ABSTRACT  Smart home is an emerging form of the Internet of things (IoT), which provides a convenient and comfortable living environment for the smart home users. With the explosive growth of smart home information, people pay more attention to the privacy protection of smart home, including the choice of privacy and effective privacy protection scheme. This paper proposes a privacy protection scheme based on information hiding, the scheme guarantees the sensitive data transmitted securely. First, the smart home Real-time sensor data are classified into sensitive data and non-sensitive data through machine learning, the process can be controlled according to the user’s preferences. Second, the sensitive data can be transmitted securely by ordinary channels using method of combination encryption with information hiding. Experiment results show that this scheme can help users to express their preferences to the privacy protection, changing the previous practice that users can only accept the privacy settings provided by the supplier. What’s more, the scheme combines encryption with information hiding, providing a double guarantee for the sensitive data during they are transmitted on ordinary channels. Comparing to the previous practice of using encryption merely, this scheme greatly improves the security of data transmission and is more suitable for the real environment.

INDEX TERMS  Smart home, privacy protection, data encryption, information hiding.

I. INTRODUCTION

With the advent of the Internet, the number of IoT connected objects surpass the population of earth’s human in 2010[1]. It is expected to reach 212 billion entities by the end of 2020 [2], and the data of IoT is estimated to account for 10 % of the entire digital universe [3]. The rapid development of wireless communication technology and wireless sensor technology greatly facilitates people’s communication and improves their perception of the world [4], [5]. Smart home is an important branch of IoT, which utilizes ubiquitous computing and information technology to anticipate and respond to the needs of residents and improve their daily living standards. One of the major factors to the rapid development of IoT is that it benefits much from the development of smart home. At the same time, IoT provides connection for smart devices in smart home [6], which in turn promotes the development of smart home. According to IDC Global Quarterly Smart Home Device Tracking Data, in the first quarter of 2019, the global shipments of smart home devices reached 168.6 million [7]. IDC predicts that from the first quarter of 2019 to the end of this year, the global market size of smart home will reach 840.7 million, and will grow to 1.46 billion sets by 2023, with a compound annual growth rate (CAGR) of 14.9 %[7]. Table 1 shows the number and market share of different types of smart home devices in 2019 and 2023. With the rapid growth of the number of smart homes, the amount of data that needs to be transmitted in the smart home network also grows rapidly. For users, smart home environment contains a lot of key and sensitive information. Therefore, it puts forward a higher requirement for data transmission [8]. The adversary’s attack on the smart home network system is generally divided into two types, one is “Passive Attack”, the other is “Active Attack”. In passive
attacks, eavesdropping, transmission, and monitoring techniques are used by the adversary [9]. They try to monitor the system and obtain system information without affecting system resources. The adversary does not modify the system information but learns something useful from it. Users’ mail, the messages and files exchanges by smart home terminal/server are all objects that can be monitored by adversaries. This type of attack is hard to detect because the adversary never changes the data, just learns something from it. In active attacks, adversaries try to change system resources or change system operations using information gathered in passive attacks [10]. Active attacks include information modification, effective system operations, and unauthorized access to system resources. When operating an active attack, the adversary would use a message modification attack to produce an unauthorized effect, this attack can modify the content of the real message or reorder the message sequence, and it also may delay the message. Due to the privacy issues in the smart home, people may concern about transmitting their sensitive information, and may hesitate to use smart devices and live in the smart home. It is necessary to design a secure and efficient privacy protection scheme for smart home network system. An effective security measure is to prevent adversaries from detecting the presence of sensitive information in smart homes network system. At present, the mechanism to solve the privacy leakage is mainly based on cryptography, such as data encryption, message authentication, integrity assurance and broadcast verification. Cryptographic-based encryption is to encrypt data into unreadable code, it is more likely to attract the attacker’s attention. Once these ciphertext is intercepted, the attacker will try to decrypt these ciphertext to obtain some useful information, so as to carry out some illegal activities. Some smart home users not only have high requirements for privacy protection, but also want to be able to express their preference in the decision-making of privacy protection, so they may want to define what information is sensitive to them and what is not.

To solve the above problems, this paper proposes a privacy protection framework based on information hiding, it guarantees the secure transmission of sensitive information. Before the data transmission, the model can classify the data into sensitive data and non-sensitive data according to the user’s preferences. Since non-sensitive data pose no threat to user privacy, it can be transmitted directly on ordinary channels without being processed. While, sensitive data need to be processed before it can be transmitted. In terms of sensitive data processing, this paper proposes a method of combining data encryption with information hiding. Sensitive data are encrypted using Data Encryption Standard (DES) encryption algorithm before it can be transmitted, making it to unreadable code. Then, a novel coverless information hiding method proposed, named the Top Three Most Significant Bit plain (TTMSB) information hiding method is used to provide a second guarantee for the security of sensitive data. In other words, the sensitive information is hidden in the multimedia carrier, so that the adversary can not notice the existence of sensitive information. TTMSB information hiding method proposed by this paper has well performance of robustness and concealing. It is suitable for data-centered sensor networks and can protect data security with lower computing overhead. The experiment shows the scheme proposed in this paper can realize the safe transmission of smart home data after they are classified into sensitive data and non-sensitive data according to users’ preferences; the combination of TTMSB information hiding method and EDS encryption can make up for the deficiency of data concealment transmission when using encryption technology only, it can realize the covert transmission of information effectively.

**Innovations of this paper:**

- This scheme enables users to define what is sensitive to them, and the model can classify the sensitive data and non-sensitive data accurately according to their preferences.
- The TTMSB information hiding method proposed in this paper is a novel coverless information hiding technology. Comparing to the traditional information hiding technology, it can realize higher performance of concealment for smart home sensitive data.
- The combination of DES encryption and TTMSB information hiding method changes the previous practice of exposing the ciphertext to the Internet directly, and realizes the double protection of data privacy in smart home network system. If an attacker tries to carry out a violent attack, it will have to break through at least two lines of defense, which seems impossible.

The arrangement of this paper is as follows: Section I introduces the development of smart home, the challenge of privacy protection and the solution proposed by this paper. Section II overviews the privacy protection research status of the smart home network. In section III, the application of the proposed scheme in smart home sensor network system is introduced. The experiments and analysis will be showed in section IV, and next is the conclusion in section V. The last part is acknowledgement in section VI.
II. RELATIVE WORK

In the past, to deal with privacy issues in smart homes network system, it was common to process all the data in a special way, such as using encryption. However, this approach will not only increase the cost of data processing, but also cause the delay during the period of data transmission. In order to reduce the unnecessary data processing, some of the data that is not sensitive to the user, needn’t to be processed before they are transmitted. A slightly better practice is that vendors set up privacy protection models according to their own experience, so as to reduce the amount of data processing. For example, although the smart home can adopts zero-conf configuration process, it has to force users to accept the vendors’ original settings [11]. These settings are not always satisfied by users, because the privacy settings provided by the vendor are not well representative of the user’s wishes. These problems are caused ultimately by users’ lack of control over data privacy. In view of this problem, many scholars have done many researches in recent years. Reference [12] is based on active learning, users need to mark some selected data before the data are classified by machine learning. Some test platforms [13] obtain the original data from different sensors in the building, processing high-level semantic information from these data, and is able to capture and enhance the privacy preferences expressed by the residents of the building. In [14], a fuzzy TOPSIS decision-making method based on demand response pricing scheme is proposed to consider the preferences of smart home users when using electrical equipment at the micro level, so as to better manage the usage of electrical equipment. Reference [15] classified data using pool-based active learning, it can help users reduce the burden of tagging instances, but there is no experimental proof in the paper. In order to make sure the smart home users fully express their wishes in privacy decisions, this paper adopts a data classification scheme based on support vector machines (SVM). This scheme can set privacy protection according to users’ preferences. Before the data transmission, the scheme can distinguish sensitive data from non-sensitive data according to the user’s labels, and the users do not need to have much knowledge about the technical aspect of smart devices.

When comes to the safe transmission of smart home information. Reference [16] adopts key management mechanism to protect users’ privacy, it can realize the self-protection of privacy to a certain extent. This kind of privacy protection scheme based on the idea of using onion route anonymous, or based on proxy server anonymous to hide users’ identity. Proxy servers often know the real identity of users, so personal home devices may leak user privacy. What’s more, this method has a complex system structure, which is the drawback of anonymous technology. The basic idea of using fake data sources is to select one or more sensor nodes to simulate the behavior of the real data source, so as to confuse adversaries [17]. In theory, the more fake sources there are, the more privacy is protected. However, false data sent by sensors frequently can increase the energy consumption of sensors and reduce the life span of sensors. Feng [18] used AES-128 encryption algorithm to encrypt smart home sensors data; In [19], WEI SHE et al. integrated homomorphic encryption with blockchain to realize the privacy protection of smart home. The above encryption algorithms can realize privacy protection of smart home to a certain extent, but they cost a lot of calculation, because of the complex algorithm and the large amount of information. Unfortunately, the computing and storage capacity of smart home devices are limited, limiting the application of the above encryption algorithm to smart home system.

In order to realize the covert transmission of sensitive information in the source-limited smart home network, it is not enough to rely solely on the key to protect the data. In particular, the readability of ciphertext is poor, it is easy to be noticed, intercepted and attacked [20], shown as figure 1. To solve this problem, a method of combination DES encryption with TTMSB information hiding method is proposed in this paper. In the scheme, sensitive data are hidden into the cover images, the attacker cannot notice the existence of sensitive data easily, shown as figure 1.

III. THE PROPOSED SCHEME IN THE SMART HOME NETWORK SYSTEM

A classic smart home system [21] is shown in figure 2. The smart home network model is generally composed of four parts, they are smart sensors, gateway node, user terminal and server. Smart sensors (such as temperature and humidity sensors, imaging sensor, healthy care sensors, etc.) are mainly responsible for collecting information of the environment and the health status of the family, and then transmitting the information to the gateway node through wired or wireless network. Gateway node is a powerful master node in the smart home network system [21], whose function is the bridge between the sensor node and the server. At general, the user terminal device is the user’s smart phone, through which the
user can remotely monitor the environment of the house or the health status of the family, as well as control the smart device. The server is a device with strong computing, communication and storage capabilities. The server is responsible for the following tasks: (a) processing and analyzing the data transmitted by the sensor, and feeding back the processing results to the smart home users. (b) user and smart device registration. (c) key management and maintenance.

As some information of the smart home is private to users, they do not want these data to be leaked, so it is necessary to put forward an effective privacy protection scheme. This paper proposes a scheme based on information hiding, it combines DES encryption algorithm with TTMSB information hiding method to protect the privacy of smart home system, during the sensitive data are transmitted. In the smart home network model, gateway is a node with relative strong computing and storage capacity, which can be used to classify data into sensitive data and non-sensitive data, encrypt sensitive data and hide them into cover images. Sensitive information is hidden into the cover image before it is transmitted by ordinary channels, it will not be found easily by the adversary. The server has a strong computing power, and it is completely trusted by the user. When it receives the data transmitted by the gateway, it extracts the ciphertext from the stego image and decrypts it into plaintext (original sensitive data). After that, the server calculates and analyzes the data and feeds the final result back to the user terminal.

The proposed privacy protection scheme is shown in figure 3. First, before the data are transmitted, they are classified into sensitive data and non-sensitive data according to the preference of users. The non-sensitive data can be transmitted directly on the ordinary channel without being processed.

While, the sensitive data is firstly encrypted, and then the ciphertext that present with garbled code state is hidden into the cover image. The stego image is same as the cover image, because it hasn’t been modified at all when hiding the sensitive data into the cover image, the attacker can hardly notice the presence of ciphertext in the stego image. After receiving the stego image, the receiver extracts the ciphertext from the stego image with the extraction key, and then decrypts the ciphertext with the decryption key to obtain the plaintext data.

![FIGURE 2. A classic smart home system.](image)

A. DATA CLASSIFICATION

In order to help the users express the definition of sensitive data better, this paper uses support vector machines (SVM) to classify data. SVM is a binary classification model [22], which was proposed by V.N. Vapnik, A.Y. Chervonenkis, c. Cortes et al in 1964. After 1990s, SVM developed rapidly and derived a series of improved and extended algorithms, they have been widely applied in pattern recognition, such as face recognition and text classification [23]. Its purpose is to find a hyper plane to segment the sample, and the principle of segmentation is to maximize the margin. Given a set of training samples \( D = \{x_1, y_1\}, \{x_2, y_2\}, \ldots, \{x_i, y_i\}\), \( y_i \in \{+1, -1\}\), the basic idea of classification learning is to find a dividing hyper plane in the sample space, it based on the training set \( D \) to separate the samples of different categories. In the sample space, the hyper plane division can be described by the following linear equation:

\[
W^T x + b = 0.
\]  

where, \( W \) is the normal vector, which determines the direction of the hyper plane, and \( b \) is the displacement, which determines the distance between the hyper plane and the origin. It is assumed that hyper plane can correctly classify training samples. For training samples \((x_i, y_i)\), the following formula is satisfied:

\[
\begin{cases} 
W^T x_i + b \geq +1 & y_i = +1 \\
W^T x_i + b \leq -1 & y_i = -1 
\end{cases}
\]  

As shown in figure 4, the sample points closest to the hyper plane are \( y_i(W^T x_i + b) = 1 \), and they are called “support vectors”. The dotted line is called the margin, and the distance between the two is called the margin.

\[
\gamma = \frac{2}{\|W\|}.
\]  

In order to find the division hyper plane with “maximum margin”, it is necessary to find the constraint hyper
parameters $W$ and $b$ in formula (2), so that $\gamma$ maximizes, that is:

$$\max_{W, b} \frac{2}{\|W\|} s.t. \ y_i(W^T x_i + b) \geq 1 \quad i = 1, 2, \ldots, m.$$  \hspace{1cm} (4)

It’s the basic form of SVM.

Data are classified before transmission. At first, users label the dataset according to their own preference, and then the classifier trains the classification model. In order to reduce the complexity of the classifier and give consideration to the universality of the classifier, this paper decides to adopt SVM as the classify model. The classification process is shown in algorithm 1: In the first line, users label the sensor dataset according to their preference. For example, if the information of health care is sensitive to the user, he only need to label “+1” to the dataset of the health care sensors. If the information of the environment is not sensitive to the user, he only need to label “-1” to the dataset of the environmental sensors. In line 2, generates a classifier after training the label dataset. In line 3, the classifier is input by Real-time sensor data, it can distinguish sensitive data and non-sensitive data. In line 4, returns sensitive data and non-sensitive data.

Algorithm 1 Data Classification

Input: Training dataset $D$, Real-time sensor data $R$
Output: S sensitive data $S$, Non-sensitive data $N_S$
1. $LabelDataset = \text{User labeling} (D)$
2. $Classifier = \text{Training}(LabelDataset)$
3. $S, N_S = \text{Classifier}(R)$
4. Return $S, N_S$

B. SENSITIVE INFORMATION ENCRYPTION AND INFORMATION HIDING

The combination of DES encryption and TTMSB information hiding method provides a double guarantee for the safe transmission of smart home sensitive data in the scheme of this paper. Because the smart home system focuses on lightweight, in order to reduce the complexity of data encryption operation and take the security of encryption into consideration, the scheme decides to use DES data encryption algorithm. DES is a symmetric cryptography system, it uses the same key during encryption and decryption. DES encryption algorithm is a block cipher algorithm [24], the block length is 64 bits, and the length of plaintext is same as ciphertext. The length of DES initial key is 64 bits, but the effective length is 56 bits, because the 8th, 16th, 24th, 32th, 40th, 48th, 56th and 64th bits are check bits, which do not participate in the encryption operation. Figure 5 shows the basic structure of DES, where the right part shows the process of using the 56-bit key. The 64-bit initial key decreases to 56-bit after the operation of permutation selection 1, and the 56-bit key changes to 48-bit subkeys $K_i \ (i = 1, 2, \ldots, 16)$ after the operation of left shift and permutation selection 2, they will be used during each iteration calculation.

The left side of figure 5 shows the encryption process for DES, it has three main phases:

1) The 64-bit plaintext is rearranged by $IP$ permutation and is divided into left block $L_0$ and right block $R_0$, the length of each block is 32-bit;
2) With the operation of subkeys $K_i \ (i = 1, 2, \ldots, 16)$, there are 16 round iterations carried out for the two blocks. Each iteration will conclude the process of permutation and substitution. The output of the last
iteration is 64-bit, which is the pre-output that generated by the interchange of the last left and last right blocks;

3) At last, the final pre-output generates a 64-bit ciphertext by the operation of \( IP^{-1} \) permutation.

The encryption process can be simplified to:

\[
L_0R_0 \leftarrow IP(<64 - bit\ plaintext>) \quad \text{at i = 1, 2, \ldots, 16}
\]

\[
R_i \leftarrow L_{i-1} \bigoplus F(R_{i-1}, K_i) \quad \text{at i = 1, 2, \ldots, 16}
\]

\[
<64 - bit\ ciphertext> \leftarrow IP^{-1}(R_{16}L_{16})
\]

Decryption process is the reverse process of encryption process:

\[
R_{16}L_{16} \leftarrow IP(<64 - bit\ ciphertext>) \quad \text{at i = 16, 15, \ldots, 1}
\]

\[
L_{i-1} \leftarrow R_i \bigoplus F(L_i, K_i) \quad \text{at i = 16, 15, \ldots, 1}
\]

\[
<64 - bit\ plaintext> \leftarrow IP^{-1}(R_0L_0)
\]

DES encryption speed is fast, according to the statistics of [25], when encrypting a 64-bit plaintext, encryption time is about 0.000137 seconds. DES encryption algorithm is simple, and the computation cost is small, so it is suitable for lightweight electronic units such as smart home sensor. Although DES has the shortcoming that plaintext data is not easy to hide, information hiding algorithm can just make up for this shortcoming, so DES encryption can be combined well with information hiding.

As for information hiding, this paper proposes a novel coverless information hiding technology [26], named the Top Three Most Significant Bit plan (TTMSB) information hiding method. It uses double keys to map the secret information with the top three bits of the cover image fragment average pixel, so as to realize the secret information hiding. Since the process of information hiding does not make any changes to cover image, the secret information will not be found by the adversary during the transmission of stego image. The algorithm framework of TTMSB information hiding method is shown in figure 6.

![FIGURE 6. The algorithm framework of TTMSB information hiding method.](image)

As well known, a digital image is made up of pixels, each of pixel has a grayscale value. The larger the grayscale value, the higher the brightness. Pixels present in the form of two-dimensional matrix elements in computer, and the gray value of pixels is usually range from 0 to 255. The gray value of a pixel can be represented by an 8-bit binary number. The bit plane on the far left is called the Most Significant Bit plane (MSB), and the bit plane on the far right is called the Least Significant Bit plane (LSB). TTMSB information hiding method is one kind of coverless information hiding technologies. It uses the double secret keys to build the mapping relation between the secret information and the average pixel values of cover image fragment, achieving the purpose of information hiding. It is worth noting that this algorithm only uses the pixel value of the image fragment to represent the secret information, when the secret information is hidden. Therefore, it doesn’t make any changes to cover images. In other words, the cover image is the same as the stego image. Adversaries cannot find secret information hidden in the stego image. The specific implementation process of this algorithm is shown as follows:

1) The sender converts the secret information text \( T \) to the binary digits \( B_1, B_2, \ldots, B_n \);

2) The sender divides the cover image into image fragments \( I_1, I_2, \ldots, I_m \), whose pixel is 5 * 5;

3) Calculate the average pixel values \( V_1, V_2, \ldots, V_m \) of \( I_1, I_2, \ldots, I_m \);

4) Map the \( B_1, B_2, \ldots, B_n \) with the top three most bit planes of \( V_1, V_2, \ldots, V_m \) according to the random order, where \( n \leq m \). The random order is called coordinate key \( K_1 \), which is decided by the sender and receiver in advance. The result of mapping is to generate the location key \( K_2 \). For example, suppose \( B_i \) and \( I_j \) were mapping according to \( K_1 \). If \( B_i \) is “1” and the top three bits of \( I_j \) are “001”, then \( B_i \) matches the third bit of \( I_j \), \( K_2 = 3 \), so the third bit of \( I_j \) represents \( B_i \). If \( B_i \) and \( I_j \) don’t match, then \( K_2 = 0 \). In this case, if \( I_j \) is “000”, then “1” is used to represent \( B_i \); if \( I_j \) is “111”, then “0” is used to represent \( B_i \);

5) The sender sends \( K_2 \) and Stego images to the receiver. Actually, the stego image is just the cover image, there’s no difference between them.

After Real-time sensor data are classified into sensitive data and non-sensitive data by the classifier. The non-sensitive data are sent to the buffer, which packages the non-sensitive data into packet payload, and when the buffer is full, the non-sensitive data are sent out. Sensitive data are encrypted in the DES encryption component, and then the ciphertext is hidden into the cover image in the TTMSB information hiding component. The process of encryption and information hiding are shown as algorithm 2. First, the sensitive data is encrypted by DES encryption technology using encryption \( Ke \) in line 1. In line 2, ciphertext \( C \) is converted to binary bit \( C' \). In line 3, \( C' \) is hidden into the cover image using the hidden key \( K_1 \), return Location key \( K_2 \) and Stego image \( I' \). Next, the stego image \( I' \) is sent to the buffer, as in line 4. In line 5, determine whether the buffer is full, and if so, send the packet payload out; If not, repeat step 1 to step 5.
Algorithm 2 Encryption and Information Hiding

| Input: Sensitive data S, Encryption key Ke, Coordinate key K1, Cover image I |
| Output: Stego image I', Location key K2 |

1. $C = \text{encrypt}(S, Ke)$
2. $C' = \text{char 2 bit}(C)$
3. for $i = 0$ to $\text{len}(C') - 1$
   
   $I' = \text{hide}(C', K1, I)$
   
   return $K2, I'$
4. Send $I'$ and $K2$ to the buffer
5. The buffer is full? send the data: repeat the step 1 to 5

C. SENSITIVE INFORMATION EXTRACTION AND DECRYPTION

The extraction and decryption are shown as algorithm 2, they are the reverse processes of information hiding and encryption. Using the coordinate key $K1$ and the location key $K2$, the binary bit $C'$ of the sensitive data is extracted from the stego image $I'$ in line 1. In line 2, the binary bit $C'$ is converted to ciphertext $C$. Then, the decryption key $Ke$ is used to decrypt ciphertext $C$ into the original sensitive data $S$, as in line 3. Finally, the sensitive data $S$ is returned, as in line 4.

Algorithm 3 Information Extraction and Decryption

| Input: Stego image I', Decryption key Ke, Coordinate key K1, Location key K2 |
| Output: Sensitive data S |

1. for $i = 0$ to $\text{size}(I')$
   
   $C' = \text{extract}(I')$
   
   return $K2, I'$
2. $C = \text{bit 2 char}(C')$
3. $S = \text{decrypt}(C, Ke)$
4. Return $S$

IV. EXPERIMENTS AND ANALYSIS

A. DATASET

The dataset in this paper is from the UMass Trace Repository, a data-sharing platform operated by the advanced systems software laboratory at the university of Massachusetts. The dataset includes the data of healthcare sensor, temperature sensor, ambient light sensor, wind speed and direction sensor, rainfall sensor and so on. In this experiment, 80% of the dataset is randomly selected as the training set, while 20% of the dataset is verification set. The experimental environment is “python3.6+cuda9.0+cudnn7.0+tensorflow1.13.1+MATLAB R2016a”.

B. EXPERIMENTAL PROCESS AND RESULTS

Firstly, the sensor data is classified into sensitive data and non-sensitive data. Before the classification, users can label the training dataset according to their preferences. The results of classification are shown in figure 7.

Secondly, the sensitive data are encrypted. Suppose we use the key “98 d0 61 ed c0 57 11 59” to encrypt the sensitive data. The result of DES encryption is shown in figure 8.

The ciphertext is present as unreadable code, as showed in figure 8. If it is directly spread on the Internet, it is easy to be noticed and attacked by the adversary. Therefore, it is necessary to hide the ciphertext into the carrier. In this experiment, we propose TTMSB information hiding method to realize information hiding, and the images are used as information hiding carrier. After the information hiding, the stego image is not changed. In other words, the cover image is same as the stego image, as shown in figure 9.

C. SAFETY ANALYSIS

Although brute-force attack is considered to be one of the most “stupid” analysis methods, in addition to brute-force attack, there has not been a complete decoding to DES. The security of DES algorithm is high, if someone uses brute-force attack to decrypt ciphertext, they will require a
In the TTMSB information hiding method, this paper adopts the cover image of 64 * 64 pixels to hide sensitive information. The order of hiding is according to the order of coordinate key \( K_1 \) and the order location key \( K_2 \). There are \(((64*64)/(5*5))! = 2.0 \times 10^{291}\) possibilities to \( K_1 \), and 4 possibilities to \( K_2 \). Taken together, there are \(4 \times 2.0 \times 10^{291} = 8 \times 10^{291}\) possibilities to hide sensitive information. Using a brute force extraction, assuming an ordinary computer could perform \(10^8\) calculations per second, it would cost \(6.3 \times 10^{275}\) years to extract secret information. If the cover image with larger pixels is used as the carrier of information hiding, it will take more time to extract secret information by brute force extraction. So it proves that TTMSB information hiding method has good security performance.

Based on the security analysis of the above two algorithms, our security strategy has the characteristics of high concealment, high difficulty in cracking and high neglect, so the security of our proposed scheme is extremely high.

**D. ROBUSTNESS ANALYSIS**

In the scheme of this paper, the sensitive data of smart home are not transmitted directly in the Internet, but indirectly through stego images and \( K_2 \). In the process of transmission, the image is always attacked by noise, which makes the image distorted. The most innovative part of the scheme is to use TTMSB algorithm to hide sensitive data of smart home, and the carrier of TTMSB information hiding is image, so it is necessary to analyze the robustness of image anti-noise in this algorithm. In this paper, gaussian noise and salt-and-pepper noise are used as noise attacks on stego images, and the final experimental results are averaged for multiple tests. Bit error rate (BER) is used as the criterion to judge robustness of TTMSB information hiding method. BER is defined as:

\[
    BER = \frac{m}{n} \times 100\%.
\]

where, \( m \) represents the number of error bits extracted from the stego image, and \( n \) represents the total number of bits extracted from stego image.

In the gaussian noise attack test, let \( \mu = 0 \), \( \sigma^2 \) from 0.1 to 1. Table 2 shows the comparison of TTMSB, CSD[28] and CBZS[29] in their ability to resist gaussian white noise attack. When \( \sigma^2 = 0.1 \), the BER of TTMSB information hiding method is the smallest, about 12%. When \( \sigma^2 = 1 \), the BER of TTMSB information hiding method is the largest, about 38%, and the BER increases with the increase of noise. The TTMSB information hiding method performs better than the other two algorithms when the noise density is low. However, it has no advantage when the noise density is high. Considering to the practical application, in the case of large noise attack, the receiver can ask the sender to resend the stego image. Therefore, it is of no practical significance to consider large noise.

In the experiment of resisting salt-and-pepper noise, the intensity of salt-and-pepper noise increase from 0.01 to 0.1. Table 3 illustrates the comparison of TTMSB, CBD[30] and CSD in their ability to resist salt-and-pepper noise. When the noise intensity of salt-and-pepper is lower than 0.05, the BER value is always 0. When the noise intensity of salt-and-pepper is larger than 0.05, the BER value also presents an increasing trend with the increase of the noise intensity, but the BER value does not exceed 4.6%. It can be seen from the data in table 3 that TTMSB information hiding method is robust, and its robustness is significantly better than the other two algorithms.

**E. CAPACITY ANALYSIS**

In the scheme, the transmission of smart home sensitive data is realized by transmitting stego image. Therefore, the more sensitive information a stego image can contain, the better. In other words, increasing the hidden capacity can reduce the number of times sending sensitive information, it is of great significance for reducing the energy consumption of smart home systems. In this paper, the number of bits hidden in per carrier (bits*carrier\(^{-1}\)) is used as the criterion to judge the hidden capacity. In the experiment of this scheme, the pixels of cover image is 64*64, and 1 bit sensitive data is hidden in one cover image fragment, its pixel is 5*5. Thus, each image can hide \((66*64)/(5*5) = 163\) bits. Table 4 shows the comparison of TTMSB method, Zhou’s method [27], Yuan’s method [31], Zheng’s method [32] and Zou’s method [33] in their ability to hide information. As can be seen from table 4, the information hidden capacity of TTMSB method is 20 times that of Zhou’s method and Yuan’s method, 9 times that of Zheng’s method and 2 times that of Zou’s method.

| BER of TTMSB, CSD and CBZS for different \( \sigma^2 \) values. |
|---|---|---|
| \( \sigma^2 \) | CSD | CBZS | TTMSB |
| 0.1 | 18% | 14% | 12% |
| 0.2 | 19% | 24% | 15% |
| 0.5 | 21% | 25% | 33% |
| 0.6 | 21% | 26% | 35% |
| 0.9 | 22% | 31% | 37% |
| 1.0 | 22% | 33% | 38% |

| The BER of TTMSB, CBD and CSD for different noise intensity of salt-and-pepper. |
|---|---|---|
| Noise intensity | CBD | CSD | TTMSB |
| 0.01 | 2.8% | 2.6% | 0% |
| 0.03 | 4.0% | 4.9% | 0% |
| 0.05 | 4.7% | 6.2% | 0% |
| 0.07 | 5.9% | 7.1% | 1.1% |
| 0.09 | 7.1% | 8.2% | 3.1% |
| 0.1 | 12.8% | 9.3% | 4.6% |
TTMSB information hiding method has great advantage in the transmission of smart home sensitive data.

V. CONCLUSION

With the rapid development of smart home system, users are eager to have a safe and efficient privacy protection scheme. In terms of privacy protection, the current practice based on data encryption is easy to attract the attention of attackers’. For these reasons, this paper proposes a privacy protection scheme combining DES encryption with TTMSB information hiding method. In the scheme, users can also express their preference to the privacy settings. This scheme classifies the smart sensor data into sensitive data and non-sensitive data, then encrypts the sensitive data and hides the ciphertext into the cover image. The receiver obtains the hidden information by coordinate key $K_1$ and location key $K_2$, and decrypts the ciphertext by the decryption key. TTMSB information hiding method combined with DES encryption method provides a double guarantee for the safe transmission of smart home data. After the sensitive information is hidden, there is no change to the stego image, so the adversary cannot notice the existence of secret information. The feasibility of the scheme has been verified by the simulation experiment. Theoretical analysis shows that the scheme with high hidden capacity is secure and can resist the attack to some extent.

ACKNOWLEDGMENT

Thanks to Prof. L. Yang for her careful help to the proposal of the scheme, as well as the guidance of the experiment. As to paper writing, thanks for the guidance of their schoolmate X. Dang.

REFERENCES

[1] D. Evans, “The Internet of Things: How the next evolution of the Internet is changing everything,” CISCO, San Jose, CA, USA, White Paper, 2011, vol. 1, pp. 1–11.
[2] J. Gantz and D. Reinsel, “The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east,” IDC iView, IDC Analyze Future, vol. 2007, pp. 1–16, Dec. 2012.
[3] V. Turner, J. F. Gantz, D. Reinsel, and S. Minton, “The digital universe of opportunities: Rich data and the increasing value of the Internet of Things,” IDC iView, Sponsored EMC, vol. 1, pp. 1–10, Apr. 2014. [Online]. Available: https://www.emc.com/leadership/digital-universe/index.html
[4] S. Li, L. Da Xu, and S. Zhao, “5G Internet of Things: A survey,” J. Ind. Inf. Integ., vol. 10, pp. 1–9, Jun. 2018.
[5] E. P. Yadav, E. A. Mittal, and H. Yadav, “IoT: Challenges and issues in indian perspective,” in Proc. 3rd Int. Conf. Internet Things, Smart Innov. Usages (IoT-SIU), New Delhi, India, Feb. 2018, pp. 1–15.
[6] A. Patel and T. A. Champaneria, “Fuzzy logic based algorithm for context awareness in Iot for smart home environment,” in Proc. IEEE Region 10 Conf. (TENCON), Nov. 2016, pp. 1057–1060.
[7] J. Ubriani and A. Wright, “In the Q1 of 2019, global smart home devices reached 168.6 million, up 37.3% year-on-year,” IDC iView, vol. 1, pp. 1–3, Jun. 2019. [Online]. Available: http://www.99it.com/archives/899123.html
[8] W. Ai, G. Dustgeer, M. Awaï, and M. A. Shah, “IoT based smart home: Security challenges, security requirements and solutions,” in Proc. 23rd Int. Conf. Autom. Comput. (ICAC), Sep. 2017, pp. 16–26.
[9] M. Abomhara and G. M. Koien, “Security and privacy in the Internet of Things: Current status and open issues,” in Proc. Int. Conf. Privacy Secur. Mobile Syst. (PRISMS), May 2014, pp. 1–8.
[10] M. Chen, “Towards smart city: M2M communications with software agent intelligence,” Multimedia Tools Appl., vol. 67, no. 1, pp. 167–178, 2013.
[11] K. J. Kaaz, A. Hoffer, M. Saiedi, A. Sarma, and R. B. Bobba, “Understanding user perceptions of privacy, and configuration challenges in home automation,” in Proc. IEEE Symp. Vis. Lang. Hum.-Centric Comput. (VL/HCC). Washington, DC, USA: IEEE Computer Society, Oct. 2017, pp. 297–301.
[12] P. Pappachan, M. Degeling, R. Yus, A. Das, S. Bhagavatula, W. Melicher, P. E. Naeini, S. Zhang, L. Bauer, A. Kobsa, S. Mehrotra, N. Sadeh, and N. Venkatsubramanian, “Towards privacy-aware smart buildings: Capturing, communicating, and enforcing privacy policies and preferences,” in Proc. IEEE 37th Int. Conf. Distrib. Comput. Syst. Workshops (ICDCSW), Jun. 2017, pp. 193–198.
[13] S. Mehrotra, A. Kobsa, N. Venkatsubramanian, and S. R. Rajagopalan, “TIPPERS: A privacy cognizant IoT environment,” in Proc. IEEE Int. Conf. Pervs. Comput. Commun. Workshops (PerCom Workshops), Mar. 2016, pp. 1–6.
[14] O. A. Sianaki and M. A. S. Masoum, “A fuzzy TOPSIS approach for home energy management in smart grid with considering householders’ preferences,” in Proc. IEEE PES Innov. Smart Grid Technol. Conf. (ISGT), Feb. 2013, pp. 1–6.
[15] M. Keshavarz and A. Anwar, “Towards improving privacy control for smart homes: A privacy decision framework,” in Proc. 16th Annu. Conf. Privacy, Secur. Trust (PST), Washington, DC, USA: IEEE Computer Society, Aug. 2018, pp. 1–3.
[16] Y. Guo, “Design and application of Internet of Things privacy protection model for smart home,” M.S. thesis, School Softw., Beijing Univ. Technol., Beijing, China, 2012.
[17] Y. Yang, M. Shao, S. Zhu, and G. Cao, “Towards statistically strong source anonymity for sensor networks,” ACM Trans. Sensor Netw., vol. 9, no. 3, pp. 1–23, May 2013.
[18] J. Feng, “Research and realization of secure communication technology in intelligent home system,” M.S. thesis, Harbin Inst. Technol., Harbin, China, 2016.
[19] W. She, Z.-H. Gu, X.-K. Lyu, Q. Liu, Z. Tian, and W. Liu, “Homomorphic consortium blockchain for smart home system sensitive data privacy preserving,” IEEE Access, vol. 7, pp. 62058–62070, 2019.
[20] J. Yang, “Research on security enhancement of information hiding in sensor networks,” M.S. thesis, Hunan University, Changsha, China, 2011.
[21] R. Spreitzer, V. Moonsamy, T. Korak, and S. Mangard, “Systematic classification of side-channel attacks: A case study for mobile devices,” IEEE Commun. Surveys Tut., vol. 20, no. 1, pp. 465–488, 1st Quart., 2018.
[22] Z. Zhou, “Support vector machines,” in Machine Learning. Beijing, China: Tsinghua Univ., 2016, pp. 121–139.
[23] J. Qin and Z.-S. He, “A SVM face recognition method based on Gabor-featured key points,” in Proc. Int. Conf. Mach. Learn. Cybern., 2005, pp. 5144–5149.
[24] Z. Cheng, C. Chen, and X. Qiu, “Cache timing attack based on Flush+Reload DES algorithm,” Comput. Eng., vol. 44, pp. 163–167, Dec. 2018.
[25] J. Zhu, “Research and implementation of QR code anti-counterfeiting algorithm based on DES improved algorithm,” M.S. thesis, Dept. Sci. Technol., Shantou Univ., 2018.
[26] Z. Zhou and Y. Luo, X. Xiang, Y. Tan, and H. Huang, “Coverless image steganography: A survey,” IEEE Access, vol. 7, pp. 171372–171394, 2019.
[27] Z. Zhou, H. Sun, R. Harit, X. Chen, and X. Sun, “Coverless image steganography without embedding,” in Proc. Int. Conf. Cloud Comput. Secur. Chan, Switzerland: Springer, 2015, pp. 123–132.
[28] J. Wu, X. Fei, and N. Wang, “Study on image zero steganography algorithm based on chaotic sequence and DCT transform,” Electron. Meas. Techn., vol. 40, pp. 180–185, May 2017.
[29] M. Bilal, S. Imitiaz, W. Abdul, S. Ghouzali, and S. Asif, “Chaos based zero-steganography algorithm,” Multimedia Tools Appl., vol. 72, no. 2, pp. 1073–1092, 2014.

TABLE 4. The hidden capacity of TTMSB method, Zhou’s method, Yuan’s method, Zheng’s method and Zou’s method.

| Methods      | Capacity(bits*carrier$^{-1}$) |
|--------------|--------------------------------|
| Zou’s method | 8                              |
| Yuan’s method| 8                              |
| Zheng’s method| 18                            |
| Zou’s method | 30                             |
| TTMSB method | 165                            |
[30] S. Singh and T. J. Siddiqui, “A security enhanced robust steganography algorithm for data hiding,” Int. J. Comput. Sci. Issues, vol. 9, no. 3, pp. 131–139, 2013.

[31] C. Yuan, Z. Xia, and X. Sun, “Coverless image steganography based on SIFT and BOF,” Internet Technol., vol. 18, no. 2, pp. 435–442, 2017.

[32] S. Zheng, L. Wang, B. Ling, and D. Hu, “Coverless information hiding based on robust image hashing,” in Proc. Int. Conf. Intell. Comput., 2017, pp. 536–547.

[33] L. Zou, J. Sun, M. Gao, W. Wan, and B. B. Gupta, “A novel coverless information hiding method based on the average pixel value of the sub-images,” Multimedia Tools Appl., vol. 78, no. 7, pp. 7965–7980, Jul. 2018.

LINA YANG received the B.S. degree in computer engineering from Shijiazhuang Railway University, Shijiazhuang, China, in 2005, the M.Eng. degree in computer science from the University of Malaya, Kuala Lumpur, Malaysia, in 2011, and the Ph.D. degree from the University of Macau, Macau, in 2015. She is currently a Lecturer with the School of Computer, Electronics and Information, Guangxi University, Nanning, China. Her research interests focus on medical biometrics, machine learning, image processing, and information security.

HAIYU DENG received the B.S. degree in thermal energy and power engineering from Shandong University, Shandong, China, in 2015. He is currently pursuing the M.S. degree with the School of Computer, Electronics and Information, Guangxi University, Nanning, China. His research interests focus on information security, image processing, and machine learning.

XIAOCUI DANG received the B.S. degree in computer science and technology from Jining University, Shandong, China, in 2017. She is currently pursuing the M.S. degree in computer science and technology with Guangxi University, Nanning, China. Her research interests focus on information security, image processing, machine learning, and bioinformatics.