ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 1 - 2026

Digital Twin-based Linkage Study of Intelligent Network Security Mapping and Power Battery Failure Prediction for New Energy Vehicles

Yuyi Chen[1,*]

[1] Jiaxing Vocational & Technical College, Jiaxing, Zhejiang, 314000, China

SUMMARY: *The rise of intelligent network-connected new energy vehicles signifies that the automobile industry is developing in the direction of intelligence, network connectivity and electrification. Based on the digital twin technology to build a multilayer system architecture, the virtual network mapping problem belongs to the combinatorial optimization problem of the optimization problem, the improved chaotic system is applied to the F function in the Feistel encryption framework in order to achieve good obfuscation and diffusion effects, and the cyclic shift control bits are added to the seed key, and a new dynamic key generation method is proposed. For new energy vehicle power battery fault prediction, a fault prediction model is established based on the LS-SVM algorithm, and the correlation calculation of power battery faults is carried out to compare the performance of SVM and LS-SVM methods. The encryption algorithm proposed in this paper has good security, plaintext sensitivity and fast encryption time; the SOC jump alarm, high-voltage interlock alarm, under-voltage alarm of on-board energy storage device, over-voltage alarm of on-board energy storage device have strong positive correlation with other faults, and the results of the Least-Squares Support Vector Machine regression prediction show that there will be a good tracking effect at the four points of time after consecutive prediction, and after the four points of time there will be a good tracking effect with real value offset, realizing new energy vehicle power battery fault prediction.*

KEYWORDS: *intelligent networked vehicle system; power battery; encryption method; battery failure prediction; support vector regression*

1 Introduction

The cyber security situation of new energy vehicle intelligent network connection is grim, with a wide variety of threats and complex mechanisms [1]. Among them, an attack on the local area network bus of the vehicle network controller will disrupt normal communication between electronic control units, which may lead to loss of control or misoperation of the vehicle [2]. Hackers use sensor spoofing to inject false data into the vehicle to confuse the vehicle's decision-making system, posing a potential danger [3]. Examining vehicle-to-circuit communication, V2X communication jamming disrupts the communication link through high-frequency signals, and this disruption leads to the failure of cooperative vehicle control [4]. False command injection, on the other hand, is even more insidious, as it spoofs protocols to transmit forged control commands, and traffic is disrupted and security risks increase dramatically as a result [5]. Cloud services are equally dangerous, with hackers stealing data to leak user privacy and malware injection paralyzing the system [6]. Intelligent networked vehicle network security protection is facing a variety of complex mechanisms and serious

*wy9785625453@163.com
https://doi.org/10.65102/is2026119
Published: 30 April, 2026

challenges.

Battery management system is the “brain” of new energy vehicles, shouldering the battery status monitoring, safety protection, energy optimization management and other important tasks [7]. With the widespread popularization of new energy vehicles, battery management system failures are increasing, seriously affecting the safety and reliability of vehicles. A single fault diagnosis technology is difficult to comprehensively cope with, and it is necessary to comprehensively integrate and apply controller LAN bus data analysis, state estimation, intelligent prediction and other technologies to establish a three-dimensional diagnostic model combining in-vehicle and cloud [8, 9]. It is necessary to revolutionize the maintenance mode and use new technological means such as big data analysis and virtual reality to realize accurate overhaul and predictive maintenance, so that the whole life cycle management of the battery management system can be more efficient and intelligent [10, 11].

Digital twin technology is an advanced technology that realizes bidirectional mapping and interaction between physical entities and virtual models by creating virtual models of physical entities and updating the state of the virtual models in real time using sensor data, historical data, etc [12-14]. The theoretical core of its application in the field of new energy vehicles lies in multi-physical field coupled modeling and real-time data-driven simulation, which accurately reproduces the dynamic behaviors of the vehicle power system, thermal management system, and electrical system [15, 16].

The technology has the following significant features: virtual and real fusion, close integration of the physical and virtual worlds, realizing real-time data interaction and synchronization, so that the virtual model truly reflects the state of the physical entity [17]. Accurate simulation, with the help of high-precision modeling and simulation technology, digital twin technology can accurately simulate the behavior and performance of physical entities, providing a reliable basis for analysis and decision-making [18]. Predictive, based on historical and real-time data, using machine learning, big data analysis and other technologies to predict the future state of physical entities, to detect potential problems and take measures in advance [19]. Can be iteratively optimized, in the virtual environment, digital twin technology can continuously optimize the design and operating parameters of physical entities, and apply the optimization results to physical entities to improve their performance [20]. New energy vehicle fault injection experiments in real vehicles are costly and risky, and it is difficult to cover all complex working conditions and extreme scenarios. Digital twin technology provides a safe and efficient verification environment for fault diagnosis and prediction algorithms by constructing high-fidelity virtual models [21].

The application of digital twin technology in Telematics systems has great potential for its business operation improvement, security enhancement and optimization of user experience [22]. He et al. explored the advantages of vehicle digital twin technology applied to self-driving vehicles, they concluded that vehicle digital twins collect a large amount of private information and exposure to open network environments can create security and privacy issues for vehicles, based on this, they proposed a strategy to avoid such issues with vehicle digital twin technology [23]. Ali et al. proposed an intelligent framework based on digital twin technology for cyber-attack detection and governance in vehicle-grid intelligent physical systems for new energy vehicles, which utilizes a long and short-term memory network to estimate the system state, and then performs cyber-attack detection based on the system state through a deep reinforcement learning network, which is capable of linking cyber-attacks to the tram within 5 seconds [24]. Lv et al. combined convolutional neural networks with support vector regression algorithms and introduced digital twin technology to construct a new security monitoring model for intelligent transportation systems, and the accuracy of the algorithms used for network

security prediction exceeded 90% [25]. Liu et al. established a car networking model combining big data and digital twins for automotive communication security, the learning rate of the model in the test increases rapidly and the loss rate gradually tends to zero, and the introduction of blockchain technology improves the privacy protection of the model for car networking [26]. In addition to this, Almeaibed et al. synthesized and analyzed the safety problem of self-driving vehicles and designed a digital twin framework for vehicle driving, which divided the study of the safety problem into three phases: data collection, data processing, and data analysis, and based on the results of the data analysis the self-driving car was able to choose the appropriate road to avoid a collision [27].

The functions of automotive battery management system cover three major areas: battery status monitoring, energy management and safety protection, and the core advantage of digital twin technology lies in the dynamic adaptability and predictive stability of the battery management system [28, 29]. Bugueno et al. outlined a methodology for the application of digital twin technology in the field of lithium batteries, emphasizing the importance of digital twins in battery management systems (BMS), especially through the combination of new technologies such as cloud computing and the Internet of Things (IoT), where digital twins can enable the active monitoring of batteries [30]. In their study, Semeraro et al. found that digital twin technology in battery energy storage systems helps in fault detection and prediction and enables real-time monitoring of the battery status, which dramatically improves the efficiency of the battery system, and introduces a formal conceptual analysis algorithm to deeply analyze the characteristics of digital twins in the battery system [31]. Yuan et al. accurately calculated the temperature difference between the sensor position and the fault through digital twin technology, summarized the fault diagnosis algorithm based on digital twin technology as well as the method of determining the threshold value of fault diagnosis, and discussed in depth from the level of accuracy, cost and efficiency of the method [32]. Eaty et al. proposed a digital twin framework for battery management in electric vehicles by predicting the battery charge health state in the cloud and then estimating the state of charge in the vehicle, and the proposed framework predicts the battery health state with a mean-square error of 0.022, which satisfies the daily trolley battery management [33].

Renold et al. systematically analyzed the current status of the application of digital twin technology in the prediction of battery health state in new energy vehicles by investigating that new energy vehicle batteries are affected by type, chemical composition, size, temperature, current, voltage, impedance, number of cycles, and driving modes, whereas the digital twin technology can break the nonlinear link between battery parameters and health state for effective prediction [34]. Pooyandeh et al. used digital twin technology for health-like monitoring of lithium-ion batteries in new energy vehicles, where battery charge state prediction is achieved by a long and short-term memory algorithm optimized by multiple trainings and three optimizers, while the digital twin monitors and predicts the operation of the battery in real time [35]. The digital twin technique used by Jafari et al. employs a different algorithm from the above for the effective management of new energy vehicle batteries, where the extreme gradient boosting model and the extended Kalman filter predicted an estimate of the state of the battery, in the estimation of the state of health of the battery by utilizing a learning-based prediction method [36]. Wang et al. advocate the use of multiple battery model modeling approaches, including equivalent circuit models, electrochemical models, and deep learning models, which are fused into a digital twin framework to more fully and accurately characterize the battery [37]. Li et al. designed an intelligent digital twin model for the battery management system of new energy vehicles, which can acquire the battery data during the actual driving of the vehicle for measurement, estimation, prediction, and diagnosis of the battery pack state, in which the prediction accuracy of the model is more than 95%, and it can

diagnose the faults in the battery management system in a timely manner [38].

In order to enhance the intelligence and efficiency of the power battery management system of new energy vehicles, a multilayer system architecture is constructed based on digital twin technology, which consists of a data acquisition layer, a digital twin model layer, an analysis and decision-making layer, and an execution layer. Secondly, a packet encryption algorithm based on Logistic chaotic system applicable to wireless sensor networks is proposed, and a new dynamic key generation method is designed, which is combined with experiments to analyze the randomness, sensitivity and encryption time of the chaotic mapping encryption algorithm, and to validate its effectiveness in the network security guarantee. On the basis of the vector machine model, a predictive analysis model and processing flow for electric vehicle power battery fault prediction based on LS-SVR are established to predict both normal and fault voltages, compare the predictive ability of SVR and LS-SVR, and derive the warning time.

2 New Energy Vehicle Cybersecurity Mapping Power Battery Digital Twin Modeling

2.1 Intelligent Networked Vehicle System

2.1.1 Basic concepts

Intelligent Connected Vehicles (ICVs), as the future of vehicles, fundamentally change the way traditional vehicles are controlled and ICVs are being implemented to achieve SAELevel-5 ICVs.Currently a variety of open applications are loaded into ICVs, and the vehicle support for the open applications requires a large amount of computational resources, as well as the performance of real-time data processing. The smart grid vehicle system model is shown in Fig. 1, where the system consists of three components: a trusted authority (TA), a vehicle with an on-board unit, and a roadside unit.

(1) TA is a trusted authorization authority with certain computing resources. When the TA finds a malicious third party eavesdropping, tampering or falsifying data, the TA can trace its real identity and update the revocation list.

(2) The in-vehicle unit has limited computational resources, mainly hosts the real identity and private key of the vehicle, and defaults that any third party cannot obtain its information.

(3) The roadside unit communicates with the vehicle through a wireless class protocol and with the wired TA, and is an intermediate node between the vehicle and the TA.

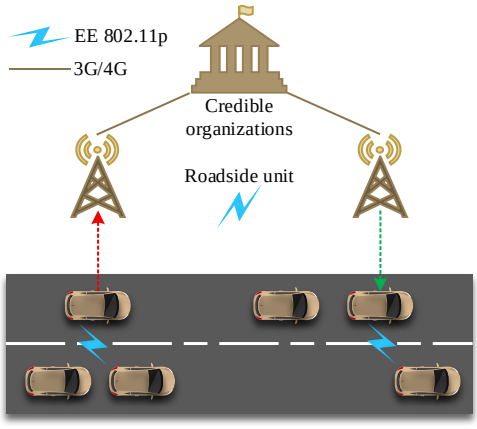


Figure 1: Model of intelligent connected vehicles system

2.1.2 Classification of certification models

Vehicle self-organizing network realizes the safe driving of vehicles in the intelligent transportation system through V2V and V2R communication, and the intelligent networked vehicle system is developed from the vehicle self-organizing network. Vehicles in the system use the vehicle's own on-board sensors to sense the surrounding environment and adjust the direction and speed of vehicle travel in real time, while using wireless communication protocols to realize real-time exchange of information between entities such as vehicle-to-vehicle and vehicle-to-curb unit, but the communication between system entities is vulnerable to eavesdropping or active attacks, passive attacks, adaptive selective messaging attacks and so on by malicious adversaries. The authentication between system entities is mainly categorized as follows:

(1) Vehicle-to-vehicle mutual authentication

The traditional intelligent networked vehicle system is similar to the network model of in-vehicle self-organizing network, and the system mainly consists of TA, roadside units and vehicles. The system vehicles share data with each other in order to reduce redundant computation, but the vehicles of the system are not fully trusted entities.

(2) Mutual authentication between vehicles and roadside units

In the traditional intelligent connected vehicle system based on 4G communication, the roadside unit acts as an intermediate node for communication between the system vehicles and the remote cloud, and within a certain range, the roadside unit vehicle broadcasts relevant information (e.g., traffic light signals at intersections, direction of travel and speed of the surrounding vehicles, etc.) or responds to the vehicle's data requests in a timely manner.

2.2 Power Battery Digital Twin Modeling

A model is a mathematical description of a physical entity, which can realize the reproduction and reconstruction of the properties, characteristics and functions of the physical entity. Models have always been the focus of power battery management research, and the existing power battery models are mainly divided into equivalent circuit models, electrochemical models and data-driven models, all of which can realize the simulation of static and dynamic characteristics of the battery or the performance simulation on the spatial and temporal scales, but all of these three types of models lack of differentiated adaptation to the physical battery and the full-life cycle evolution, such as the electrochemical model has a great dependence on the initial value and parameters; The equivalent circuit model not only relies on a large number of calibration experiments, but also fails to satisfy the full-life cycle parameter adaption and has poor accuracy; the data-driven model has a great dependence on the training set, and has limited generalization performance for scenarios beyond the training set.

The digital twin model absorbs the advantages and shortcomings of the above modeling, and abstracts or extracts the actual operating environment and working conditions of the physical entity by portraying the reaction process or mapping relationship inside the battery, so that the model is in line with the actual scenario of the environment into the virtual reality, so that the model infinitely approximates the physical entity and evolves with the entity “at the same time in the air”, thus realizing the digital simulation of the physical battery. Thus realizing the digital simulation, monitoring, prediction, optimization, etc. of the entity battery.

2.3 System Architecture Design

The architectural design of the power battery management system for new energy vehicles has a more intelligent and efficient management capability with the support of digital twin technology. The system architecture consists of a data acquisition layer, a digital twin model

layer, an analysis and decision-making layer, and an execution layer. The data acquisition layer monitors the key data of the power battery, such as temperature, voltage, current and other parameters, in real time through sensors, IoT and other technologies to ensure the timeliness and accuracy of the data. These data are then transmitted to the digital twin model layer. This layer can accurately simulate the battery's operating state and its health condition in a virtual environment by constructing a virtual model that is consistent with the physical battery. The digital twin model not only dynamically reflects the actual state of the battery, but also predicts future performance changes, providing data support for subsequent optimization and management.

3 Secure encryption of virtual network mapping for smart networks

3.1 Mathematical modeling of virtual network mapping

(1) Description of virtual network mapping

Virtual network mapping refers to mapping a virtual network $G^j\left(N^j,E^j\right)$ onto a subgraph of a physical network $G^j\left(N^j,E^j\right)$ that satisfies the constraints that each physical node can only be mapped by at most one virtual node, and that each physical path (in case of an indivisible path) can only accept one virtual link construction request.) can only accept construction requests for one virtual link, and other constraints. In addition, other constraints such as $D_j^v, c\left(n_j^i\right), b\left(e_j^i\right)$ must be satisfied.

(2) The set of physical nodes that a virtual node can map

For each virtual node $n_j^i\left(i\epsilon\left[1,\left|N^j\right|\right]\right)$ of the jth virtual network $G^j\left(N^j,E^j\right)$ define the set $\Omega\left(n_j^i\right), \Omega\left(n_j^i\right)$ denotes the set of all physical nodes that satisfy the requirement of accepting the mapping of virtual nodes n_j^i. The $\Omega\left(n_j^i\right)$ has to satisfy the following two conditions: 1) the distance between the physical node and the virtual node n_j^i is less than or equal to D_j^v, the distance refers to the actual Euclidean distance, latency, etc.; and 2) the physical node's remaining CPU capacity is sufficient to support the virtual node's n_j^i CPU resource requirement of the virtual node n_j^v.

(3) Residual network

The network remaining after physical network $G^0\left(N^0,E^0\right)$ has completed the mapping of virtual nodes and links is referred to as the remaining network $G_{res}^0\left(N^0,E^0\right)$ of G^0, where the virtual nodes may be from different virtual networks and the virtual links may be from different virtual networks. The remaining node CPU capacity of the ith physical node in the remaining network G_{res}^0 is denoted as $r_n\left(n_0^i\right)$, and the remaining bandwidth of the jth physical link is denoted as $r_e\left(e_0^j\right)$.

(4) Physical network augmentation map

For the jth virtual network $G^j\left(N^j,E^j\right)$, its corresponding physical network augmentation graph is based on the remaining network of the physical network $G_{res}^0\left(N^0,E^0\right)$, for each virtual

node $n_j^i\left(i\epsilon\left[1,\left|N^j\right|\right]\right)$ by adding a corresponding meta-node $\mu\left(n_j^i\right)$ to $\mu\left(n_j^i\right)$ and to all meta-edges with infinite bandwidth are added between the physical nodes belonging to $\Omega\left(n_j^i\right)$, and the CPU capacity of the meta-node $\mu\left(n_j^i\right)$ is equal to $c\left(n_j^i\right)$. The physical network augmentation graph of the jth virtual network $G^j\left(N^j,E^j\right)$ can be represented as $G^{0^{r,}}\left(N^{0^{r,}},E^{0^{r,}}\right)$ where $N^{0^{r,}}=N^0\bigcup\left\{\mu\left(n_j^i\right)\middle|\, n_j^i\in N^j\right\}$; $E^{0^{r,}}=E^0\cup\left\{\mu\left(n_j^i\right),n_0^x\middle|\, n_j^i\in N^j,n_0^x\in\Omega\left(n_j^i\right)\right\}$.

(5) Objective function

When the jth virtual network mapping is performed, its optimization objective is to maximize the virtual network mapping gain with the highest cost-effectiveness, that is, to minimize the virtual network mapping cost under the premise of trying to obtain the jth virtual network mapping gain (completing the construction of the jth virtual network). Since the network structure of the jth virtual network has already been determined when the request is issued, its mapping gain is also determined, so the specific optimization objective is to minimize the equilibrium cost Eq. (1), and Eq. (2) is the range of values of the weights.

$$cost\left(i^{\#}\right)=\sum_{\omega\in N^0}^{\left|N^0\right|}\left(\frac{\beta_\omega}{r_n(\omega)+\sigma}*c(m)\right)+\sum_{(u,v)\in E^0}^{\left|E^0\right|}\left(\frac{\alpha_{u,v}}{r_e(u,v)+\sigma}*\sum_{i\in\left[1,\left|E^j\right|\right]}f_{u,v}^{j,i^*}\right) \tag{1}$$

$$\frac{r_e(u,v)}{\left|E^0\right|}\le\alpha_{u,v}\le r_e(u,v),\frac{r_n(\omega)}{\left|N^0\right|}\le\beta_\omega\le r_n(\omega) \tag{2}$$

(6) Constraints

$$\sum_{i\in\left[1,\left|E^j\right|\right]}\left[\left(x_{u,v}^{j,i}+x_{v,u}^{j,i}\right)*b\left(e_j^i\right)\right]\le r_e(u,v),\forall u,v\in N^{0^{r,}} \tag{3}$$

$$r_n(\omega)\ge\sum_{i\in\left[1,\left|E^j\right|\right]}\left[\frac{\left(x_{m,\omega}^{j,i}+x_{\omega,m}^{j,i}\right)}{\left|E_m^j\right|}*c(m)\right],\forall m\in\frac{N^{0^{r,}}}{N^0},\forall\omega\in N^0 \tag{4}$$

Constraint set (3) ensures that the sum of the bandwidths of all virtual links of the jth virtual network to be mapped on any physical link is less than or equal to the remaining bandwidth of that physical link, thereby ensuring that there are sufficient bandwidth resources on the physical paths mapped by the virtual links to be used for providing the mapping service. The constraint set (4) ensures that the remaining CPU resources of the physical node are greater than or equal to the CPU capacity requirement of the jth virtual network to be mapped.

3.2 Encryption Algorithm Design

3.2.1 Logistic Mapping Characterization and Improvement

Chaotic systems are highly sensitive to initial conditions, a property that can be mapped to confusion and diffusion in cryptography. Also, chaotic systems have complex and deterministic chaotic behavior. Before applying chaotic mapping as a security component for wireless sensor networks, its principles and possible improvements need to be discussed. In this chapter, Logistic mapping is chosen as the object of study because Logistic mapping is the most commonly used and simply defined chaotic system in cryptographic algorithms. It is defined

as:

$$x_{n+1} = \mu x_n \left(1 - x_n\right), \mu \in (0,4) x_n \in [0,1] \tag{5}$$

where μ is the system parameter and x_n is the iteration sequence. When $\mu \in (3.57,4)$, the mapping exhibits chaotic nature.

However, the embedded systems of most wireless microsensor network nodes have relatively low precision and do not have units to handle floating point operations. The study of discrete chaotic systems will help the application of chaotic mapping in wireless sensor nodes. The discrete chaotic Logistic mapping equation is defined as equation (6):

$$z_{n+1} = 4z_n - \frac{2}{m} z_n^2 \tag{6}$$

where Z_n is in the range $[0,2m], m = 2^{L-1}$ (L represents the word length of the computer). Then Z_n can represent all unsigned integers of the computer word length, i.e., $Z_n \in [0,2m]$. When the initial value of Eq. (6) is 0 or $2m$, the value of all iteration results will be 0, and then the weak key problem occurs. To solve the problem, Eq. (6) needs to be further improved as:

$$z_{n+1} = 4z_n - \frac{2}{m} z_n^2 - 1, \left(z_n = 0 or 2m\right) \tag{7}$$

3.2.2 Feistel Cryptographic Framework

Symmetric encryption algorithm has the advantages of low computation, fast encryption speed and high encryption efficiency, which can well meet the needs of a lightweight encryption algorithm, and can fit the resource-constrained situation and practical needs of the sensor network, so in this paper, we choose to design the symmetric encryption algorithm applicable to the sensor network. There are three types of network structures that are often used in designing symmetric packet encryption algorithms, which are Feistel network structure, variant Feistel network structure and SP network structure. In this chapter, the traditional Feistel structure with improved discrete Logistic system is used to design the packet encryption algorithm, and the improved discrete Logistic system is utilized to design the F-function for confusion and diffusion, supported by the framework of Feistel structure.

The Feistel structure is shown in Fig. 2, where the wheel function F acts on the plaintext block several times for the purpose of obfuscation and diffusion, and outputs the ciphertext block at the end. The encryption process of the algorithm can be described as (8), where L_i, R_i represent the left and right halves of the plaintext, respectively, and K_i is the i th round key.The F-function does not need to be invertible, but it usually has to be nonlinear. The main purpose of the F function is to maximize the avalanche effect between plaintext and ciphertext. One of the problems to be addressed in this paper is how to quickly make the system obfuscate and diffuse to the desired effect.

$$\begin{cases} R_{i+1} = L_i \oplus F\left(K_i, R_i\right) \\ L_{i+1} = R_i \end{cases} \tag{8}$$

Since the basic processing unit of a sensor node is usually 8 bits, the F function designed

in this paper also uses 8bits as the smallest processing unit. In addition, in order to increase the chaos and diffusion speed of the algorithm, a modified chaotic system is used in the F function.

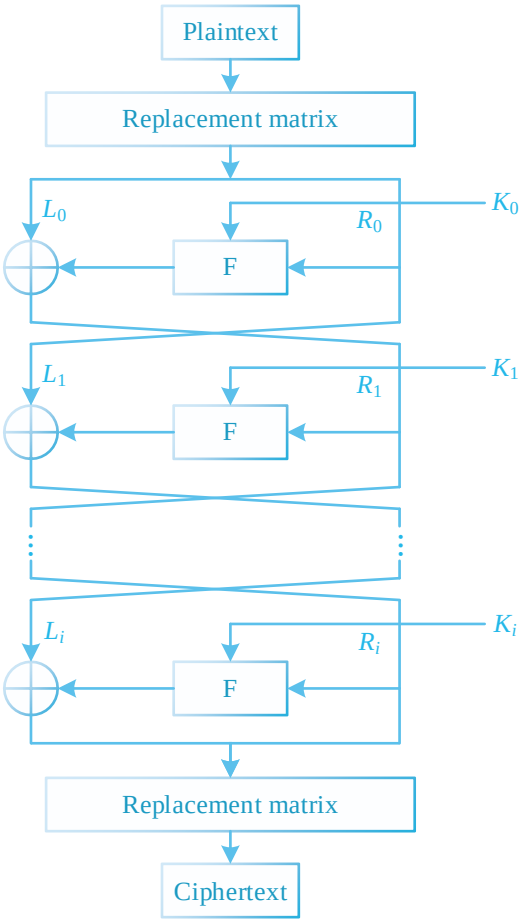


Figure 2: Feistel encryption structure

3.2.3 Seeded Keys and Key Extension Algorithms

The three components of the seed key are given in Figure 3, where the first 4 bits represent the cyclic control bits of the key expansion algorithm. The middle 16 bits are obtained by processing the plaintext after hashing operation, being used as the initial value of the parameter A in the wheel function F. The last 32 bits are generated by the chaotic system, which is extended by the key extension algorithm, and the wheel key for encryption can be obtained. When designing the encryption algorithm, the security of the key is very important, and the complexity of the seed key and encryption key needs to be extended as much as possible to ensure the security of the system. Dynamic key can further improve the security characteristics of the algorithm, and two common ways of applying dynamic key are adding perturbation method and ciphertext feedback method. However, these two methods require additional operation steps, thus this paper provides a new dynamic key generation method, which adds a cyclic shift control bit to the seed key, so that different round keys can be generated when the seed key passes through the key expansion algorithm.

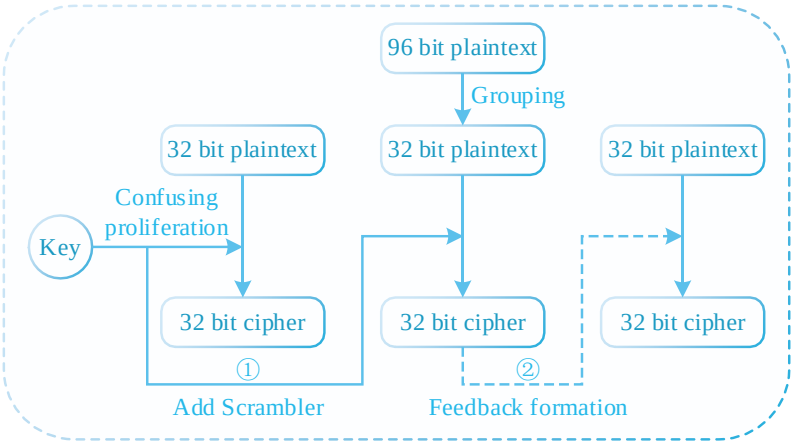


Figure 3: illustrates two dynamic key generation methods

3.2.4 Decryption steps

When a 32-bit plaintext is entered, a bit-level substitution operation needs to be performed first. The relationship between the plaintext and the ciphertext can be damaged to some extent by the substitution operation. The substitution matrix is randomly generated and the matrix must be invertible to enable the decryption process. The inverse matrix $\mathbf{P}^{-1}$ is used for substitution during decryption. The matrix $\mathbf{P}$ and its inverse matrix $\mathbf{P}^{-1}$ used in this paper are designed as follows:

$$\mathbf{P} = \begin{bmatrix} 10 & 7 & 12 & 9 & 16 & 13 & 3 & 14 \\ 11 & 5 & 1 & 15 & 4 & 6 & 8 & 2 \\ 26 & 23 & 28 & 25 & 32 & 29 & 19 & 30 \\ 27 & 21 & 17 & 31 & 20 & 22 & 24 & 18 \end{bmatrix} \tag{9}$$

$$\mathbf{P}^{-1} = \begin{bmatrix} 11 & 16 & 7 & 13 & 10 & 14 & 2 & 15 \\ 4 & 1 & 9 & 3 & 6 & 8 & 12 & 5 \\ 37 & 32 & 23 & 29 & 26 & 30 & 18 & 31 \\ 20 & 17 & 25 & 19 & 22 & 24 & 28 & 21 \end{bmatrix} \tag{10}$$

In Feistel structure, the encryption and decryption keys are the same, but the steps of encryption and decryption are reversed.The Feistel structure can be encrypted repeatedly in multiple rounds, which makes the Feistel structure provides strong security and makes it an irreplaceable encryption method of other current cryptographic algorithms.

This is because the operations used during encryption and decryption are the same. This feature allows the Feistel structure algorithm to recover from the ciphertext to the original plaintext at any time, which is very important for the person who receives the confidential information after decryption.The encryption of the Feistel structure is very flexible, and it can be used to improve the security by using multiple rounds of operations, each round of operations can be used with a different key, and at the same time, by adjusting the number of rounds to increase or decrease the number of rounds to change the algorithm's complexity, so as to adapt to the different needs of the application scenarios. The complexity of the algorithm can be changed by increasing or decreasing the number of rounds so that it can be adapted to different application scenarios. In the algorithm, we can reuse the registers in the process of code

realization to reduce the consumption of computing resources, and the basic encryption method of Feistel structure is as follows:

$$\begin{cases} L_i = R_{i+1} \oplus F\left(K_i, R_i\right) \\ R_i = L_{i+1} \end{cases} \tag{11}$$

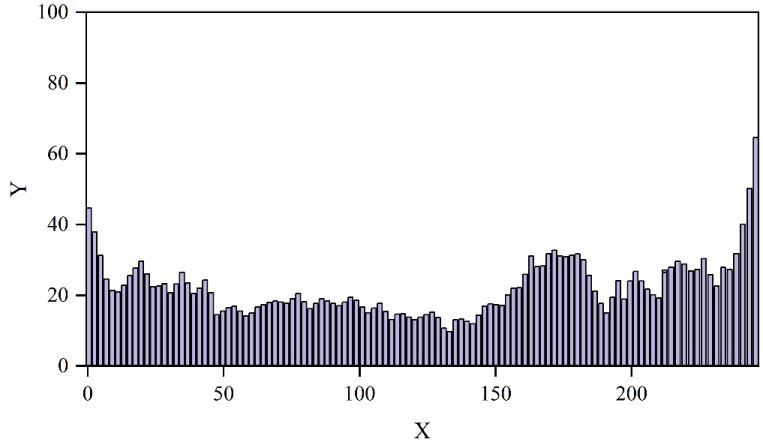
3.2.5 Message encryption algorithms for two-way authentication of identity

In the process of transmitting data, the transmission links applied are consistent, and the encryption requirements of the network nodes are applied to the transmission links, and it is necessary to apply encryption technology to ensure the security of the header and routers and other key equipment, which plays an important role in the subsequent stabilization of the output information. Setting up a new key management mechanism, in the process of assembling for the key, the composition of the corresponding random key, the sensitivity of the mechanism to the parameter a is accurate to the last 11 decimal places or so, the corresponding parameter in the 12th decimal place or so. The cryptographic connection device coincides with the same node, and the corresponding data encryption techniques allow the ciphertext in the mechanism to achieve a constant update during the reset process. These encryption means have different forms of transmission, and because of the special requirements for header and routing information, the granularity of encryption is defined, i.e., a single key can only encrypt plaintext data of a certain length, and the key must be updated when the length exceeds the granularity. After negotiating the granularity of encryption, the keys required for encryption and authentication must be calculated. The authentication information is stored in the client's USB Key, the chaos equations and their parameters are stored in the server, and the user only needs to memorize his authentication identification Uid. If a USB Key or Uid is lost, there is no loss to the user or server.

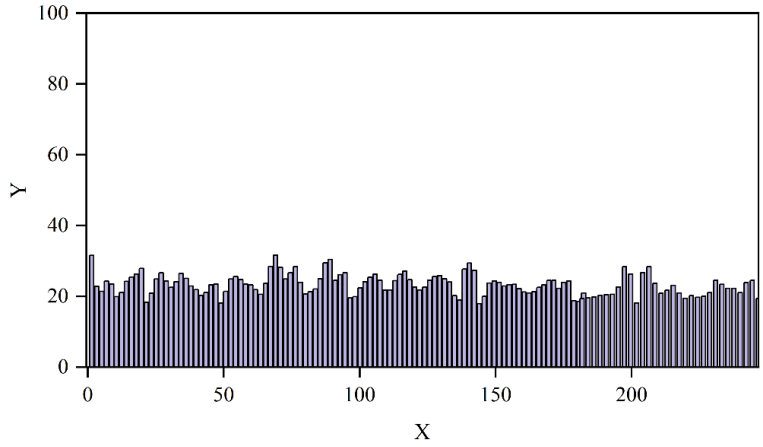
3.3 Experimentation and Analysis

3.3.1 Implementation of Chaotic Mapping Group Encryption Algorithm

In order to test the security performance of the designed encryption algorithm in this paper and to show the encryption and decryption effect, the encryption and decryption algorithm is implemented in code here, and the results are obtained by encrypting an image experimentally, as shown in Figure 4. From the encryption effect of Fig. 4 on the network communication information data, it can be seen that the encryption algorithm has achieved better results, and any relevant information of the original plaintext cannot be obtained from the encryption result at all. Even if the ciphertext is decrypted using a key with a small difference, no information about the plaintext can be obtained. Due to the complexity of the key expansion algorithm, even if a part of the key can be obtained, it is not possible to deduce the whole key. Even if the whole key is obtained in some case, the key expansion algorithm cannot be inferred, so the encryption key as well as the shift space for each round will not be known. Even if the wrong decryption key has a small difference from the correct decryption key, no information about the original ciphertext will be obtained.



(a) Encrypt before



(b) Encrypt after

Figure 4: Histogram of information before and after encryption

3.3.2 Ciphertext Distributability and Randomness Analysis

An important index to measure the performance of the data encryption algorithm for network communication messages with multiple chaotic mappings is the distribution characteristics of the plaintext and ciphertext as well as the randomness of the ciphertext's 0-1 binary sequence, if the distribution of the ciphertext is not sufficiently random or homogeneous, in which case the decoder can fully utilize this to crack the encrypted file, and then decrypt it. In order to reflect the performance of the encryption algorithm as accurately as possible, this paper encrypts an English text with a size of 8KB, and Fig. 5 and Fig. 6 show the respective histograms of the plaintext and the ciphertext obtained after the action of the encryption algorithm, respectively.

From the figure, it can be clearly seen that the spatial distribution of the ASCII values of the plaintext and ciphertext are very different, due to the fact that the original plaintext data has large statistical characteristics, while the encrypted data shows an average homogeneous characteristics, so it can be very good to hide the information in the data, and thus it can be very well protected against ciphertext-only attacks.

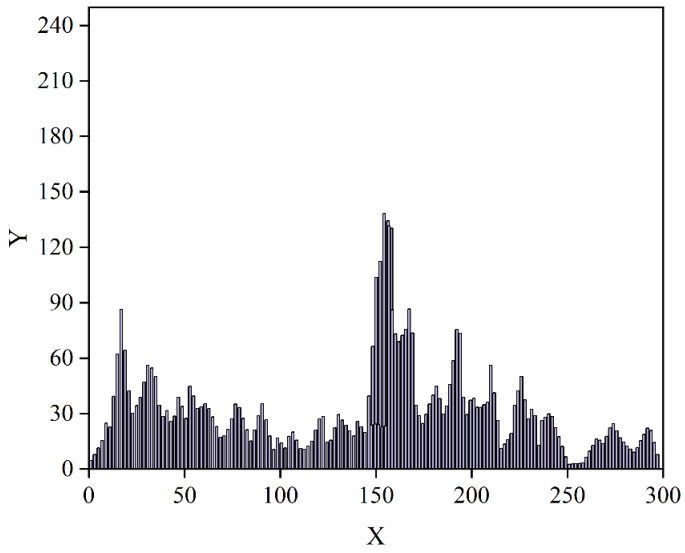


Figure 5: Plain Text Histogram

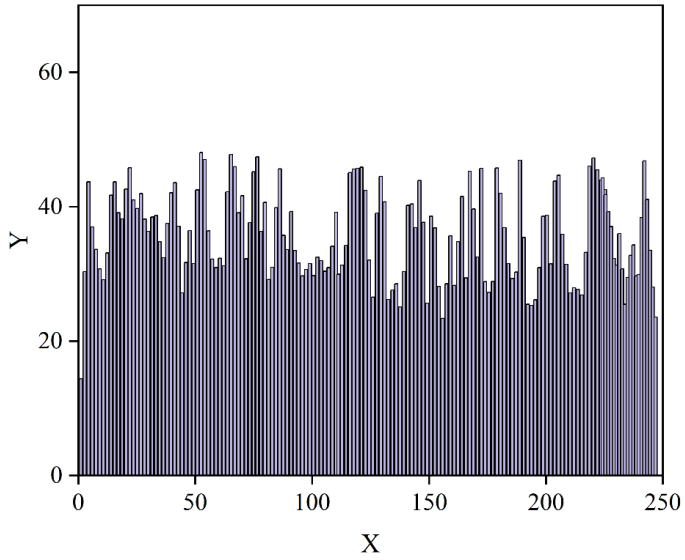


Figure 6: Ciphertext histogram

3.3.3 Explicit sensitivity

For the comparison of the plaintext sensitivity analysis between the chaotic mapping-based network communication information data encryption algorithm proposed in this paper and the network communication information data encryption algorithm without multiple chaotic mappings, two sets of 250-byte plaintexts with small differences are selected as follows:

The ciphertexts generated by M1 and M2 under the action of the proposed algorithm and the action of the network communication information data encryption algorithm without chaotic mapping, respectively, are differential, and the results obtained from the differential are used to measure the distributivity and randomness with histograms, and the obtained ciphertext differences are shown in Figs. 7 and 8.

The results obtained from the difference represent the difference between the two encrypted ciphertexts, and the large difference indicates that the cryptographic method is more sensitive to the plaintext; the results show that the processed ciphertexts have better consistency and random consistency. Fig. 7 shows that the encryption algorithm without chaotic mapping has

the same encryption result as long as the plaintext is the same; Fig. 8 shows that the encryption algorithm with chaotic mapping proposed in this paper can have very different encryption results even if the plaintext is the same. Comparing Fig. 7 and Fig. 8 shows that the encryption algorithm with chaotic mapping proposed in this paper strengthens the plaintext obfuscation and has better plaintext sensitivity and resistance to differential analysis.

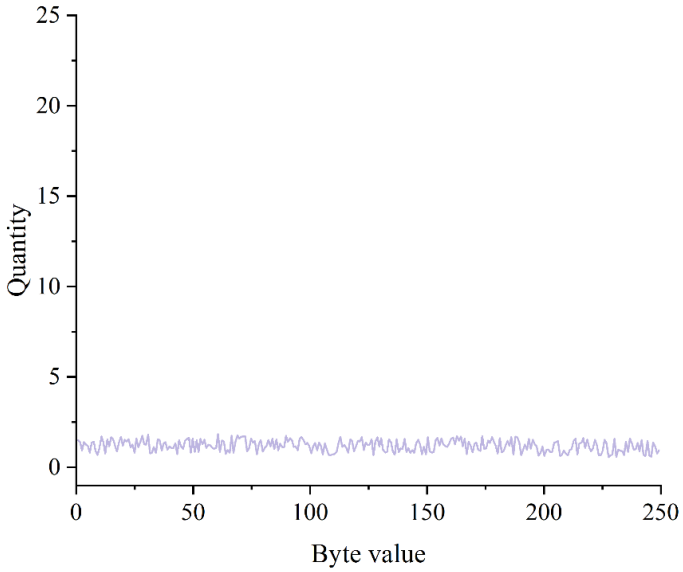


Figure 7: Histogram of ciphertext differences for unencrypted encryption algorithms

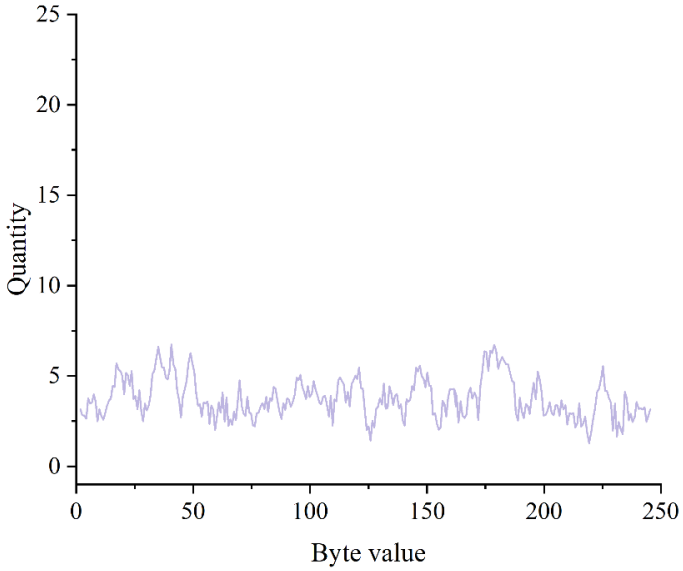


Figure 8: Histogram of ciphertext differences for a chaotic mapping-based encryption algorithm

3.3.4 Encryption time analysis

In order to verify the effectiveness of the method proposed in this paper, the typical SkipJack encryption algorithm and RC5 encryption algorithm are applied and tested together with the method of this paper.The SkipJack algorithm and RC5 algorithm are encryption algorithms with good performance for application in wireless sensor networks, which can well satisfy the requirement of encryption time in wireless sensor networks. Three encryption algorithms are

used to encrypt 5000 bytes of plaintext data respectively. In these three encryption algorithms, these 5000 bytes are divided into a group of 8 bytes for 550 encryptions and the time consumption for encryption and decryption is shown in Fig. 9.

From the time comparison graph, it can be seen that encrypting 5000 bytes of data, SkipJack encryption algorithm, RC5 algorithm, and chaotic mapping encryption algorithm use 0.883s, 0.573s, and 0.346s respectively. Compared with the other two encryption algorithms, the multi-chaotic mapping group encryption algorithm is much faster than the SkipJack encryption and RC5 algorithm encryption/decryption speed. Therefore, the encryption speed of multi-chaotic mapping group encryption algorithm can still satisfy the application of intelligent network security information data better.

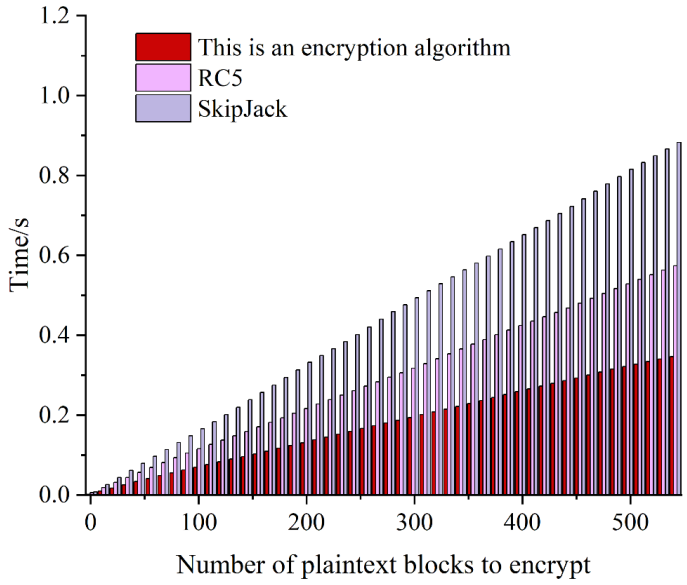


Figure 9: Compare encryption times for three encryption algorithms

4 Power Battery Failure Prediction for New Energy Vehicles

4.1 Power Battery Failure Correlation Analysis

4.1.1 Computational model for correlation analysis

According to the research preliminary selection and gearbox oil temperature related variables, ambient temperature, wind speed, gearbox shaft end temperature, output power and other related variables, the selected parameter variables for the wind turbine gearbox oil temperature can play a direct or indirect impact, using the Pearson correlation coefficient [39] correlation analysis, select and gearbox oil temperature correlation with the large input variables, in the statistics, the Pearson product moment correlation coefficient is often expressed as r or ρ, and takes a range of values between $(-1,+1)$. The Pearson product-moment correlation coefficient between two variables is defined by the quotient of the product of the covariance of the two variables and the standard deviation of the two, and the specific expression is:

$$\rho_{XY} = \frac{\operatorname{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E\left(X-\mu_X\right)\left(Y-\mu_Y\right)}{\sigma_X \sigma_Y} \tag{12}$$

where ρ denotes the overall correlation coefficient and X and Y are denoted as data samples. If the covariance and standard deviation of the data samples are used instead of the overall covariance and standard deviation, the correlation coefficient r can be expressed as:

$$r = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sqrt{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}\sqrt{\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2}} \tag{13}$$

The equivalent formula for the Pearson correlation coefficient is:

$$r = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{X_i - \bar{X}}{s_X}\right)\left(\frac{Y_i - \bar{Y}}{s_Y}\right) \tag{14}$$

where $\frac{X_i - \bar{X}}{s_X}, \bar{X}$ and s_X are the standardized variables, the sample mean and the sample standard deviation, respectively.

4.1.2 Analysis of correlation between faults

In this section, the data of a model of new energy vehicle in Beijing is selected for correlation calculation and data visualization operation. Table 1 shows the fault-related data and codes of new energy vehicles, and the heat map of fault frequency correlation analysis is shown in Figure 10, whose horizontal and vertical coordinates [1,2,...20] are attribute terms, and the more the color of the squares in the map tends to be red, the stronger the positive correlation is, and the more it tends to be blue, the stronger the negative correlation is. From Figure 10, it can be seen that the middle diagonal is autocorrelation, and its value is 1. By the nature of the correlation coefficient, we know that the correlation between X and Y is equal to the correlation between Y and X. Therefore, the correlation coefficient is symmetric about the diagonal. In addition to the effect of time and mileage, the rest is the autocorrelation effect between fault types, the analysis of the triangle on the figure shows that the overall color of the SOC jump alarm, high-voltage interlock alarm, under-voltage alarm of the on-board storage device, over-voltage alarm of the on-board storage device is inclined to the red color, which is more positively correlated with other faults, while the overall color of the battery single unit consistency alarm and the single unit over-voltage alarm is more inclined to the purple color, which is more negatively correlated with other faults. The overall color of the battery single unit poor consistency alarm and single unit battery over-voltage alarm is more inclined to purple, with stronger negative correlation with other faults.

Table 1: Fault data and codes for new energy vehicles

Fault data	Code
Overcharge of the on-board energy storage device	S1
High-voltage interlock status alarm	S2
DC-DC status alarm	S3
DC-DC temperature alarm	S4
Insulation alarm	S5
Poor cell consistency alarm	S6
SOC Jump Alarm	S7
SOC alarm is too high	S8
Single cell under-voltage alarm	S9
Overvoltage alarm for individual cells	S10
Low SOC alarm	S11
Under-voltage alarm for on-board energy storage device	S12
Overvoltage alarm for on-board energy storage device	S13
Battery overheating alarm	S14
Temperature Difference Alert	S15
Mileage	S16
Moon	S17

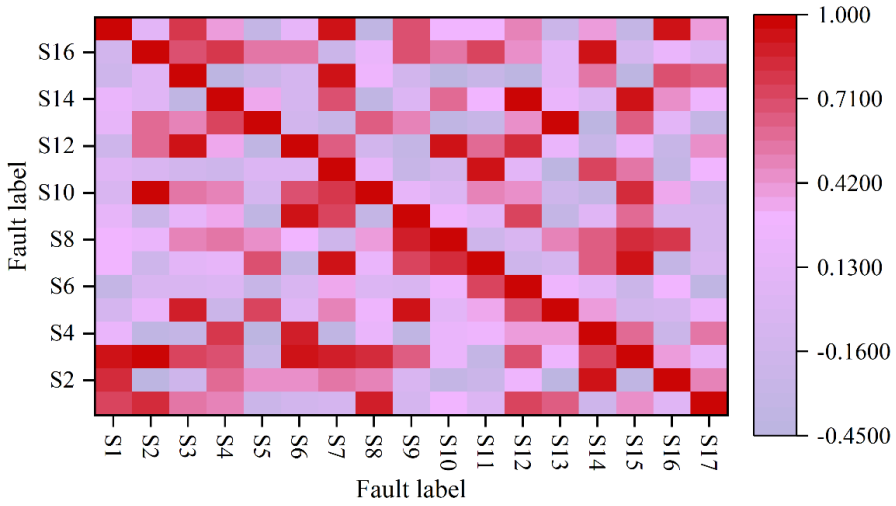


Figure 10: Fault frequency correlation analysis heatmap

4.1.3 Failure Correlation Analysis of the Mileage Dimension

This section quantitatively analyzes the correlation between vehicle mileage and fault frequency by extracting the mileage dimension fault correlation vector.

The vector is first extracted for the target study attribute (mileage):

$$Corr_{mile} = \left[corr_{1,m}^{T} \quad \ldots \quad corr_{n,m}^{T} \right]^{T} \tag{15}$$

The data visualization operation is performed on the mileage dimension fault correlation vector to obtain the mileage correlation coefficient curve, as shown in Fig. 11. From the figure, it can be seen that the positive correlation between the mileage dimension and the battery monomer poor consistency fault is relatively high, and the correlation coefficient reaches 0.229;

the negative correlation with the temperature difference fault is high, and the correlation coefficient reaches -0.289; the negative correlation with the monomer overvoltage fault is high, and the correlation coefficient reaches -0.203; the negative correlation with the SOC excessive fault is high, and the correlation coefficient reaches -0.355; and the negative correlation with the high-voltage interlock fault is high. 0.355; the negative correlation with the high-voltage interlock fault is high, and the correlation coefficient reaches -0.203.

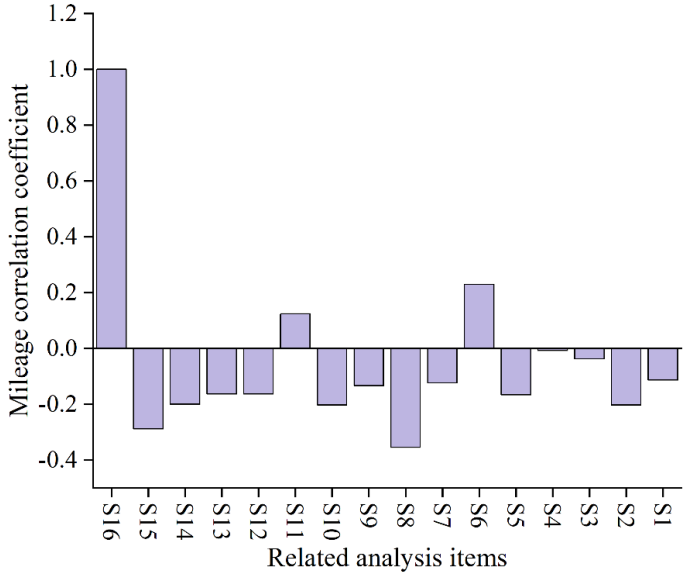


Figure 11: Distance correlation coefficient curve

4.2 Battery Failure Prediction Based on LS-SVM

4.2.1 SVM algorithm

Support Vector Machine (SVM) belongs to one of the machine learning algorithms, which can be mainly divided into SVM classification model and SVM regression model according to the specific application, which can better realize the idea of structural risk minimization, and is now widely used in the field of machine learning.

(1) SVM principle

Set the training sample $\{(x_1, y_1),(x_2, y_2),\cdots,(x_t, y_t)\}$, where $x_t \in R_n, y_i \in \{1,-1\}$, $i = 1,2,\cdots,t$. H, H_1, H_2 can be expressed as:

$$H : w \bullet x + b = 0 \tag{16}$$

$$H1 : w \bullet x_i + b \geq 1, y_i = 1 \tag{17}$$

$$H2 : w \bullet x_i + b < -1, y_i = -1 \tag{18}$$

where w is the normal vector, b is the offset value, and $w \bullet x$ is the inner product of w and x.

There are no data points between H_1 and H_2, and all points outside of it conform to $|w \bullet x + b| \geq 1$, and when $|w \bullet x + b| = 1$, it is the closest point to the Optical Hyper Plane, and the distance of the point x to the hyper plane H is denoted as:

$$d\left(w,b,x_i\right)=\frac{\left|w\bullet x_i+b\right|}{\left\|w\right\|} \tag{19}$$

The distance between the spaced surfaces where the support vectors are located is:

$$\rho(w,b)=\frac{2}{\left\|w\right\|} \tag{20}$$

The optimal hyperplane is also the hyperplane that maximizes $\rho(w,b)$, viz:

$$\min_w=\frac{\left\|w\right\|}{2} \tag{21}$$

SVM is essentially to find the maximum interval between H_1,H_2, which is computed by minimizing $\left\|w\right\|^2$ to get the maximum interval hyperplane satisfying the constraints $y_i\left(w\bullet x_i+b\ge1\right),i=1,2,\cdots t$, i.e.:

$$\begin{cases}\min\limits_w\phi(w)=\dfrac{\left\|w\right\|^2}{2}\\ s.t\qquad y_i\left(w\bullet x_i+b\right)\ge1,i=1,2,\cdots t\end{cases} \tag{22}$$

The convex quadratic optimization problem with constraints is transformed into an optimization problem without constraints as shown in Eq. (23):

$$L(w,b,\alpha)=\frac{1}{2}\left\|w\right\|^2-\sum_{i=1}^{t}\alpha_i\bullet\left\{y_i\left[\left(w\bullet x_i\right)+b\right]-1\right\} \tag{23}$$

where α_i is the Lagrange multiplier. $L(w,b,\alpha)$ takes the partial derivatives of w and b to find the maximum value.

(2) SVM regression

It is insensitive when the error is less than ε and can be considered as 0. The expression is as follows:

$$\left|y-f(x)\right|=\begin{cases}0, & \left|y-f(x)\right|\le\varepsilon\\ \left|y-f(x)\right|-\varepsilon, & \left|y-f(x)\right|>\varepsilon\end{cases} \tag{24}$$

In the linear problem, the training samples $\left\{\left(x_1,y_1\right),\left(x_2,y_2\right),\cdots,\left(x_t,y_t\right)\right\},x_i\in R_n$ refer to n-dimensional input vector, $y_i\in R$ as the corresponding objective value, and the expression of the linear function $f(x)$ is:

$$f(x)=w\cdot x+b \tag{25}$$

It is further transformed into a regression problem by introducing the slack variable ξ with the aim of improving the accuracy and constructing the fault-tolerant penalty coefficient C, at

which point the SVM takes the form:

$$\begin{cases} \min \quad \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{t}\left(\xi_i + \xi_i^*\right) \\ s.t.\begin{cases} y_i - \left(w \bullet x_i\right) - b \le \varepsilon + \xi_i \\ \left(w \bullet x_i\right) + b - y_i \le \varepsilon + \xi_i^* \quad i = 1, 2, \cdots, t \\ \xi_i, \xi_i^* > 0 \end{cases} \end{cases} \tag{26}$$

The Lagrangian function is introduced and transformed into an optimization problem without constraints, which in turn is calculated in the following equation (27):

$$\begin{aligned} L\left(w, b, \alpha, \alpha^*\right) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{t}\left(\xi_i + \xi_i^*\right) - \sum_{i=1}^{t}\alpha_i \bullet \left[\varepsilon + \xi_i - y_i + \left(w \bullet x_i\right) + b\right] \\ - \sum_{i=1}^{t}\alpha_i^* \bullet \left[\varepsilon + \xi_i^* - y_i + \left(w \bullet x_i\right) - b\right] - \sum_{i=1}^{t}\left(\eta_i \xi_i + \eta_i^* \xi_i^*\right) \end{aligned} \tag{27}$$

where $\alpha, \alpha^* \ge 0$ is the Lagrange multiplier. The optimal solution satisfies the KKT condition by taking the partial derivative of $L\left(w, b, \alpha, \alpha^*\right)$ to find the optimal value, and the following equation is obtained:

$$\begin{cases} \frac{\partial L}{\partial w} = 0 \\ \frac{\partial L}{\partial b} = 0 \\ \frac{\partial L}{\partial \xi_i} = 0 \\ \frac{\partial L}{\partial \xi_i^*} = 0 \end{cases} \Rightarrow \begin{cases} w - \sum_{i=1}^{t}\left(\alpha_i - \alpha_i^*\right)x_i = 0 \\ \sum_{i=1}^{t}\left(\alpha_i - \alpha_i^*\right) = 0 \\ C - \alpha_i - \eta_i = 0 \\ C - \alpha_i^* - \eta_i^* = 0 \end{cases} \tag{28}$$

using the dyadic form of the Lagrangian optimization problem, reduces to:

$$\begin{cases} \max W\left(\alpha_i, \alpha_i^*\right) = -\frac{1}{2}\sum_{i=1}^{t}\left(\alpha_i - \alpha_i^*\right)\left(\alpha_j - \alpha_j^*\right)\left(x_i \bullet x_j\right) \\ \qquad\qquad + \sum_{i=1}^{t}\left(\alpha_i - \alpha_i^*\right)y_i - \sum_{i=1}^{t}\left(\alpha_i + \alpha_i^*\right)\varepsilon \\ s.t.\begin{cases} \sum_{i=1}^{t}\left(\alpha_i - \alpha_i^*\right) = 0 \qquad i = 1, 2, \cdots, t \\ 0 \le \alpha_i, \alpha_i^* \le C \end{cases} \end{cases} \tag{29}$$

It is further solved to obtain the optimal solution and then w, b is determined. The final regression function is shown in (30):

$$f(x) = \sum_{i=1}^{n}\left(\alpha_i - \alpha_i^*\right)\left(x_i \bullet x_j\right) + b^* \tag{30}$$

When dealing with nonlinear training data samples, a definite nonlinear mapping ϕ can

be used to map the input vectors from the low-dimensional space to the high-dimensional space, so as to carry out a linear regression in the high-dimensional space, and the process does not need to know the specific form of the nonlinear mapping ϕ, and utilizes the introduction of the kernel function K to carry out the operation, which in turn ultimately results in a nonlinear regression function:

$$f(x)=\sum_{i=1}^{n}\left(\alpha_i-\alpha_i^*\right)K\left(x\bullet x_i\right)+b^* \tag{31}$$

4.2.2 Principles of LS-SVM

In large sample data processing, ordinary SVM is computationally complex and has slow convergence speed. Least squares support vector machine, using least squares linear system instead of traditional SVM, adopts quadratic programming method to solve the function estimation problem, this method has the advantages of global optimization, strong generalization ability, higher convergence speed and accuracy.

Suppose the wind turbine SCADA data has n training data sample space $D=\left\{\left(x_i,y_i\right)\mid i=1,\cdots,n\right\}$, where the input training samples $x_i\in R^N$ and the output training samples $y_i\in R,N$ is the dimension of the input space and the least squares support vector machine regression function is:

$$f(x)=w^T\cdot\varphi(x)+b \tag{32}$$

where: w is the weight vector of the classifier, $w\in R^N$; b is the bias, $b\in R$.

In order to combine the complexity and error of the function and make the minimum value obtained, its optimization problem can be described as:

$$\min J(w,e)=\frac{1}{2}w^Tw+\frac{1}{2}\gamma\sum_{i=1}^{n}e_i^2 \tag{33}$$

$$y_i=w^T\cdot x_i+b+e_i \tag{34}$$

The next step is to introduce a Lagrangian function to solve the above optimization problem, denoted as follows:

$$L(w,b,e,\alpha)=J(w,e)-\sum_{i=1}^{n}\alpha_i\left(w^T\cdot x_i+b+e_i-y_i\right) \tag{35}$$

where α is the Lagrange multiplier and e_i is the training sample error term.

According to the KKT condition, the partial differentiation of w,b,e,α respectively, i.e:

$$\begin{cases} \dfrac{\partial L}{\partial w_i} = 0 \\ \dfrac{\partial L}{\partial b} = 0 \\ \dfrac{\partial L}{\partial e_i} = 0 \\ \dfrac{\partial L}{\partial \alpha_i} = 0 \end{cases} \tag{36}$$

Solving the system of equations, and hence the optimal solution, yields the regression function of the least squares support vector machine, i.e:

$$y = \sum_{i=1}^{n} \alpha_i \cdot xx_i + b = \sum_{i=1}^{n} \alpha_i K\left(x, x_i\right) + b \tag{37}$$

where: $K\left(x, x_i\right)$ is the kernel function.LS-SVM saves computer memory resources, simplifies the computational process, and effectively improves the computational efficiency compared with SVM, while relatively few parameters need to be optimized, and this algorithm has better generalizability.

4.3 Effectiveness of model application

4.3.1 Data selection

Voltage fault data includes three categories: first, overvoltage fault data, the data includes 100 groups of sampling points, overvoltage occurs at the 50th sampling point; second, undervoltage fault data, undervoltage fault data is a fault generated when the total voltage is lower than a certain threshold, usually caused by rapid acceleration, the data includes 120 groups of sampling points; normal voltage data, including 120 groups of sampling points.

4.3.2 Results of predictive analysis

The first d-1 of d consecutive time series values are taken as the training set and the 2nd to the dth are taken as the test set from the original training sample set G. The LS-SVR overvoltage true and predicted values as well as the prediction error are shown in Fig. 12 and Fig. 13. The error is very small, around ±0.2V, when the voltage value is rising and falling. During a sudden drop in voltage, the error increases and can reach a maximum of -0.5V.

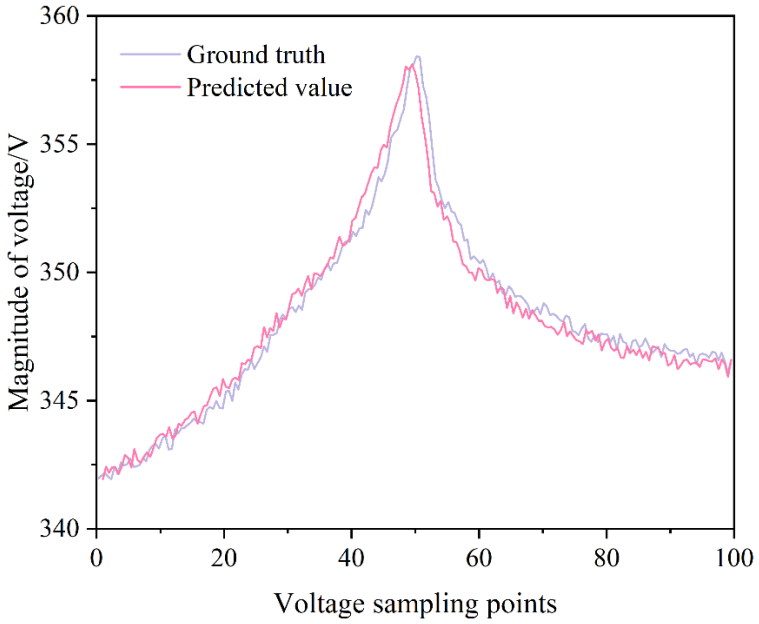


Figure 12: Actual vs. predicted overvoltage values for the LS-SVR

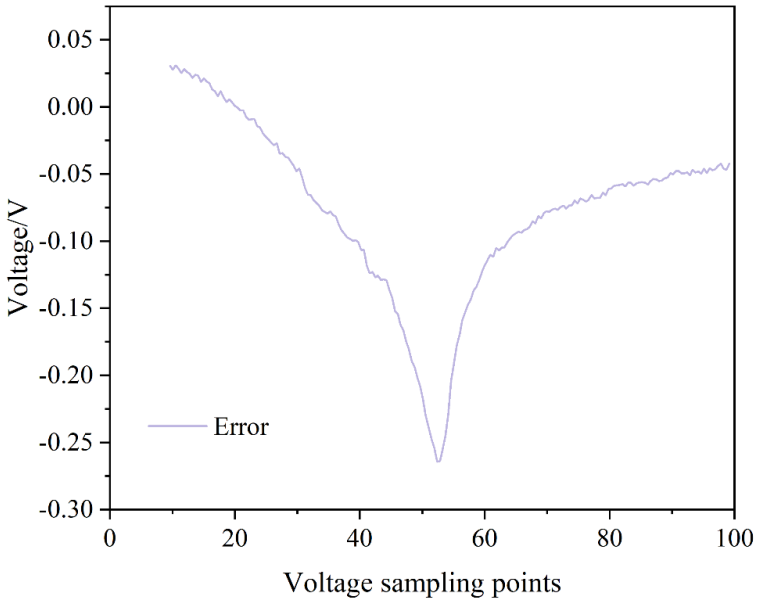


Figure 13: LS-SVR overvoltage prediction error

The above analyzed total voltage overvoltage data is predicted with the algorithm of least squares support vector machine. The same approach is taken below to predict the total voltage undervoltage data. Fig. 14 and Fig. 15 show the LS-SVR prediction plots as well as the error plots under total voltage undervoltage, respectively. In the case of relatively large voltage fluctuations, the prediction value will have a large error, the maximum will be 9.3% error, which is normal. If the voltage continues to undervoltage discharge for a period of time, the predicted value will immediately track the original curve, and when the voltage has a stabilizing trend, the predicted value quickly approaches the actual value.

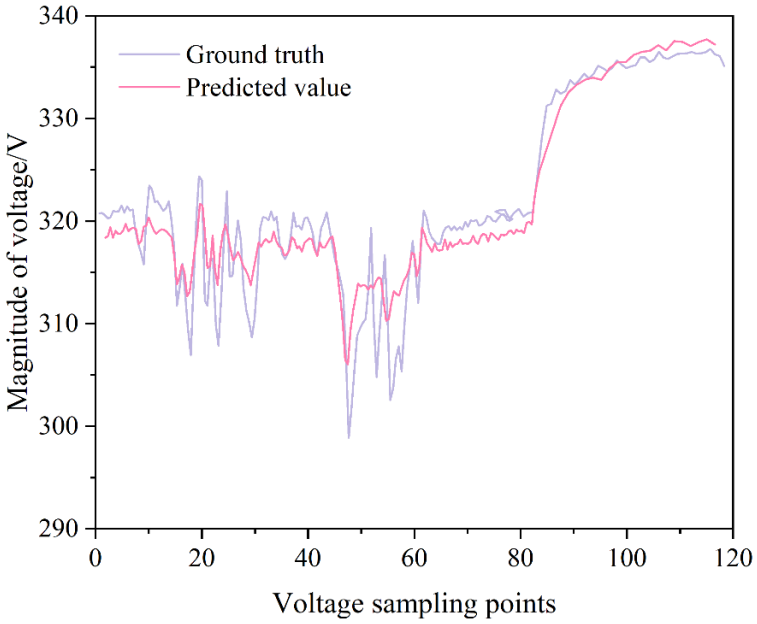


Figure 14: LS-SVR Total Voltage Undervoltage Prediction Chart

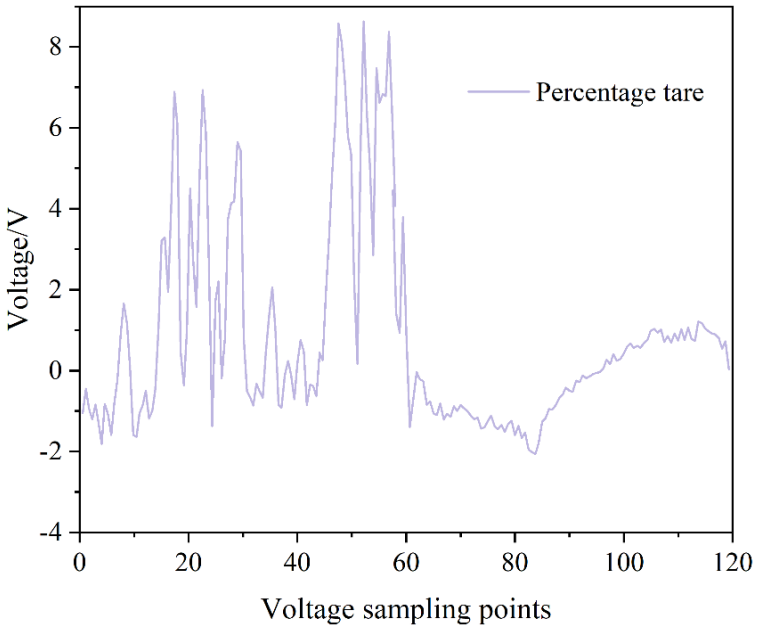


Figure 15: LS-SVR Total Voltage Undervoltage Prediction Error Chart

Fig. 16 and Fig. 17 show the LS-SVR prediction plots as well as the error plots under normal voltage, respectively. As can be seen from the plots under normal voltage, the error is slightly larger only when there is a sudden voltage change, but the foot duration is not long, while in other voltage continuous moments, the prediction bits of the least squares support vector regression machine are able to track the original data very well. Under the LS-SVR model, the maximum error is only 0.8 V. The prediction can be made in any state of the battery and there is no overvoltage false alarm.

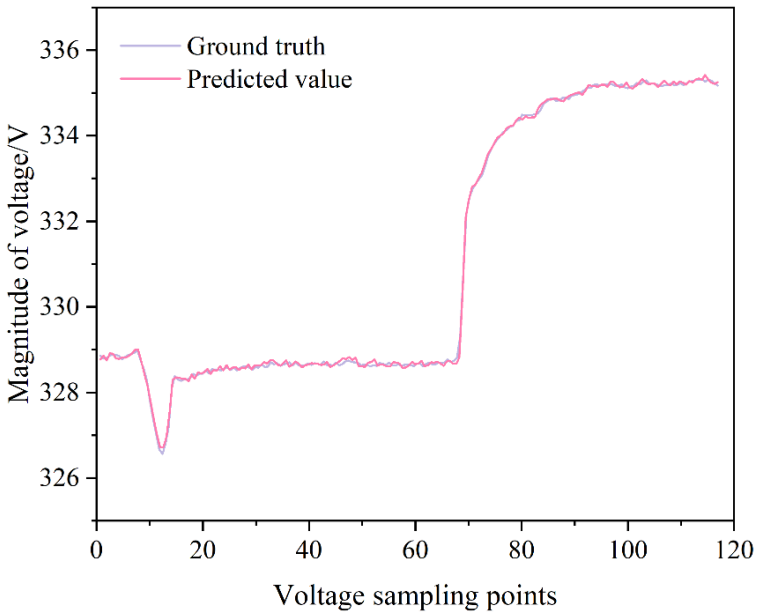


Figure 16: LS-SVR normal electric bed prediction chart

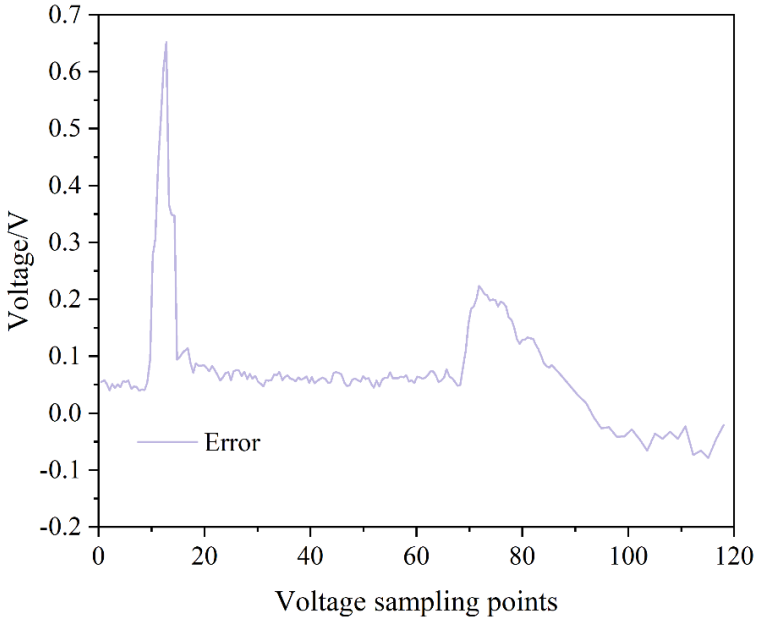


Figure 17: LS-SVR Normal Voltage Prediction Error Chart

4.3.3 Analysis of operational results

1. Comparison of SVR and LS-SVR prediction effectiveness

In regression prediction error metrics Mean Square Error (MSE) and Mean Percentage Error (MAPE) are introduced in order to evaluate the approximation ability of the support vector machine and the metrics to generalize the error performance.The expressions of MSE and MAPE are given below:

$$MSE = \frac{1}{n}\sum_{i=1}^{n} w_i \left(y_i - \hat{y}_i \right)^2 \tag{38}$$

$$MAPE = \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times \frac{100}{n} \tag{39}$$

According to the prediction effect of SVR and LS-SVR above, both of them can accurately track the trend of the actual value, and only in the case of a sudden change in voltage, there will be a slightly larger error. However, the maximum voltage will not exceed 1.2V, and the general fault threshold voltage is set to be more than 10V greater than the normal operating voltage, which fully meets the actual situation, and basically will not appear in the case of false alarms.

Comparison of SVR and LS-SVR overvoltage operation results are shown in Table 2, the root mean square error result of SVR is 0.0784, and the average percentage error is 0.0571, while the error result of LS-SVR is even smaller, which is 0.0124 and 0.0259, respectively.The two prediction algorithms have a better fit, but the root mean square error and average percentage error of the least squares support vector machine are better than those of the support vector machine. error are smaller than those of the Support Vector Machine and the running time is faster.LS-SVR simplifies the complexity of the problem, the algorithm converges quickly, requires less computational resources, and is more suitable for use in situations where real-time requirements are high.

Table 2: Comparison of overvoltage operation results between SVR and LS-SVR

	MSE	MAPE	Performance period/s
SVR	0.0784	0.0571	1.372541
LS-SVR	0.0124	0.0259	0.321832

The comparison results of SVR and LS-SVR total voltage undervoltage operation are shown in Table 3, the root mean square negotiation result of SVR is 12.2841, and the average percentage error is 0.773, while the results of LS-SVR are small, which are 7.8754 and 0.722, respectively. Least Squares Support Vector Machines have smaller root mean square error and average percentage misinfection than Support Vector Machines and run faster. It can also be demonstrated that LS-SVR has less prediction error, better prediction results and faster runtime than SVR.

Table 3: Comparison of Undervoltage Operation Results Between SVR and LS-SVR

	MSE	MAPE	Performance period/s
SVR	12.2841	0.773	4.392842
LS-SVR	7.8754	0.722	0.356737

The normal voltage operation results of SVR and LS-SVR are normalized as shown in Table 4, which also proves that LS-SVR has less prediction error than SVR, better prediction results, and faster running time officials. It also shows that the prediction error at normal voltage is also very small.

Table 4: Comparison of normal voltage operation results between SVR and LS-SVR

	MSE	MAPE	Performance period/s
SVR	0.0672	0.0351	2.197528
LS-SVR	0.0248	0.0314	0.735193

2. Result prediction

Since LS-SVR predicts better results, so it is chosen to use LS-SVR as the fault prediction model. According to LS-SVR to predict the last 5, 10, 15 time points of data respectively, the first N-5, N-10, N-15 data as training data, and the last 5, 10, 15 data as prediction data, to get the real value and prediction value of the comparison of the graphs, as shown in Fig. 18-Fig. 20. The error in the prediction to 1~4 time points is relatively small, usually around 0.4V relative to the total voltage of 309V is basically negligible. It can be set to be more accurate after predicting four time points.

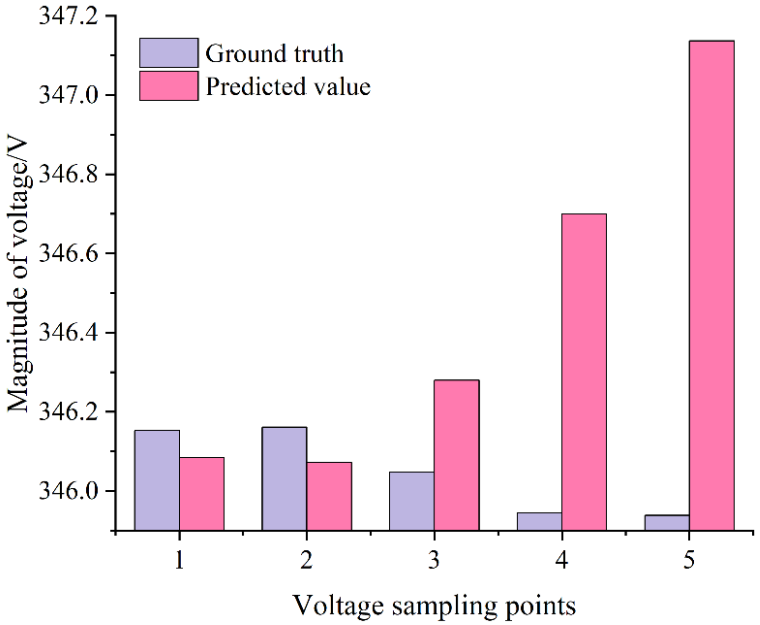


Figure 18: The true and predicted values of the last 5 moments

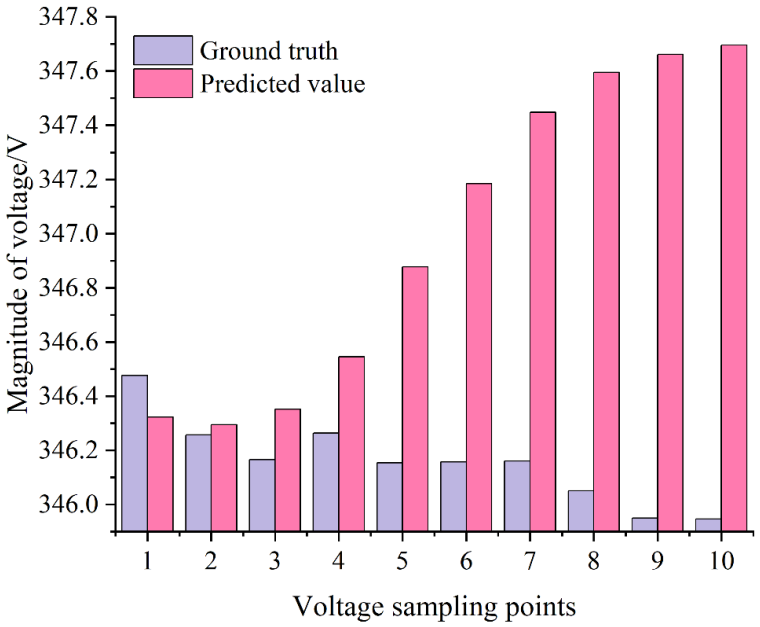


Figure 19: The true and predicted values for the last 10 moments

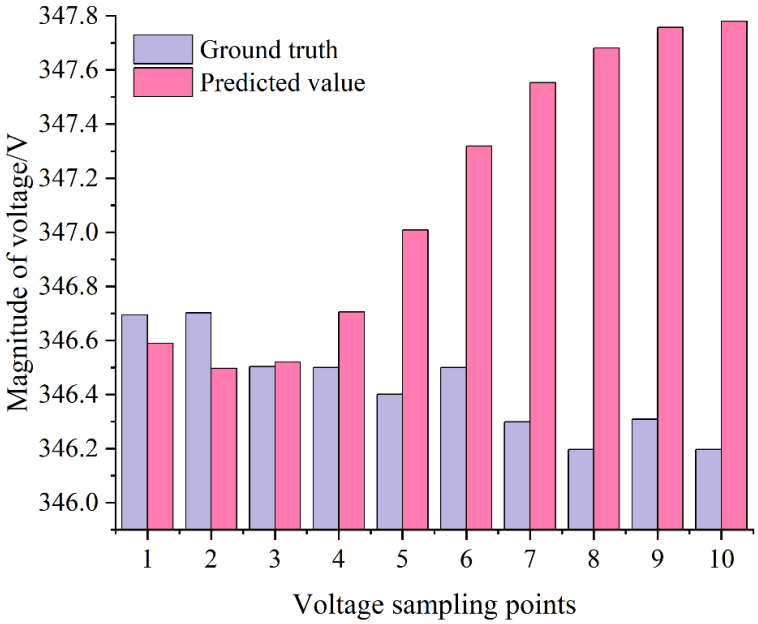


Figure 20: The true and predicted values for the last 15 moments

According to the above proof of prediction after four points in time is more accurate, Table 5 for the prediction of the fault alarm before the four points in time the data in the actual situation of the total voltage to reach the peak, this time for the 358.2V will produce a failure to report the alarm and in advance of the four points in time can be predicted to get the basic and the actual alarm value coincides with the accurate prediction of overvoltage faults as a result of the faults.

Table 5: Compare the actual values with the predicted values at the first four time points before the fault

Ground truth	355.4	356.2	357.4	358.2
Predicted value	355.48	356.27	357.37	358.62
Error percentage/%	0.0122	0.0984	0.114	0.111

5 Conclusions and outlook

5.1 Conclusion

The research centers around the intelligent networked vehicle system, power battery twin model and designing the power battery management system for new energy vehicles, a new dynamic key generation method is designed by adding cyclic shift control bits in the seed key, which generates the dynamic wheel key with dynamic cyclic shift, effectively reducing the extra operation, for this reason, this paper proposes a group based on Logistic Chaos System encryption algorithm; at the same time, for new energy vehicle power battery failure prediction, LS-SNM prediction model is constructed based on SVM algorithm, and correlation calculations are carried out for the power battery failure and failure, failure and driving mileage dimensions, comparing the performance of SVR and LS-SVR. The experimental results show that the information encryption method of two-way authentication of identity has strong security and practicality, followed by a relatively high degree of positive correlation between the mileage dimension and the battery monomer consistency poor failure, with a correlation coefficient of 0.229, and the prediction results show that the LS-SNM prediction model will have a very good

tracking effect at four time points after consecutive prediction, and will be shifted from the true value after four time points.

5.2 Outlook

This study integrates technical points in the field of power battery of intelligent network and new energy vehicles in multiple dimensions, and forms methods and models with practical value in network security encryption and battery failure prediction, providing theoretical support and technical reference for the safe and efficient operation of new energy vehicles in the intelligent network, and the verification of the analysis results of the power battery voltage outliers is limited due to the different battery arrangements in different models of electric vehicles We expect more OEMs to provide data analysis result verification environment. It is expected that more OEMs will provide a validation environment for the data analysis results, and further research can be carried out in the directions of algorithm optimization, multi-scene adaptation and model iteration to enhance the application efficiency in practical engineering.

References

[1] Wu, H. T., & Horng, G. J. (2017). Establishing an intelligent transportation system with a network security mechanism in an Internet of vehicle environment. Ieee Access, 5, 19239-19247.

[2] Lokman, S. F., Othman, A. T., & Abu-Bakar, M. H. (2019). Intrusion detection system for automotive Controller Area Network (CAN) bus system: a review. EURASIP Journal on Wireless Communications and Networking, 2019(1), 1-17.

[3] Veres, S. M., Molnar, L., Lincoln, N. K., & Morice, C. P. (2011). Autonomous vehicle control systems—a review of decision making. Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering, 225(2), 155-195.

[4] Noor-A-Rahim, M., Liu, Z., Lee, H., Khyam, M. O., He, J., Pesch, D., ... & Poor, H. V. (2022). 6G for vehicle-to-everything (V2X) communications: Enabling technologies, challenges, and opportunities. Proceedings of the IEEE, 110(6), 712-734.

[5] Raouf, B., & Mousavian, S. (2025). False data injection to conceal load-altering attacks via electric vehicles. Sustainable Energy, Grids and Networks, 42, 101640.

[6] Yan, G., Wen, D., Olariu, S., & Weigle, M. C. (2012). Security challenges in vehicular cloud computing. IEEE Transactions on intelligent transportation systems, 14(1), 284-294.

[7] Gade, A. R. (2021). The new battery management system in electric vehicle. IJERT-International Journal of Engineering Research Technology, 7.

[8] Kosuru, V. S. R., & Kavasseri Venkitaraman, A. (2023). A smart battery management system for electric vehicles using deep learning-based sensor fault detection. World Electric Vehicle Journal, 14(4), 101.

[9] Liu, H., Song, X., & Zhang, F. (2021). Fault diagnosis of new energy vehicles based on

improved machine learning. Soft Computing, 25(18), 12091-12106.

[10] Karmawijaya, M. I., Haq, I. N., Leksono, E., & Widyotriatmo, A. (2019, November). Development of big data analytics platform for electric vehicle battery management system. In 2019 6th international conference on electric vehicular technology (ICEVT) (pp. 151-155). IEEE.

[11] Li, S., & Zhao, P. (2021). Big data driven vehicle battery management method: A novel cyber-physical system perspective. Journal of Energy Storage, 33, 102064.

[12] Botín-Sanabria, D. M., Mihaita, A. S., Peimbert-García, R. E., Ramírez-Moreno, M. A., Ramírez-Mendoza, R. A., & Lozoya-Santos, J. D. J. (2022). Digital twin technology challenges and applications: A comprehensive review. Remote Sensing, 14(6), 1335.

[13] Yang, D., Cui, Y., Xia, Q., Jiang, F., Ren, Y., Sun, B., ... & Yang, C. (2022). A digital twin-driven life prediction method of lithium-ion batteries based on adaptive model evolution. Materials, 15(9), 3331.

[14] Fuller, A., Fan, Z., Day, C., & Barlow, C. (2020). Digital twin: enabling technologies, challenges and open research. IEEE access, 8, 108952-108971.

[15] Chen, Y. (2022). Research on collaborative innovation of key common technologies in new energy vehicle industry based on digital twin technology. Energy Reports, 8, 15399-15407.

[16] Vandana, Garg, A., & Panigrahi, B. K. (2021). Multi-dimensional digital twin of energy storage system for electric vehicles: a brief review. Energy Storage, 3(6), e242.

[17] Barricelli, B. R., Casiraghi, E., & Fogli, D. (2019). A survey on digital twin: Definitions, characteristics, applications, and design implications. IEEE access, 7, 167653-167671.

[18] Fang, X., Wang, H., Liu, G., Tian, X., Ding, G., & Zhang, H. (2022). Industry application of digital twin: from concept to implementation. The International Journal of Advanced Manufacturing Technology, 121(7), 4289-4312.

[19] Minerva, R., Lee, G. M., & Crespi, N. (2020). Digital twin in the IoT context: A survey on technical features, scenarios, and architectural models. Proceedings of the IEEE, 108(10), 1785-1824.

[20] Yao, J. F., Yang, Y., Wang, X. C., & Zhang, X. P. (2023). Systematic review of digital twin technology and applications. Visual computing for industry, biomedicine, and art, 6(1), 10.

[21] Cao, H., Zhang, D., & Yi, S. (2023). Real-Time Machine Learning-based fault Detection, Classification, and locating in large scale solar Energy-Based Systems: Digital twin simulation. Solar Energy, 251, 77-85.

[22] Dai, Y., & Zhang, Y. (2022). Adaptive digital twin for vehicular edge computing and networks. Journal of Communications and Information Networks, 7(1), 48-59.

[23] He, C., Luan, T. H., Lu, R., Su, Z., & Dong, M. (2022). Security and privacy in vehicular

digital twin networks: Challenges and solutions. IEEE Wireless Communications, 30(4), 154-160.

[24] Ali, M., Kaddoum, G., Li, W. T., Yuen, C., Tariq, M., & Poor, H. V. (2023). A smart digital twin enabled security framework for vehicle-to-grid cyber-physical systems. IEEE Transactions on Information Forensics and Security, 18, 5258-5271.

[25] Lv, Z., Li, Y., Feng, H., & Lv, H. (2021). Deep learning for security in digital twins of cooperative intelligent transportation systems. IEEE transactions on intelligent transportation systems, 23(9), 16666-16675.

[26] Liu, J., Zhang, L., Li, C., Bai, J., Lv, H., & Lv, Z. (2022). Blockchain-based secure communication of intelligent transportation digital twins system. IEEE transactions on intelligent transportation systems, 23(11), 22630-22640.

[27] Almeaibed, S., Al-Rubaye, S., Tsourdos, A., & Avdelidis, N. P. (2021). Digital twin analysis to promote safety and security in autonomous vehicles. IEEE Communications Standards Magazine, 5(1), 40-46.

[28] Khawaja, Y., Shankar, N., Qiqieh, I., Alzubi, J., Alzubi, O., Nallakaruppan, M. K., & Padmanaban, S. (2023). Battery management solutions for li-ion batteries based on artificial intelligence. Ain Shams Engineering Journal, 14(12), 102213.

[29] Merkle, L., Segura, A. S., Grummel, J. T., & Lienkamp, M. (2019, May). Architecture of a digital twin for enabling digital services for battery systems. In 2019 IEEE international conference on industrial cyber physical systems (ICPS) (pp. 155-160). IEEE.

[30] Bugueño, V., Barbosa, K. A., Rajendran, S., & Díaz, M. (2022, October). An overview of digital twins methods applied to lithium-ion batteries. In 2022 IEEE International Conference on Automation/XXV Congress of the Chilean Association of Automatic Control (ICA-ACCA) (pp. 1-7). IEEE.

[31] Semeraro, C., Aljaghoub, H., Abdelkareem, M. A., Alami, A. H., & Olabi, A. G. (2023). Digital twin in battery energy storage systems: Trends and gaps detection through association rule mining. Energy, 273, 127086.

[32] Yuan, Z., Pan, Y., Wang, H., Wang, S., Peng, Y., Jin, C., ... & Ouyang, M. (2023). Fault data generation of lithium ion batteries based on digital twin: A case for internal short circuit. Journal of Energy Storage, 64, 107113.

[33] Eaty, N. D. K. M., & Bagade, P. (2023). Digital twin for electric vehicle battery management with incremental learning. Expert Systems with Applications, 229, 120444.

[34] Renold, A. P., & Kathayat, N. S. (2024). Comprehensive review of machine learning, deep learning, and digital twin data-driven approaches in battery health prediction of electric vehicles. IEEE Access, 12, 43984-43999.

[35] Pooyandeh, M., & Sohn, I. (2023). Smart lithium-ion battery monitoring in electric vehicles: An AI-empowered digital twin approach. Mathematics, 11(23), 4865.

[36] Jafari, S., & Byun, Y. C. (2022). Prediction of the battery state using the digital twin

framework based on the battery management system. IEEE Access, 10, 124685-124696.

[37] Wang, W., Wang, J., Tian, J., Lu, J., & Xiong, R. (2021). Application of digital twin in smart battery management systems. Chinese Journal of Mechanical Engineering, 34(1), 57.

[38] Li, H., Kaleem, M. B., Chiu, I. J., Gao, D., Peng, J., & Huang, Z. (2024). An intelligent digital twin model for the battery management systems of electric vehicles. International Journal of Green Energy, 21(3), 461-475.

[39] Qimin Zhou, Yingcang Ma, Zhiwei Xing & Xiaofei Yang. (2025). Pearson correlation coefficient-guided large-scale fuzzy cognitive maps learning algorithm. Fuzzy Sets and Systems, 519,109523-109523.