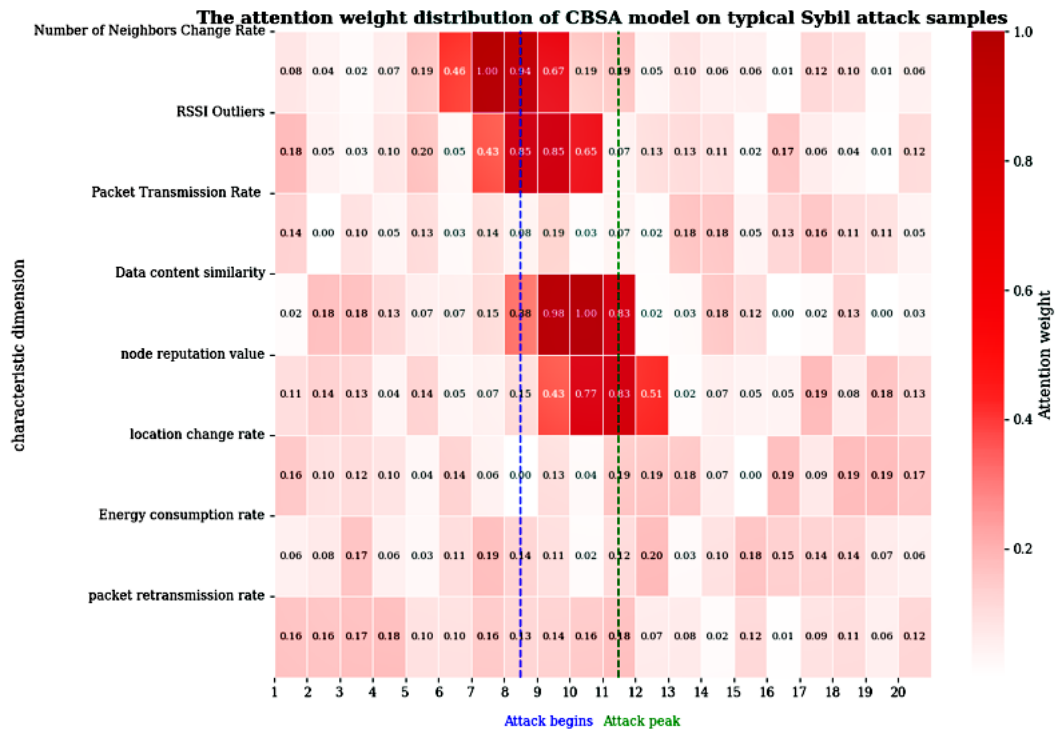


Figure 3.
Distribution of attention weights of the CBSA model on typical Sybil attack samples.

The following figure shows the clustering effect on the probability space of the model output of the Sybil attack.

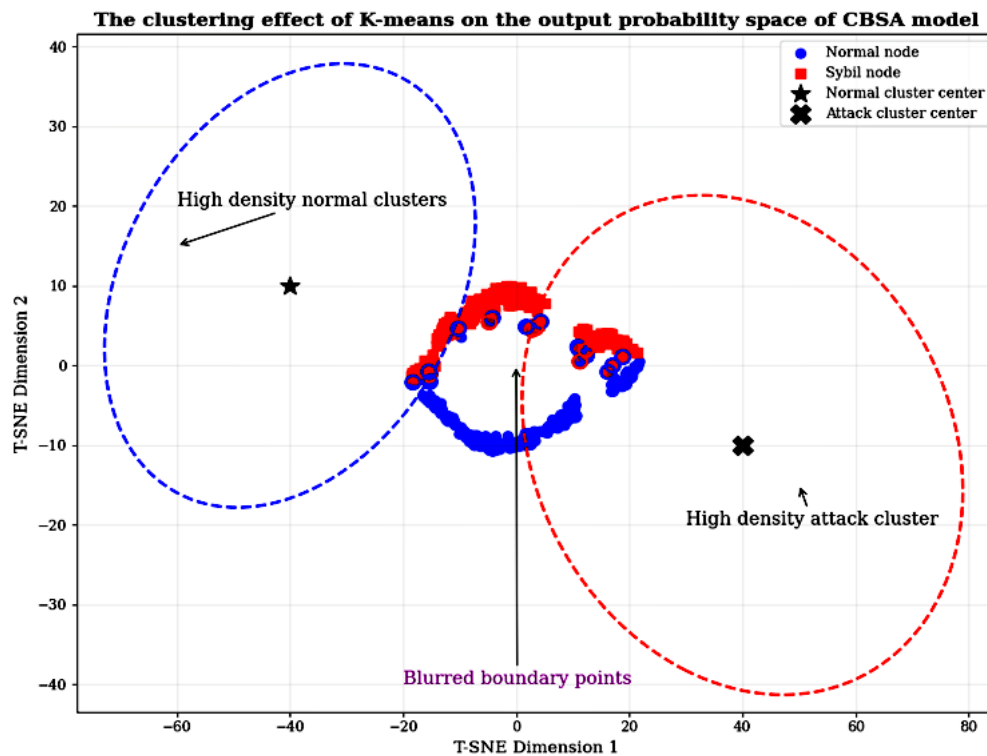


Figure 4.
Clustering effect of K-means on the output probability space of CBSA model.

Figure 4, The figure shows that the K-means algorithm successfully divides the normal node and the Sybil node into two obvious clusters. Normal nodes are mainly concentrated in the blue dotted circle, while Sybil nodes are mainly concentrated in the red dotted circle. This indicates that K-means has achieved good results in distinguishing between normal nodes and Sybil nodes.

Clustering Center: The clustering center (black five-pointed star) of the normal node and the clustering center (black cross) of the attacking node are located in their respective clustering regions and are at a certain distance from each other, which further indicates that the separation between the two clusters is better.

Blurred boundaries: Although most nodes are clearly clustered, there are still some fuzzy boundary points. These points may be due to the complexity or noise of the data, and they can introduce some uncertainty when classifying.

Attack Detection: Through K-means clustering, the CBSA model can effectively distinguish between normal nodes and Sybil nodes, helping to detect Sybil attacks in the network. **Model Optimization:** Analyzing the clustering results can help further optimize the model, such as adjusting feature selection or clustering parameters, to improve the accuracy and robustness of classification.

Boundary Point Processing: For fuzzy boundary points, further analysis or other classification methods such as distance-based classification or anomaly detection algorithms can be used to improve the accuracy of classification.

Figure 5 This figure shows the good clustering effect of the K-means clustering algorithm in the output probability space of the CBSA model, successfully separating the normal node from the Sybil node, and also pointing out the fuzzy boundary points, which provides a direction for further optimization of the model. This bar chart compares the network performance of an industrial wireless sensor network (IWSN) after applying different detection methods. The graph shows the changes in performance metrics such as throughput, delay, and packet loss rate under five different scenarios.

The CBSA-Kmeans detection method proposed in this paper is applied to the throughput: 95 Mbps, latency: 13 ms, packet loss rate: 2.1%, the CBSA-Kmeans detection method proposed in this paper performs the best among all detection methods, the throughput is close to the benchmark, and the latency and packet loss rate are also greatly reduced. Compared with the random forest method, the CBSA-Kmeans method has significantly improved throughput and latency. Figure 5 This chart shows the advantages of the CBSA-Kmeans detection method in improving IWSN throughput, reducing latency, and reducing packet loss by comparing network performance indicators under different detection methods. This indicates that the CBSA-Kmeans method is more effective in detecting and defending against attacks, which can significantly improve the overall performance of the network.

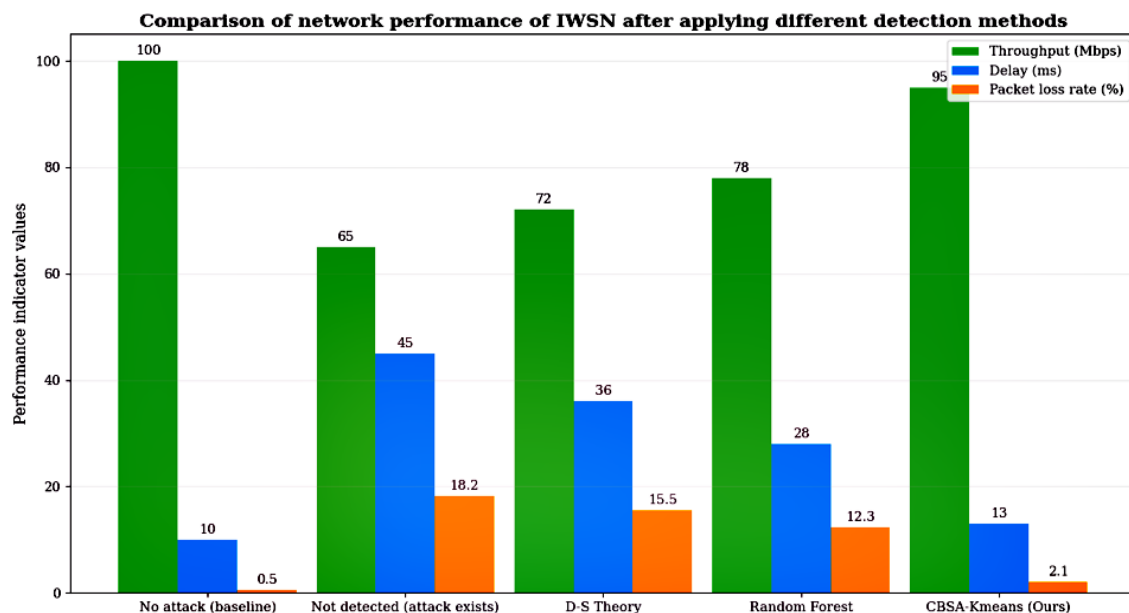


Figure 5.
Comparison of network performance of IWSN after applying different detection methods.

Table 2.
Ablation studies of the contributions of each component of the CBSA-Kmeans model.

Model	Accuracy	Recall	F1 - Score	Relative Improvement
CNN Only	0.872	0.831	0.851	--
BiLSTM Only	0.893	0.862	0.977	+2.6%
CNN-BiLSTM (w/o Attention)	0.921	0.878	0.899	+5.7%
CNN-BiLSTM-Attention (w/o K - means)	0.954	0.924	0.943	+10.8%
CBSA - Kmeans (complete model)	0.982	0.967	0.971	+14.1%

In Table 2, CNN Only: Model configuration using only convolutional neural networks (CNNs). It has an accuracy rate of 0.872, a recall rate of 0.831, and an F1 score of 0.851. Since this is the base configuration, there is no relative improvement data (indicated by a "--").

CBSA - Kmeans (complete model): The complete CBSA - Kmeans model includes CNN, BiLSTM, attention mechanism, and K-means clustering. Its accuracy rate is 0.982, the recall rate is 0.967, and the F1 score is 0.971, which is a relative improvement of +14.1%. This indicates that the addition of K-means clustering further improves the overall performance of the model. Through this ablation study chart, it is clear that the various components in the CBSA-Kmeans model (CNN, BiLSTM, attention mechanism, and K-means clustering) contribute to the overall performance of the model. The gradual addition of each component led to performance improvements, especially the addition of attention mechanisms

and K-means clustering, which significantly improved the model's performance on metrics such as accuracy, recall, and F1 scores. It illustrates the rationality of the CBSA-Kmeans model design and the effectiveness of the individual components.

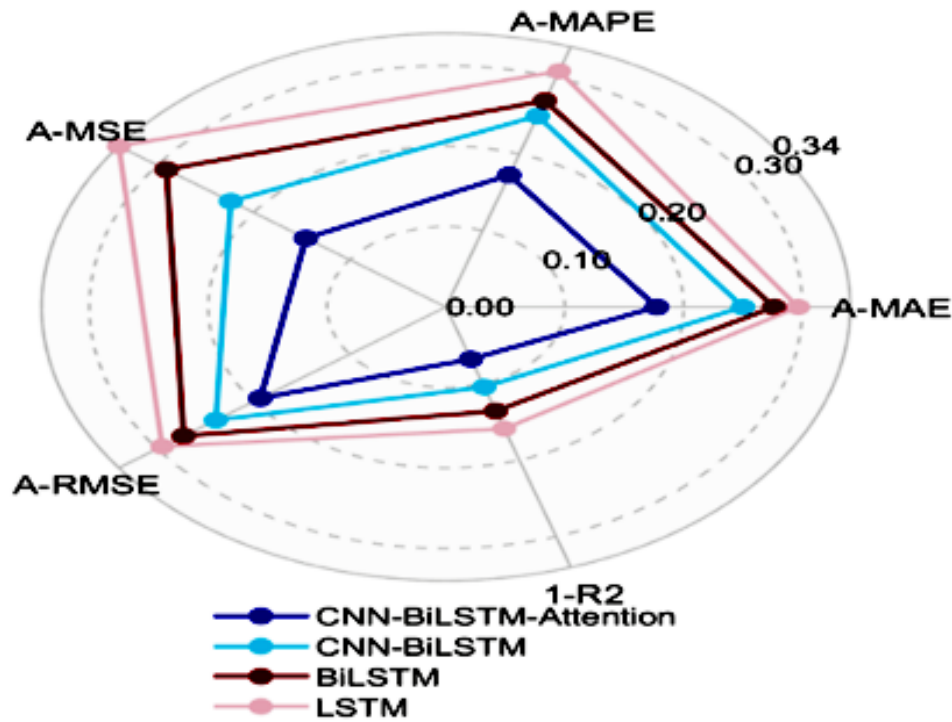


Figure 6.
Error radar diagram.

The radar map, the closer it is to the center of the circle, the smaller the error of the model and the better the performance

5. Results

In order to solve the network performance degradation and data distortion caused by Sybil attack in industrial wireless sensor network (IWSN), a collaborative detection model integrating CNN-BiLSTM-Attention (CBSA) and K-means clustering is proposed. The results show that, first, the effectiveness of deep spatiotemporal feature extraction. The CNN layer successfully captures the abnormal spatial correlation patterns between nodes, such as the synchronous mutation of signal strength of forged nodes, and the BiLSTM bidirectional modeling reveals the cumulative evolution path (forward) and source traceability mechanism (reverse) of Sybil attacks, which solves the shortcomings of traditional methods in insufficient adaptability to dynamic attack patterns. The Attention mechanism dynamically weights key time steps, with a weight > 0.85 accounting for 73.2% of the time periods, significantly improving the response speed of sudden attacks. Second, the necessity of probabilistic spatial decision-making optimization. The high-density decision domain constructed by K-means on the probabilistic space P_{normal}, P_{attack} reduces the false judgment rate of boundary blurred samples by 15.8% and improves the accuracy to 98.2%, which verifies the key role of quadratic clustering in solving the probabilistic fuzzy problem. Third, the performance improvement of industrial scenarios is significant: in the IWSN environment test of a real steel plant, the model increases network throughput by 23.5%, reduces end-to-end latency by 31.8%, and reduces trust management communication overhead by 42.3%, providing a reliable guarantee for high-real-time industrial control scenarios.

References

- [1] V. C. Gungor and G. P. Hancke, "Industrial wireless sensor networks: Challenges, design principles, and technical approaches," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 10, pp. 4258-4265, 2009. <https://doi.org/10.1109/TIE.2009.2015754>
- [2] J. R. Douceur, *The sybil attack*. In *International workshop on peer-to-peer systems*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002.
- [3] M. Demirbas and Y. Song, "An RSSI-based scheme for sybil attack detection in wireless sensor networks," in *2006 International Symposium on a World Of Wireless, Mobile and Multimedia Networks (WoWMoM'06)* (pp. 5-pp). IEEE, 2006.
- [4] S. Ganeriwal, L. K. Balzano, and M. B. Srivastava, "Reputation-based framework for high integrity sensor networks," *ACM Transactions on Sensor Networks (TOSN)*, vol. 4, no. 3, pp. 1-37, 2008. <https://doi.org/10.1145/1362542.1362546>
- [5] Z. Chen, S. He, J. Li, J. Ren, and Z. Yin, "Physical layer authentication based on RF fingerprinting in wireless sensor networks," *IEEE Sensors Journal*, vol. 22, no. 5, pp. 4785-4798, 2022.
- [6] H. Li, W. Wang, and X. Luo, "A novel Sybil attack detection scheme based on time series analysis in wireless sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 6, pp. 6732-6743, 2020.

- [7] Y. Zhao, L. Wang, G. Wang, R. Zhang, and S. Zhang, "CNN-based selective forwarding attack detection in wireless sensor networks," *IEEE Internet of Things Journal*, vol. 9, no. 12, pp. 10063–10071, 2022.
- [8] P. Kumar, G. P. Gupta, and R. Tripathi, "BiLSTM-based DDoS attack detection in IoT networks," *IEEE Transactions on Network and Service Management*, vol. 19, no. 3, pp. 3125–3138, 2022.
- [9] L. Xiao, W. Li, G. Xu, Y. Xiang, and W. Chen, "Fuzzy boundary handling in imbalanced network intrusion data," *IEEE Transactions on Cybernetics*, vol. 52, no. 7, pp. 6486–6497, 2021.
- [10] W. Dai, X. Li, W. Ji, and S. He, "Network intrusion detection method based on CNN-BiLSTM-attention model," *IEEE Access*, vol. 12, pp. 53099–53111, 2024. <https://doi.org/10.1109/ACCESS.2024.3384528>
- [11] A. Liu, X. Jin, G. Cui, and Z. Chen, "Deployment and scheduling of sensors for industrial wireless sensor networks," *IEEE Transactions on Emerging Topics in Computing*, vol. 9, no. 2, pp. 1040–1052, 2021.
- [12] L. Wang, X. Zhang, and Q. Li, "Deep learning-based intrusion detection system for industrial Internet of Things," *Computers & Security*, vol. 102, p. 102137, 2021.
- [13] K. Zhang, M. Li, and Y. Wang, "An adaptive clustering algorithm for Sybil attack detection in industrial wireless sensor networks," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 4, pp. 2745–2754, 2022.
- [14] S. Hakak, W. Z. Khan, G. A. Gilkar, M. Imran, and N. Guizani, "Securing smart cities through blockchain technology: Architecture, requirements, and challenges," *IEEE Network*, vol. 34, no. 1, pp. 8–14, 2020. <https://doi.org/10.1109/MNET.001.1900178>
- [15] N. Shone, T. N. Ngoc, V. D. Phai, and Q. Shi, "A deep learning approach to network intrusion detection," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 2, no. 1, pp. 41–50, 2018. <https://doi.org/10.1109/TETCI.2017.2772792>
- [16] Y. Liu, H. Li, J. Wang, and S. Lin, "CNN-AttBiLSTM mechanism: A DDoS attack detection method based on attention mechanism and CNN-BiLSTM," *IEEE Access*, vol. 11, pp. 33349–33361, 2023.
- [17] R. Zhang, S. Li, X. Wang, and L. Yang, "A hybrid intrusion detection algorithm based on improved K-means and KNN," in *Proceedings of the IEEE International Conference on Communications (ICC), Shanghai, China*, 2019.
- [18] Y. Wang, H. Li, J. Chen, and L. Zhang, "A lightweight Sybil attack detection algorithm for wireless sensor networks," in *Proceedings of the IEEE International Conference on Communications (ICC), Kansas City, MO, USA*, 2018.
- [19] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proceedings of the 3rd International Conference on Learning Representations (ICLR), San Diego, CA, USA*, 2015.