electronics

MDPI

Article

A Hybrid Cryptography Scheme for NILM Data Security

Ruijue Feng [1], Zhidong Wang [2,*], Zhifeng Li [1], Haixia Ma [1], Ruiyuan Chen [3], Zhengbin Pu [2], Ziqiu Chen [1] and Xianyu Zeng [1]

[1] School of Electrical Engineering, Guangzhou College of South China University of Technology, Guangzhou 510800, China; fengrj@gcu.edu.cn (R.F.); lizf@gcu.edu.cn (Z.L.); mahx@gcu.edu.cn (H.M.); ziqiu918@gmail.com (Z.C.); zengxianyu2020@gmail.com (X.Z.)

[2] Research Center of Smart Energy Technology, School of Electric Power, South China University of Technology, Guangzhou 510640, China; puzhengbin@gmail.com

[3] College of Engineering, South China Agriculture University, Guangzhou 510642, China; 201721010103@stu.scau.edu.cn

* Correspondence: zdwang@scut.edu.cn; Tel.: +086-1353-504-0794

Received: 11 June 2020; Accepted: 8 July 2020; Published: 10 July 2020

Abstract: Using fine-grained data analysis, non-invasive load monitoring (NILM) can reveal the detail of electricity customers' habits, which is helpful in the improvement of refined management and better user experience. However, the possibility of electricity customers' privacy leak is also gradually increasing, and the security of NILM data has become a priority problem to be solved. To protect the privacy disclosure of NILM data, this paper analyzes the NILM privacy leak problems and ways in which information leak occurs faced by NILM data. On the basis of the comprehensive survey of cryptographic algorithms to choose the most appropriate data security method for NILM, a hybrid cryptography scheme was proposed to protect the data security. In the scheme, symmetric algorithm AES (Advanced Encryption Standard) was used to encrypt data for high efficiency, and asymmetric algorithm RSA (Rivest-Shamir-Adleman) was used to encrypt AES key for identity authentication. The classical algorithm HMAC-SHA1 (Hash Message Authentication Codes-Secure Hash Algorithm 1) was further developed to guarantee the integrity of data. By transplanting the algorithm into STM32 MCU (STMicroelectronics 32 bit Micro Controller Unit) for performance test and using Visual studio 2017 + QT tools to develop the test interface, one optimal operation mode was selected for the scheme. At the same time, the effectiveness of the scheme was verified, and the scheme computing cost depended on the efficiency of encryption and decryption, or signature and verification of the RSA algorithm.

Keywords: NILM; privacy; cryptography; hybrid cipher algorithm

1. Introduction

Non-invasive load monitoring (NILM) was first proposed by Hart in the 1980s, and it only needs to install one monitoring equipment at the power entrance, which can analyze the operation state of all connected electrical devices. By decomposing the electric quantity of the monitored equipment, the type and operation of the single load for each electrical device in the load group can be obtained [1]. Accurate electricity information on each/type of equipment has great significance for power companies to optimize grid planning and operation. It will also help electricity customers to realize intelligent power consumption and the whole society to implement the awareness of ecological civilization into specific actions. Thanks to the advantages of NILM technology, such as convenient implementation, less equipment cost, wide application range, and so on, it has gradually been paid attention to by the power academic and industrial community.

NILM technology is conducive to mastering the customer's electricity use behavior and improving the user's electricity use habits and experience. However, it is also a double-edged sword in the electricity customers' electricity information being more transparent and easy to be exposed. There will be a huge risk of public social security if customers' electricity behavior information is exposed or even maliciously used. Consumer information leakage caused by the user's payment, electricity, and other behaviors has gradually aroused widespread concern and become the focus of all countries. Taking mainland China as an example, relevant research of smart grid privacy protection has been attached great importance, which is fundamentally guaranteed in the form of policies and regulations. In June 2017, "network security law of the people's Republic of China" was officially implemented. The promulgation of this law has provided a legal basis for protecting important data of key infrastructure and citizens' personal privacy information. In the same year, the state internet information office issued the "report on the construction and development of Digital China", which proposed to strengthen the network security protection of information infrastructure. Violations of citizens' personal privacy and other criminal acts were severely cracked down on.

In addition to policies and regulations, some progress has been made in technical protection. In the current smart grid, there are roughly four kinds for common methods to protect the electricity customers' privacy in the real-time communication and data interaction between power supplier and electricity customers. The first one is the data aggregation scheme. In references [2–4], the scheme is based on homomorphic encryption, which solves the risk of the gateway being hijacked or eavesdropping in the communication link when the original data of the smart meter are uploaded to the regional gateway. In references [5,6], an improved method is proposed to solve the problems of fault tolerance, security, and communication protocol in data aggregation. The second one is identity anonymity technology. In reference [7], one anonymous certificate protocol based on Camenisch–Lysyanskaya signature certificate is proposed. This protocol was linkable, and the control center can verify the power consumption data and track the bad data without knowing the real identity of the user. In reference [8], on the basis of the remote anonymous authentication technology in trusted computing, the concept of trusted smart meters is proposed to protect users' private information from being affected. The third one is the privacy protection method based on incentive demand response. This kind of method is to compensate the user's privacy loss to a certain extent by designing one reasonable incentive mechanism. Its essence is to trade the user's privacy as a kind of marketable commodity [9,10]. The fourth one is a cryptographic scheme for secure communication in open channels [11,12]. In reference [13], on the basis of the hybrid cryptographic algorithm, a new scheme of authentication and key protocol is proposed to ensure the safe operation of information.

The above four kinds of methods are mainly applied to the data of smart meters. At present, the network transmission frequency of smart meter data is relatively low (commonly 15 min). Compared with the NILM technology with fine-grained data, the amount of electrical data is much less, and the real-time requirements for security algorithms are relatively less stringent. Furthermore, the above security methods rarely cover all the main goals of power information security, such as accessibility, integrity, and confidentiality. The fine-grained data analysis of NILM technology can fully grasp the user behavior by analyzing the user's power consumption. NILM technology should pay more attention to the privacy behavior of electrical users. Unfortunately, current NILM technology research is majorly focused on event detection, feature extraction, and load identification [14–18]. There are only a few studies on the communication and information processing in the NILM scenario. In order to solve the communication congestion between smart meters and cloud computing centers, reference [19] propose an improved NILM method based on IP and device characteristics. In reference [20], an NILM system is designed for real-time remote monitoring of load operation status and identification of load types. However, these NILM studies on the communication and information processing do not provide complete research on NILM information security. To the best of our knowledge, there is still a lack of the specific public literature on NILM communication and information security.

Aiming at the fact that NILM technology is faced with a greater possibility of power consumption information disclosure, this paper used a cryptographic algorithm to implement the

data privacy protection scheme suitable for NILM [21,22]. In the scheme, the main technologies of the security algorithm, including key management, identity authentication, and integrity and efficiency, were considered. Through the test of the proposed scheme on STM32 single-chip microcomputer platform, an optimal operation mode was selected for the scheme. The validity and computing cost of the scheme were finally verified.

A. Contributions of the paper

At present, NILM technology mostly focuses on the research of power load identification and decomposition, without public literature study on the NILM data security method. The major contributions of this paper can be summarized as follows:

- First, we seek a method suitable for NILM data, which mainly belongs to the field of practical engineering application innovation.
- Second, after detailed analysis on the different classical cryptography techniques applied to NILM data, we conclude the most suitable scheme for NILM data.
- Finally, we present a performance analysis of security algorithms suitable for NILM data through practical examples.

B. Organization of the paper

This paper is organized as follows. In Section 2, we analyze the characteristics and application of NILM technology, and point out that its data collection volume is tens of thousands of times of the current smart meter. In Section 3, we analyze the privacy leakage of NILM data and the manner in which hackers attack data. In Section 4, we analyze the application of different encryption and decryption algorithms in data protection from the perspective of cryptography, and proposed one hybrid cryptographic scheme for the application of NILM system. In Section 5, we test the performance of the hybrid cryptographic scheme and verify the effectiveness and the computing cost of the scheme. Section 6 presents the conclusions.

2. Technology of Non-Intrusive Load Monitoring

In the process of smart grid construction, it is important for reasonable dispatching and efficient utilization of power to get fine-grained data of electricity customers, which is conducive to fine electricity management. As one of fine management key technologies, NILM identifies the type and operation situation of each electrical device in the load concentration group by installing monitoring equipment at the user's power supply entrance to monitor the voltage and current signal and using a pattern recognition algorithm. For example, working period, power size, and electricity consumption of various types of electric load, such as refrigerator, air conditioner, washing machine and so on, can be perceived in real time through NILM. A large number of fine-grained electro-data are transmitted to the data processing center of the power supplier. After data mining, extraction, and analysis, it provides important data support for smart power consumption behavior analysis, energy conservation service demand response, ladder electricity, electricity price system, and other refined power consumption businesses. Moreover, the abnormal power consumption behavior of users can be effectively mined, such as stealing electricity, leakage, abnormal power off, and so on. Thus, the safe operation and economic benefits of the power grid together with the consumer can be ensured. On this basis, electricity users can independently analyze household energy consumption points by feeding back electricity consumption information to power suppliers, and independently adjust and optimize electricity consumption behavior [23,24]. Compared with a traditional intelligent ammeter, which only realizes total active power detection and electric energy measurement, NILM is more intelligent and further optimizes the user's power experience.

NILM can realize the perception of electricity load working state and energy consumption level information by analyzing metering data resources. The NILM field acquisition equipment mostly uses an embedded controller, and its storage space and computing capacity are relatively limited. In order to achieve load type full identification and energy full decomposition, electricity customers should send real-time electro-data to electric power company for operation and management via the communication network, as shown in Figure 1. At present, sampling frequency of real-time data in

the NILM system is divided into high frequency and low frequency [25,26]. High frequency usually achieves a better load identification effect, but also puts forward higher requirements for data processing capacity. REDD (reference energy disaggregation dataset) was led and pushed out by Zico Kolter from Massachusetts Institute of Technology (MIT) to analyze the collected data with a high frequency of 16.5 KHz, which was familiar with the industry [27]. Leading high frequency sampling companies, such as Sense energy company, equip high performance sample devices, where the sampling frequency is from KHz to MHz, and collected data in one sinusoidal ac power cycle are as high as 1000. High frequency sampling data of NILM also mean that a large amount of power data will be generated continuously. However, owing to the limited performance of on-site monitoring equipment, these large amounts of data need to be transmitted to remote power companies for processing. In the process of data transmission, it may go through the data collection of multiple different nodes and communication networks.

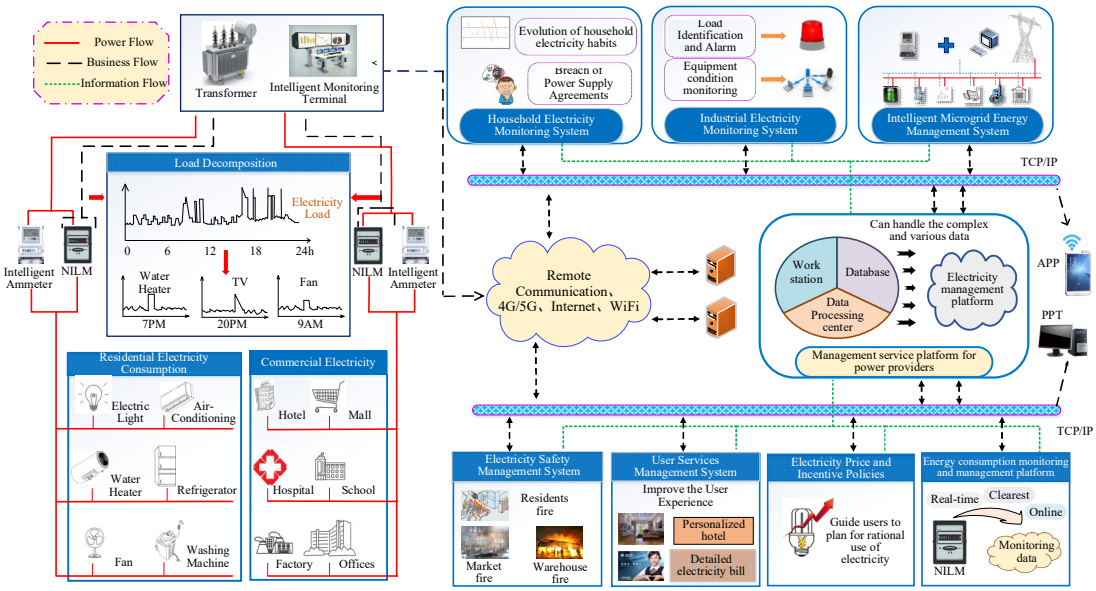


Figure 1. Non-intrusive load monitoring (NILM) system.

3. NILM Data Transmission Security

The process of NILM data transmission from the electricity customers' house to the power company brings a great challenge to network communication transmitting NILM data. On the one hand, the communication network needs to have enough bandwidth to meet the transmission requirements of large NILM data. With the gradual application of 4G, 5G, and other high-speed communication networks, the broadband problem of nm big data is expected to be solved. On the other hand, the security problem in the process of NILM data transmission will become increasingly prominent. The main power grid (namely transmission and substation system) generally adopts an independent power dedicated communication network, which is mainly based on optical fiber network and is not shared with the outside world, so the security of the information transmitted by the dedicated communication network of the main power grid is relatively good. However, because the distribution network has the characteristics of a wide distribution range with too many nodes, the cost of laying a fiber optic network for communication is extremely high. The current solution is to use a flexible communication method by using various networks, including RS 485/422, GPRS (General Packet Radio Service), ZigBee, 4G, and 5G networks. However, it has also brought some problems compared with the use of the power dedicated network; that is, the information security problem of NILM data transmission on the distribution network side is more serious, as shown in the Figure 2.

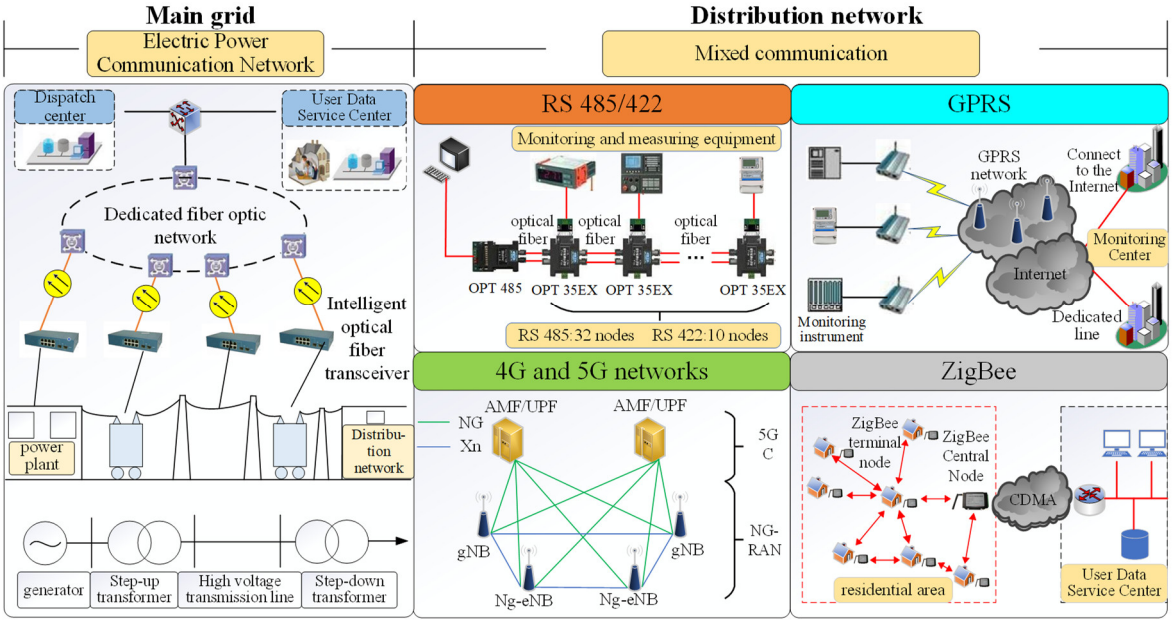


Figure 2. Main grid and distribution network.

For power application scenarios, power information security mainly includes three important goals, which are availability, integrity, and confidentiality [28,29]. For NILM scenarios, it is only used as a load analysis technology to provide refined power management. The NILM is not directly related to protection, and those control functions affect the stability of the power system. Compared with the protection and control functions of power system, real-time availability demands are relatively low. Integrity guarantees that the data of NILM will not be tampered with, which affects the authenticity and reliability of data. Confidentiality is of significance to NILM. Information leakage and behavior exposure of electricity customers that involves privacy sensitive topics will directly influence the commercial value and business ethics of NILM application. The encryption and decryption methods in cryptography for the confidentiality of NILM data require a lot of computing time, and with the increase of NILM data, the computing time will be multiplied. The monitoring equipment of NILM is mainly composed of embedded chips with relatively limited storage and computing performance, which needs to find an efficient encryption and decryption algorithm on the basis of ensuring the security of NILM data.

Although NILM technology provides convenience for information interaction between electricity customers and power supplier, it is more possible than ever that personal information and privacy will be exposed. With the intelligent development of household appliances, a large number of intelligent household electrical appliance will be accessed in user terminals, more closely connecting electricity customers with the information world. Measurement data of NILM system involves a lot of user privacy, including address, account number, meter data, real-time bill, historical bill, home LAN (local Area Network), and so on [30,31].

However, in the process of collecting and recording the power consumption information of electricity customers and communicating with the power supplier, the privacy information is threatened, as shown in Figure 3. Hackers may attack electro-data from the smart home appliances, the bidirectional network between the user side and the power supply side, and the data collection and processing center on the power supply side. Firstly, it is vulnerable to eavesdropping attacks. By obtaining data content and connecting power use time with different loads through NILM technology, the attacker steals electricity customers' real identity, living habits, and behavior pattern. It seems that electricity customers are activity under the attacker's "monitoring", such as getting up, resting, and watching TV, whether at home, in personnel composition, or in economic situation, and so on. Personal behavior privacy will be exposed to the attacker. Secondly, it is vulnerable to impersonation attack. Through intercepting user identity information, the attacker pretends to be a legal user,

accesses the power information system through the transmission channel, and controls the use of power consumption load. Thirdly, it is vulnerable to tampering attack. Illegal users may tamper with electro-data and conduct consumer fraud out of the lure of economic benefits [32].

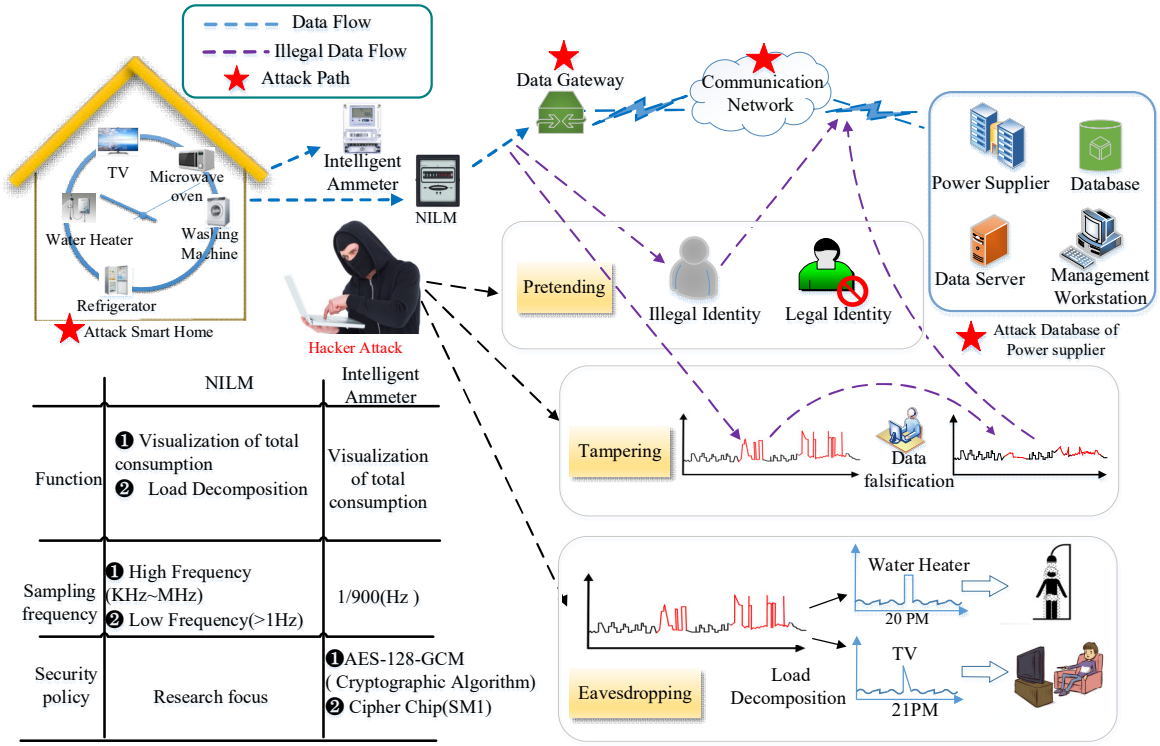


Figure 3. The manner of NILM attack compared with intelligent ammeter.

In order to prevent hacker attacks, to protect the information security of data, to maintain the interests and privacy of electricity consumers, the cryptography technology is widely used to ensure that, even if the information is hijacked, useful information cannot be obtained by the hacker. Intelligent ammeters, which have been gradually applied in the field of electric power, are the closest example to NILM security. For the information security of an intelligent ammeter, AES-128-GCM (Galois/Counter Mode) encryption mode is used to protect the data security and each intelligent ammeter has an authenticated key and a unique PIN (Personal Identification Number) key [33]. At the Cyber Tech 2016 conference, however, Israel's minister of energy and water infrastructure disclosed that Israel's electricity board suffered a serious cyber attack on 25 January 2016. After that, Israeli authorities were forced to shut down the infected computers of electric power facilities [34]. In China, using embedded security module chip ESAM (embedded secure access module), through the hardware integration of the national secret algorithm SM1, data encryption and decryption are realized, which are used to store key data such as remaining meters and rates in the meter [35]. The security policy and function of this mode are relatively fixed, lack of flexibility, and cannot meet the security requirements of NILM large data and fine-grained data. Small-scale pilot verification work has been carried out in Jiangsu, Tianjin, Nanjing, and other places in China for meters with NILM function, but there is still no report on NILM data security research. The security of relevant operational data is facing great challenges, therefore, it is necessary to study a secure transmission scheme of power utility information in NILM environment to provide scheme and technical support for user identity privacy protection and electro-data security.

4. NILM Data Privacy Protection Scheme Based on Cryptographic Algorithm

4.1. Data Cryptography Technology

Although NILM data information have rich connotations, there are unsafe factors in the transmission course, which can easily divulge users' sensitive information in plain text transmission. Therefore, it is urgent to find a suitable information security method for NILM data. As the most basic security technology, cryptography is an effective method to protect the secure transmission and storage of data. Safe and rational application of cryptographic algorithm can provide core support for the stable and efficient operation of an NILM system. The encryption algorithm for NILM data security should be easy to use, and even transparent to electricity customers. This is important seeing as thousands of electricity customers do not master cryptography technology, and they are not even willing to remember the password generally involved in the cryptography technology. Another important point is that, owing to the large amount of NILM data, the algorithm efficiency based on security is the key point. The most common cryptographic algorithms applied to NILM data can be classified into the following three basic algorithms: symmetric encryption, asymmetric encryption, and hash algorithm, as shown in Figure 4.

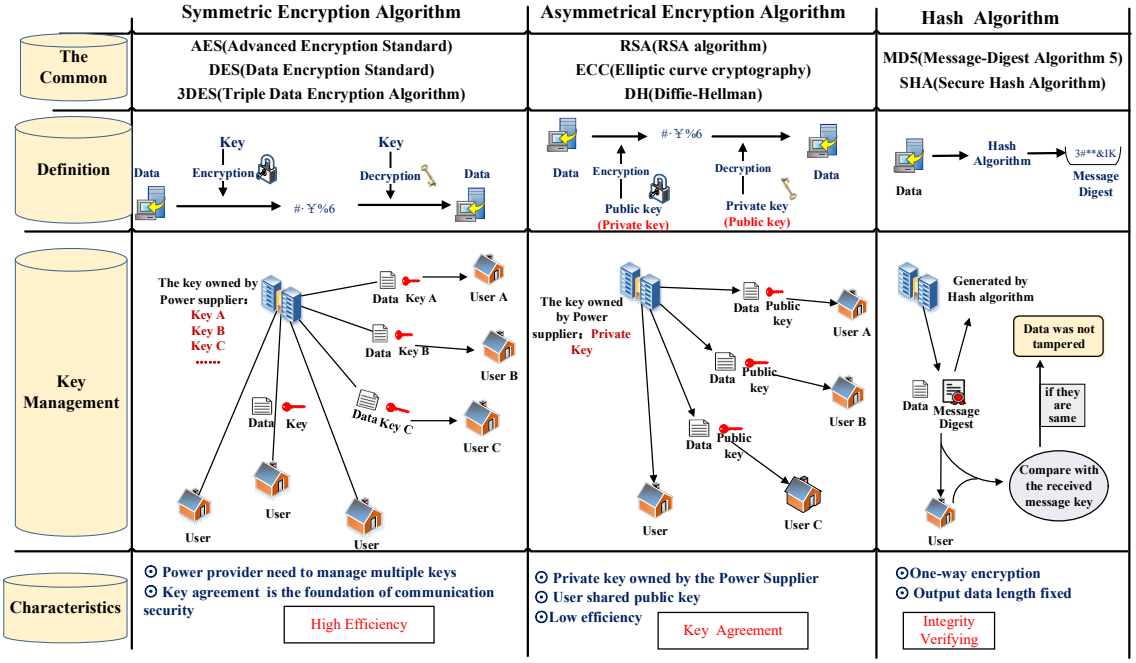


Figure 4. Classification of cryptographic algorithms.

Symmetric encryption algorithm is also known as shared key encryption. The electricity customers and power supplier use the same key for encryption and decryption. This kind of algorithm is frequently used to encrypt large-scale sensitive data in electric power industry owing to its high efficiency of encryption and decryption [36]. Confidentiality is the foundation of NILM data transmission. As the electricity customers and power supplier use the same key, how to distribute the key to electricity customers is the key to security assurance. In addition, the number of electricity customers will be more and more if the NILM system is widely used. For example, when communicating with one electricity customer using symmetric encryption, the power supplier must use the unique key that is unknown by any other electricity customer. Thus, the number of keys owned by the power supplier will increase at geometric series, and key management will become the burden between the power supplier and electricity customers. Symmetric encryption algorithms mainly include the following: AES, DES, 3DES, RC2, and RC5, among others. Among them, AES, whose safety is higher than that of DES and 3DES, is the standard of electronic data encryption of technology institutes in America. In order to adapt to different requirement and situations, AES

provides five different kinds of working modes, which are ECB (Electronic Code book Book), CBC (Cipher Block Chaining), CTR (Counter), CFB (Cipher Feed Back) and OFB (Output Feed Back), and AES has become the most popular algorithm in the electric power industry [37].

Compared with symmetric encryption technology, asymmetric encryption algorithm does not need to share the common key. Because it belongs to a double key system, and uses a public key and private key, it has two different ciphers for encryption and decryption. When using asymmetric encryption algorithm data, only one matching pair of public key and private key can complete the encryption and decryption of data. The public key can be made public and the private key should be self-reserved. Even if the public key may be intercepted in the transmission and publishing process, it is nonsense for the attacker as there is no private key paired with it. Thus, the distribution and key management of asymmetrical encryption algorithm is relatively simple and easy. However, owing to the complexity of this algorithm, its speed of encryption and decryption is relatively slow. Moreover, the generated keys are complex, so the fixed key mode can only be used and this algorithm is not suitable for NILM system in real-time collection of electronic data. RSA encryption is the representative of asymmetrical encryption algorithm, which is the most influential and most commonly used asymmetric encryption algorithm at present. It is recommended as the asymmetrical encryption standard by ISO (International Organization for Standardization) because it can resist most known cryptographic attacks [38,39].

Digital signature is another application of RSA. Combined with the one-way hash algorithm, it can realize user identity authentication and data integrity calibration by extracting NILM fingerprint information data.

4.2. NILM Data Privacy Protection Scheme Based on Hybrid Cipher Algorithm

Encrypting data using cryptographic algorithms is an active safety-protection strategy. Compared with sending plain text directly, sender need a certain amount of system time to encrypt data, and the receiver should also need extra time to decrypt data. Asymmetrical encryption algorithm will increase algorithm size and take up more network resources, so the processing efficiency of the system will be decreased by encryption and decryption. For frequent real-time and interaction in the NILM system, the loss of efficiency should be minimized to the greatest extent. In addition to efficiency loss, the security of the data encryption system is based on key confidentiality. The security of symmetric encryption algorithm is completely dependent on the key and needs a secure channel to distribute. Moreover, electricity customers of an NILM system are very large and key management is complex. Considering the requirements of efficiency, key management, and integrity of NILM system, a new privacy protection scheme of NILM data based on hybrid cipher algorithm was proposed.

Scheme description:

(1) Hybrid cipher algorithm. The scheme description is shown in Figure 5. Aiming at the threats of eavesdropping, tampering, and falsification in electro-data, three kinds of encryption algorithms, AES, RSA, and HMAC-SHA1, were used in this scheme. RSA algorithm protects the key of AES algorithm and AES algorithm encrypts the users' large data. It can not only realize the safe and convenient key management, but also ensure the speed of data encryption. However, there are still loopholes in key management of RSA as the public key is public to the outside. If user B masquerades user A to send data to the power supplier by its public key, the power supplier cannot recognize the true identity of the sender. In order to solve this problem, hash algorithm HMAC-SHA1 was introduced. HMAC-SHA1 is a one-way encryption algorithm. Users can generate one unique digital digest with a specific length for real-time electro-data and combine it with RSA algorithm to generate users' digital signature, and the integrity of data transmission and identity authentication of the sender can be guaranteed.

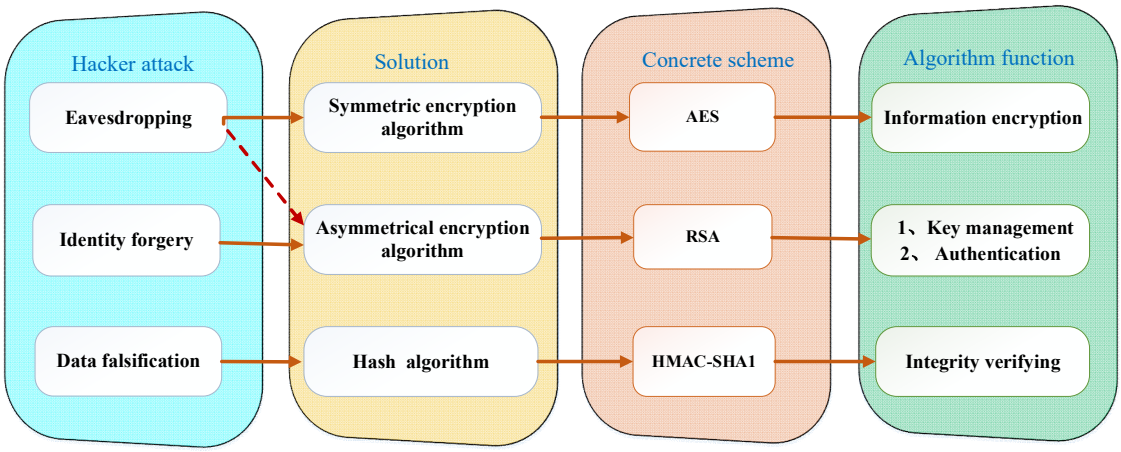


Figure 5. Hybrid cipher algorithm.

(2) Key distribution and management. The electricity customers newly load NILM and both sides conduct network authentication. The power supplier pre-installs the public key (PK) on the user side. NILM has configured factory initialization key (public key = PK11, private key = PS11). The power supplier has the key pairs public key = PK and private key = PS. The key distribution is shown in Figure 6.

A. Power supplier sends the public key PK to electricity consumers. At this time, hackers can intercept PK;
B. Electricity customers received PK and compared it with the public key pre-installed by the power supplier in NILM. If they are consistent, the public key is proved to be legal and the next step will be executed. If they are inconsistent, the public key is considered to be illegal and the next step will not be executed;
C. Electricity consumers use the public key PK to encrypt the public key PK11 initialized in the NILM device and transmit it to the power supplier. At this time, the hacker can intercept the ciphertext and know that it is encrypted through PK, but because the hacker does not know the private key PS, he cannot decrypt the ciphertext;
D. Power supplier receives the ciphertext, decrypts it with private key PS, and gets PK11;
E. RSA key pair (public key = PK1, private key = PS1) is randomly generated by the key management center of the power supplier, encrypted with PK11, and transmitted to the user. Even if hackers intercept ciphertext, they cannot decrypt it. Electricity consumers decrypt the ciphertext with PS11 and obtain the public key PK1 and PS1 distributed by power supplier. The key distribution is complete;
F. Electricity consumers need to secure private key PS1, and keep PK and PK1. The power supplier should keep the PS secret and have all users' public key sequences. The key based on AES algorithm is generated randomly every time data are sent, so that it can be used up and discarded without saving and management, which effectively reduces the difficulty of key management.

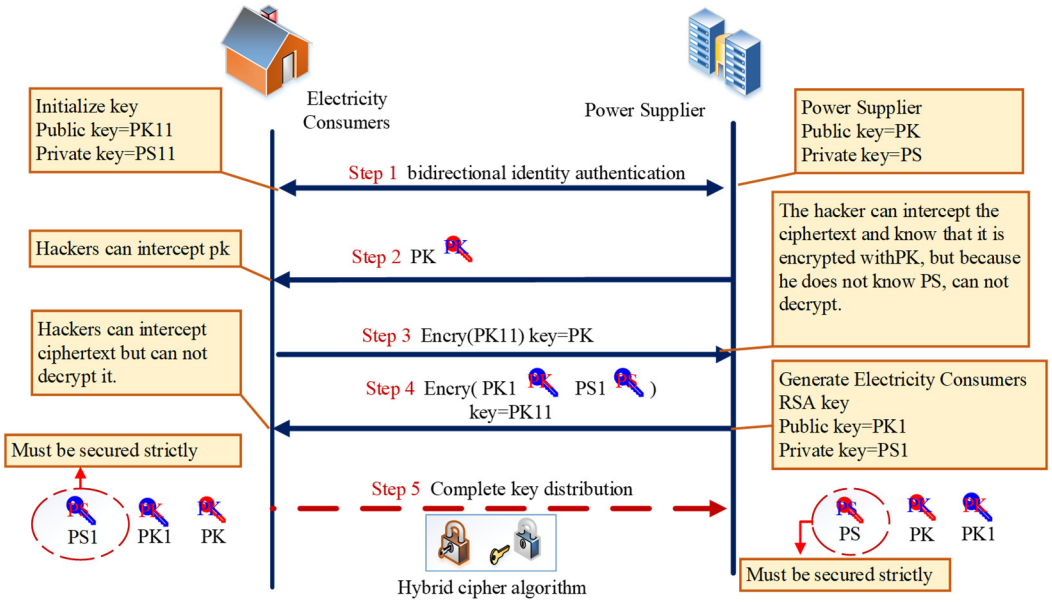


Figure 6. Key distribution and management.

(3) Algorithm flow. After the electricity customer newly installs the non-intrusive load detection device and officially runs online, the power supplier side immediately initializes the electricity customer's meter data, and identifies the electricity customer identity as A. Electricity consumers obtain PS1, PK1, and public key PK sent down by the power supplier. If electricity consumers need to send electricity data (start with identification A) to the power supplier, AES key should be randomly generated to encrypt the plaintext data and generate the ciphertext block of plaintext. Public key PK of power supplier is used to encrypt AES key to get the key ciphertext block. HMAC-SHA1 algorithm generates the digital digest for data, and the PS1 private key of electricity consumers encrypts it to form a digital signature. When the power supplier receives data sent from electricity consumers, the key ciphertext block with private key PS is firstly decrypted to get the AES key. Then, the AES key is used to decrypt the plaintext ciphertext block to obtain plain text. Finally, the AES key is discarded. According to the identification of user A in plain text, choose public key PA of user A to decrypt the digital signature. If there is no solution, it is shown that electricity customers' identity information has been masqueraded and then discarded. Otherwise, extract the corresponding digital summary. The power supplier uses HMAC-SHA1 algorithm to get one new digital digest with the decrypted plain text and compares it with the received digital digest. If the two parts are consistent, it indicates that data integrity is maintained and the data have not been tampered with. The process of data transmission and reception is over. Similarly, if the power supplier needs to send information to the electricity customers, the same process should be adopted. The flow is shown in Figure 7.

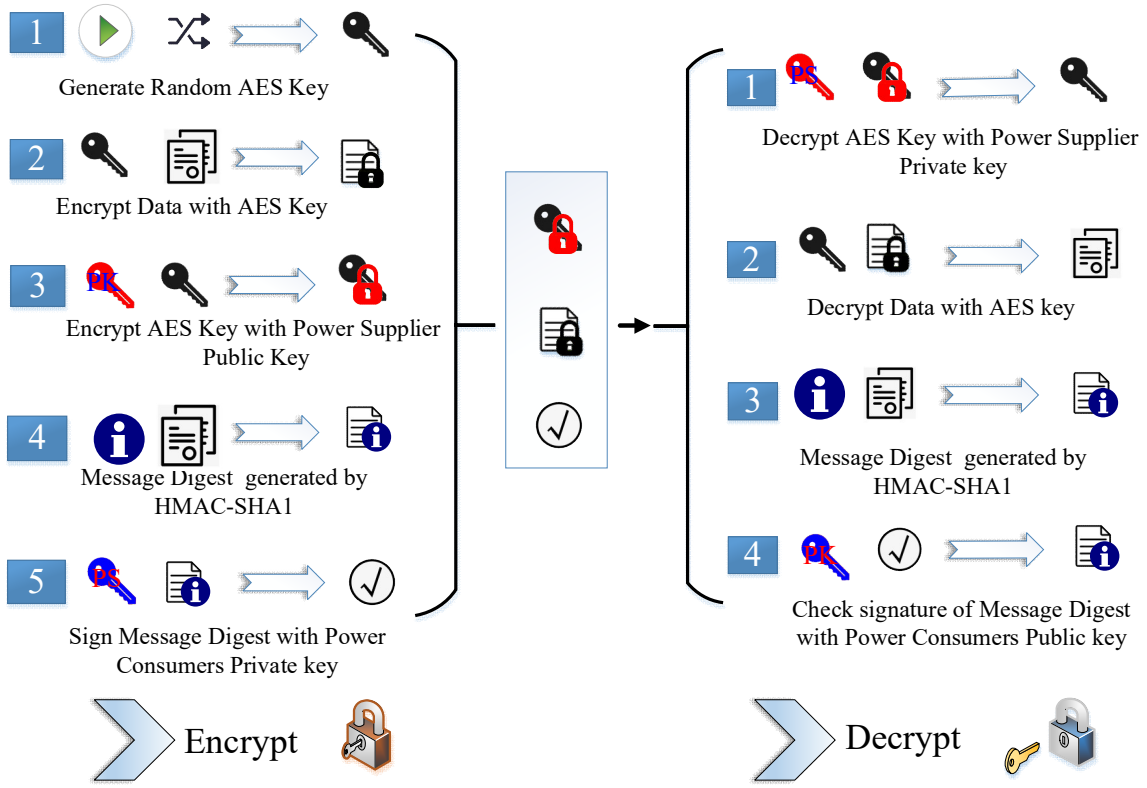


Figure 7. Algorithm flow.

5. Performance Analysis

Because a large amount of NILM data encryption and decryption requires a lot of calculation time, the efficiency of the algorithm proposed in this paper is a key factor to be considered in practice. This paper will establish a test platform to fully verify the efficiency of this method in different cryptography scenarios. At present, one of the most classical platforms for the encryption and decryption algorithm is based on an ARM (Advanced RISC Machine) operating system, such as Linux system. It involves calling OpenSSL library functions to conduct an analysis of encryption and decryption performance. However, there is less analysis of the algorithm running on bare metal alone. To provide a performance reference scheme for bare metal and experimental data for the selection of data encryption card hardware in the NILM system, the cost-effective STM32 MCU was used as the carrier to run mixed encryption algorithm for scheme test [40]. Introduction of the cipher algorithm will lead to extra delay for the transmission of electric data. Therefore, it is necessary to evaluate the efficiency of encryption and decryption of cryptographic algorithms, and to fully consider the computing resources occupied by data encryption and decryption while ensuring data confidentiality. This paper analyzed the performance of the scheme from the computing cost of encryption and decryption. The test scheme is shown in Figure 8 below.

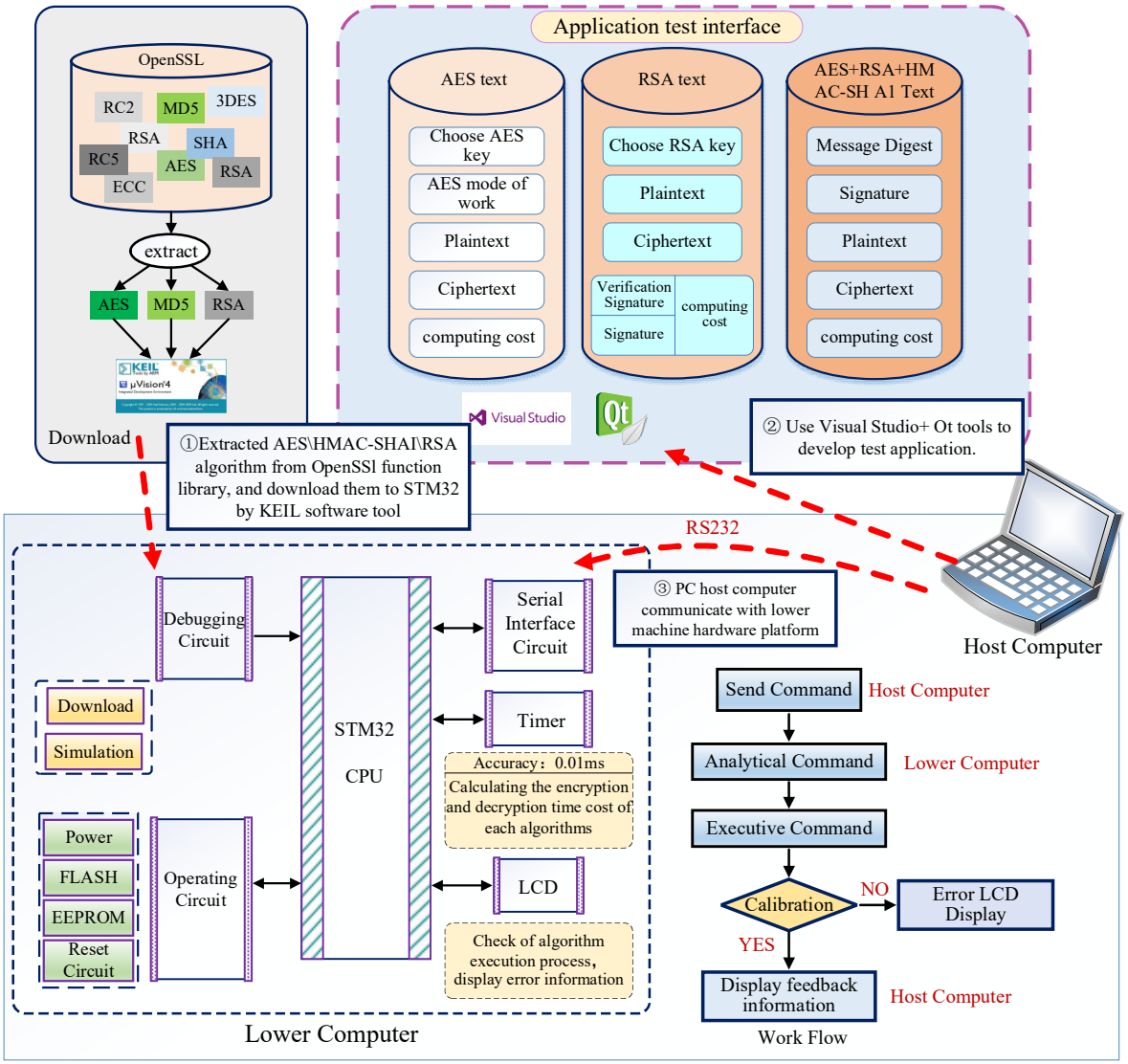


Figure 8. Test equipment diagram.

The test platform was composed of a PC host computer and MCU lower machine hardware platform. (1) The interface of the PC host computer was developed by Visual Studio 2017 and Qt5 tools, whose function was sending commands and receiving feedback information from the lower machine through the serial port. The sending commands included the following: a. encryption algorithm type; b. length of key; and c. electricity data. The received information was the required test performance of various algorithm modes running in MCU under the setting conditions. (2) The MCU platform of the lower computer was established with the ARM chip of STM32H743IIT6 model, which is the high-performance 32 bits ARM Cortex-M7MCU chip. Its working frequency can reach 480 MHz. It involves extracting the relevant encryption algorithms in the OpenSSL function library and transplanting them to the MCU, including AES, RSA, and HMAC-SHA1 algorithm. In this test, the low machine was responsible for running the encryption and decryption algorithm according to the instructions sent by the upper computer, and communicating with the upper computer software through the serial port.

5.1. Analysis of Encryption and Decryption Efficiency of AES Algorithm

AES uses a block cipher system. The size of each cipher block is 128 bit and the allowed key length is 128 bit, 192 bit, and 256 bit, respectively. Electric data to be encrypted are changed in real time, and the data format may vary. There are usually five work modes of AES, which are ECB, CBC,

OFB, CFB, and CTR. In order to use AES algorithm safely and efficiently in the NILM system, the correlation among electric data length, key length, and working mode was tested. Among them, data length was at intervals of 5000 byte, and the encryption and decryption computing cost of 29 group data under different modes and different key lengths was recorded. Finally, experimental data were processed and analyzed.

5.1.1. Computing Cost Analysis of Encryption in Different Working Modes

Encryption computing time under different working modes with 128 bit key length is shown in Figure 9 below as an example. In the diagram, computing time under different working modes increased in all cases with the increase of power consumption information, and presented a linear correlation, $R^2 = 1$. The time cost of CFB1 and CFB8 modes was obviously longer than that of any others. From the calculations, CFB1 was 8 times that of CFB8, and CFB8 was 16 times that of CFB128. As the NILM system needs to transmit data in real time and too long a computing cost will affect the efficiency of the system, the two longer time modes were not selected for use. Except for CTR mode, the testing curves of the other four modes showed a high coincidence and computing cost was the same, while the time of computing cost of the CRT mode was 2.5 ms longer than that of any other mode in the test of 145,000 byte. However, if the NILM system uses low frequency as the frequency of real-time data acquisition, the CTR mode is basically the same as the calculation cost of the other four modes. In the end, the test results show that the trend of encryption computing cost of 192 bit and 256 bit key length was the same as that of 128 bit.

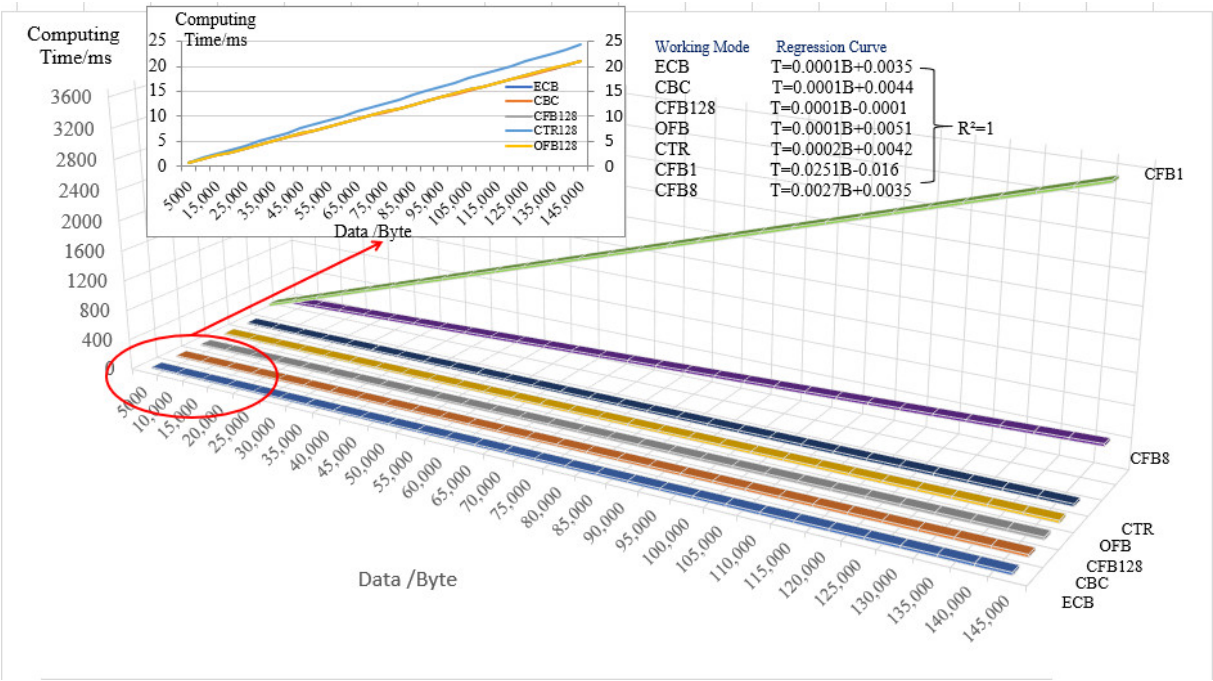


Figure 9. Encrypted plaintext computing cost of 128-bit key in different encrypting modes.

5.1.2. Comparative Analysis of Computing Cost between Encryption and Decryption

Table 1 showed the comparison and analysis of the computing cost of the encryption and decryption of each mode. From the table, the encryption and decryption computing cost interval of CBC, CBF1, and CFB8 was large, and difference of other modes was less than 0.2 ms. The absolute difference value of CTR mode was 0.01 ms, considering that counting period of the timer in STM32 was set to be 0.01 ms, so the time of encryption and decryption of the CTR mode was the same. According to the principle of the CTR mode, there is no error in encryption. However, the ciphertext data will be distorted if there is a hacker attack or network transmission problem in the network transmission. Because AES is a grouping algorithm, the decryption result will only have errors in the

distorted data blocks, and the decryption result for the blocks without data distortion is consistent with the plaintext. Thus, the CTR mode was selected as the AES encryption and decryption mode in this scheme.

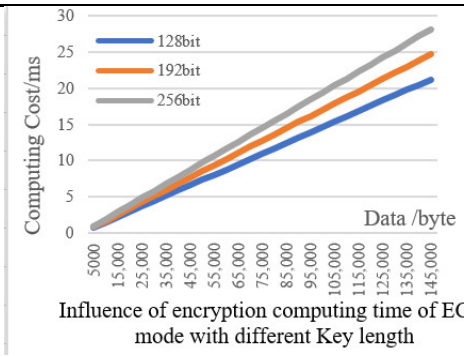
Table 1. Computing cost analysis and comparison of AES encryption and decryption.

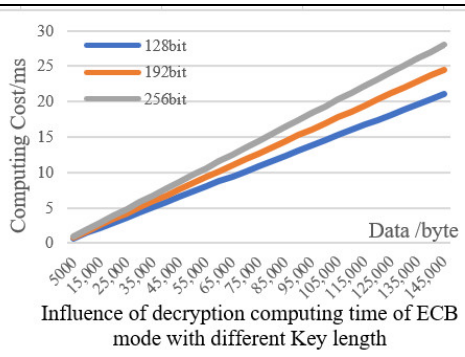
Working Mode	Difference range of encryption and decryption computing time in different key length/ms			Conclusion
	128 bit	192 bit	256 bit	
ECB	−0.01~−0.16	0~−0.16	0~−0.16	Decryption a little longer than encryption
CBC	−0.03~−0.87	−0.04~−0.88	−0.03~−0.87	Decryption a little longer than encryption
CFB1	0.26~7.24	0.19~6.64	0.15~8.43	Encryption longer than decryption
CFB8	0.03~0.5	0.01~0.5	0.01~0.51	Encryption longer than decryption
CFB128	0~0.05	0.01~0.05	0~0.05	Encryption a little longer than decryption
OFB	0~0.02	0~0.01	0~0.02	Encryption and decryption same time
CTR	0~0.01	0~0.01	0~0.01	Encryption and decryption same time

5.1.3. Influence on Encryption and Decryption Computing Cost with Different Key Lengths

Table 2 shows the encryption and decryption efficiency of different key lengths. From the data in the table, it can be seen that, with the increase of the key length, the encryption and decryption efficiency of AES algorithm was obviously reduced. In the hybrid cryptosystem, AES algorithm is mainly used to encrypt data, thus efficiency is the first factor to be considered. After comprehensive consideration of security, interoperability, and computing cost, AES-128-CTR was selected as the encryption and decryption algorithm in the hybrid cryptosystem.

Table 2. Encryption and decryption efficiency with different key lengths.





Working Mode	ECB	CBC	CFB1	CFB8	CFB128	OFB	CTR
Conclusion	Computing cost of encryption and decryption increases with the increase of key length	The trend table of the six modes is the same as that of ECB, and computing cost of encryption and decryption increases with the increase of key length					

5.2. Encryption and Decryption Efficiency of RSA Algorithm of Different Key Lengths

Because the encryption speed of asymmetric algorithm RSA is too slow to fit the large data collection of the NILM real-time system, it is usually used for encryption and decryption of the encryption key and digital signature and verification. The electric power industry usually uses it as

longitudinal encryption authentication gateway, network security isolation device (inverse), and device management, among others. Common RSA key length is 1024 bit and 2048 bit. It could be found in the test that the encryption data size was 128 byte and 256 byte, respectively, with the key length of 1024 bit and 2048 bit, respectively. Although the data size is limited by the length of the key, as the auto-fill mode is used during the encryption process, the encryption and decryption times are exactly the same on the maximum allowed data length. At the same time, the effect of key length on encryption and decryption efficiency was tested as shown in Table 3. From the table, the speed of RSA public encryption was fast and private encryption was slow. The speed of data signature was slow and data verification was fast. Owing to the difficulty of factorization based on large integers, it is difficult to decrypt the RSA algorithm. With the development of cryptography technology, RSA-768-bit algorithm was already declassified in the year 2009. Although the amount of calculation for decrypting a 1024-bit key is more than 1000 times that of a 768-bit key, an expert estimated the former will be declassified in the next 10 years. Application of 2048-bit key length will be the mainstream of RSA algorithm. At present, the NILM system is only in pilot use in Nanjing. With the wide popularization in future, its use period will extend and the security of the cryptography algorithm will be guaranteed. Therefore, RSA algorithm with a key length of 2048 bit was selected.

Table 3. RSA computing time with different key lengths.

Key length bit	Encryption time (times/s)	Decryption time (times/s)	Signature time (times/s)	Verify times (times/s)
1024	212.3	5.7	5.6	213
2048	57.9	0.87	0.87	57.7

5.3. Hybrid Cryptography Scheme Test

According to the analysis above, symmetric algorithm AES-126-CTR was selected to encrypt data. Asymmetric algorithm RSA-2048 was selected to encrypt the AES key and digital signature. In order to ensure the accuracy of the data, HMAC-SHA1 was used in the mixed encryption and decryption scheme to check the integrity. It is agreed that the initial counter value of AES-CTR algorithm module on the side of power supplier and power consumers is set to be 0. In the case of transmission error, if the integrity check fails, the sender will be immediately informed to resend the data. At this time, the receiver reinitializes the AES-CTR module and resynchronizes the counter, which is set to 0. The hybrid encryption scheme of NILM electro-data was composed with these three parts. The effectiveness and computing cost of the scheme were tested as shown in Figure 10 below.

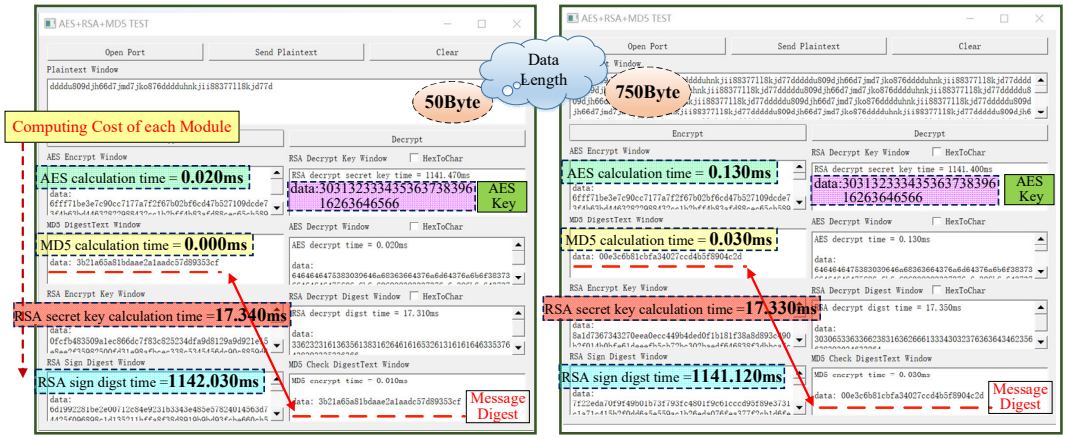


Figure 10. Scheme test interface.

The encryption and decryption computing time of each module in the scheme was tested, including AES encryption data, HMAC-SHA1 data message digest generation, RSA encryption of AES key, and signature of HMAC-SHA1 digest generation. In order to guarantee the effectiveness of

the scheme, the encrypted and decrypted data and ciphertext were displayed on the interface for analysis and comparison. After many repetitions of the test, the encrypted data were consistent with the decrypted data and message digest of user and power supplier were consistent, so the NILM hybrid cryptography algorithm was effective. From the test, the whole scheme computing cost depended on the RSA decryption and digital signature, as shown in Figure 11 below.

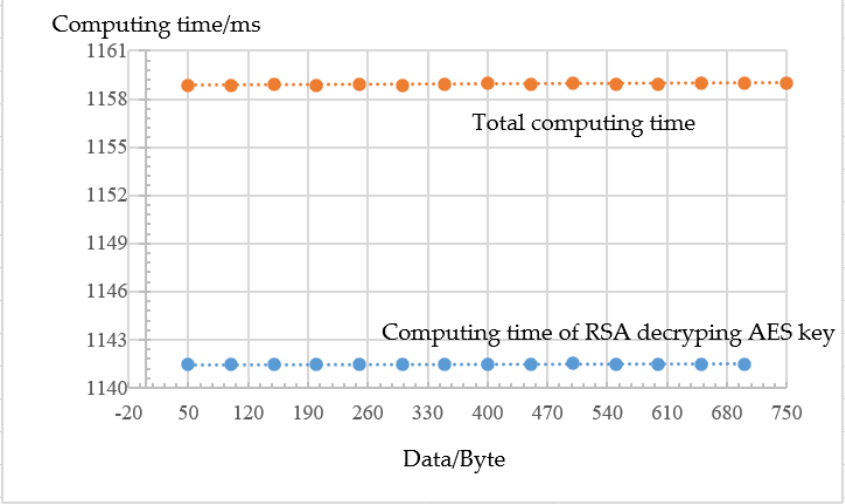


Figure 11. Scheme computing time test result.

6. Conclusions

By being installed at the power entrance of the user, NILM collects the electricity customers' power utility information to achieve better power consumption experience and value added-services. Although frequent data collection brings convenience to electricity customers, it is inevitable that disclosure of sensitive information of electricity customers is involved. Therefore, NILM data security, especially concerning the privacy protection of electricity users, becomes an important security requirement in the NILM system. Taking into full consideration the NILM system characteristics, this paper proposed a hybrid cryptographic scheme. In the scheme, symmetric algorithm AES was used to encrypt electro-data. Asymmetric algorithm RSA was used to encrypt AES key to achieve efficient key management, which was also used for authentication, and hash algorithm HMAC-SHA1 guaranteed the integrity of data. The computing cost, key management, authentication, and integrity of data were considered in the scheme. The requirements of data security and confidentiality were satisfied. By transplanting the algorithm to STM32 MCU and developing the test interface, the performance of the scheme computing cost was mainly tested.

A. Five working modes of AES algorithm, including ECB, CBC, OFB, CFB, and CTR, were tested with different key lengths of 128 bit, 192 bit, and 256 bit, respectively. The test results showed that the computing cost of the five working modes was linear with the data length and the key length. In addition to CTR mode, the computing cost of the other four modes was highly coincident. When the data quantity of CTR mode was 145,000 byte, the computing cost of mode was only 2.5 ms higher than that of other modes, so it can be drawn that the computing cost of the five working modes is the same. From the test, the encryption and decryption time of the CTR mode were the same. On the basis of the real-time requirement of the NILM system, the AES-128-CTR mode was selected as the working mode of the hybrid cipher scheme.
B. The efficiency of RSA algorithm with the key lengths of 1024 bit and 2048 bit was tested. The test results showed that speed of RSA public encryption was fast and private encryption was slow. The speed of data signature was slow and data signature verification was fast. Considering the security requirements of NILM, RSA-2048 was selected as the working mode of the hybrid cipher scheme.

C. The effectiveness and efficiency of the hybrid cipher algorithm scheme were tested. By testing many times on the developed interface, the encrypted data were consistent with the decrypted data. Information digest of the power consumers and power supplier was consistent; therefore, the NILM hybrid cipher algorithm was effective. The computing cost of RSA decryption and signature basically determined the computing cost of the total scheme.

Author Contributions: Conceptualization, R.F. and Z.W.; methodology, R.F.; software, R.F.; validation, R.C., H.M., and Z.L.; formal analysis, R.F.; investigation, H.M.; resources, R.F.; data curation, X.Z. and Z.C; writing—original draft preparation, R.F.; writing—review and editing, Z.P.; visualization, R.C.; supervision, Z.W.; project administration, R.F.; funding acquisition, Z.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Doctoral Research Startup Project of Guangdong Science and Natural Science, Grant Number 2017A030310288, and partially by outstanding youth teacher project by Guangzhou College of South China University of Technology: NO. JQ180001.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hart, G.W. Nonintrusive appliance load monitoring. *Proc. IEEE* **1992**, *80*, 1870–1891.
2. Chen, Y.; Martínez-Ortega, J.-F.; Castillejo, P.; López, L. A homomorphic-based multiple data aggregation scheme for smart grid. *IEEE Sens. J.* **2019**, *19*, 3921–3929.
3. Abdallah, A.; Shen, X.S. A lightweight lattice-based homomorphic privacy-preserving data aggregation scheme for smart grid. *IEEE Trans. Smart Grid* **2018**, *9*, 396–405.
4. He, D.; Kumar, N.; Zeadally, S.; Vinel, A.; Yang, L.T. Efficient and privacy-preserving data aggregation scheme for smart grid against internal adversaries. *IEEE Trans. Smart Grid* **2017**, *8*, 2411–2419.
5. Knirsch, F.; Eibl, G.; Engel, D. Error-resilient masking approaches for privacy preserving data aggregation. *IEEE Trans. Smart Grid* **2018**, *9*, 3351–3361.
6. Sui, Z.; de Meer, H. An efficient signcryption protocol for hop-by-Hop data aggregations in smart grids. *IEEE IEEE J. Sel. Areas Commun.* **2020**, *38*, 132–140.
7. Diao, F.; Zhang, F.; Cheng, X. A privacy-preserving smart metering scheme using linkable anonymous credential. *IEEE Trans. Smart Grid* **2015**, *6*, 461–467.
8. Zhao, J.; Liu, J.; Qin, Z.; Ren, K. Privacy protection scheme based on remote anonymous attestation for trusted smart meters *IEEE Trans. Smart Grid* **2018**, *9*, 3313–3320.
9. Gong, Y.; Cai, Y.; Guo, Y.; Fang, Y. A privacy-preserving scheme for incentive-based demand response in the smart grid. *IEEE Trans. Smart Grid* **2016**, *7*, 1304–1313.
10. Ratliff, L.J.; Dong, R.; Ohlsson, H.; Cárdenas, A.A.; Sastry, S.S. Privacy and customer segmentation in the smart grid. In Proceedings of the 53rd IEEE Conference on Decision and Control, Los Angeles, CA, USA, 15–17 December 2014.
11. Li, H.; Liu, D.; Dai, Y.; Luan, T.H.; Yu, S. Personalized search over encrypted data with efficient and secure updates in mobile clouds. *IEEE Trans. Emerging Top. Comput.* **2018**, *6*, 97–109.
12. Fu, Z.; Ren, K.; Shu, J.; Sun, X.; Huang, F. Enabling personalized search over encrypted outsourced data with efficiency improvement. *IEEE Trans. Parallel Distrib. Syst.* **2016**, *27*, 2546–2559.
13. Kumar, P.; Gurtov, A.; Sain, M.; Martin, A.; Ha, P.H. Lightweight authentication and key agreement for smart metering in smart energy networks. *IEEE Trans. Smart Grid* **2019**, *10*, 4349–4359.
14. He, K.; Jakovetic, D.; Zhao, B.; Stankovic, V.; Stankovic, L.; Cheng, S. A generic optimisation-based approach for improving non-intrusive load monitoring. *IEEE Trans. Smart Grid* **2019**, *10*, 6472–6480.
15. Rahimpour, A.; Qi,H.; Fugate, D.; Kuruganti, T. Non-intrusive energy disaggregation using non-negative matrix factorization with sum-to-k constraint. *IEEE Trans. Power Syst.* **2017**, *32*, 4430–4441.
16. He, J.; Zhang, Z.; Zhu, L., Zhu, Z.; Liu, J.; Gai, K. An efficient and accurate nonintrusive load monitoring scheme for power consumption. *IEEE Internet Things J.* **2019**, *6*, 9054–9063.
17. Morais, L.R.; Castro, A.R.G. Competitive autoassociative neural networks for electrical appliance identification for non-Intrusive load monitoring. *IEEE Access* **2019**, *7*, 111746–111755.
18. Rehman, A.U.; Lie, T.T.; Vallès, B.; Tito, S.R. Event-detection algorithms for low sampling nonintrusive load monitoring systems based on low complexity statistical features. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 751–759.

19. Liu, Y.; Sun, Y.; Li, B. A modified IP-based NILM approach using appliance characteristics extracted by 2-SAX. *IEEE Access* **2019**, *7*, 48119–48128.
20. Chen, P.; Zhong, S. A non-intrusive load monitoring and identification system based on cloud platform. *Appl. Electron. Tech.* **2018**, *44*, 91–95.
21. Wang, Z.; Wang, G.; Li, Y.; Tong, J. An encryption method for IEC 61850-9-2LE packet based on tiny encryption algorithm. *Autom. Electr. Power Syst.* **2016**, *40*, 121–127.
22. Kolter, J.Z.; Johnson, M.J. REDD: A public data set for energy disaggregation rese arch. In proceedings of the KDD workshop on Data Mining Applications in Sustainability (SustKDD 2011), San Diego, CA, USA, 21–24 August 2011.
23. Cheng, X.; Li, L.; Wu, H. Ding, Y.; Song, Y.; Sun, W. A survey of the research on non-intrusive load monitoring and disaggregation, *Power Syst. Technol.* **2016**, *40*, 3108–3117.
24. Pereira, L.; Nunes, N. Performance evaluation in non-intrusive load monitoring: Datasets, metrics, and tools-A review. *Wires Data Min. Knowl. Discovery* **2018**, *8*, doi:10.1002/widm.1265.
25. Huang, B.; Knox, M.; Bradbury, K.; Collins, L.M.; Newell, R.G. Non-intrusive load monitoring system performance over a range of low frequency sampling rates, In proceedings of the 2017 IEEE 6th International Conference on Renewable Energy Research and Applications (ICRERA), San Diego, CA, USA, 5–8 November 2017.
26. Kelly, J.; Knottenbelt, W. The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes. *Sci. Data* **2015**, doi:10.1038/sdata.2015.7.
27. Wang, Z.; Ni, Y.; Zhang, Z.; Wang, G.; Chen, Z.; Deng, F.; Pu, Z.; Yang, L.; Zhang, Y.; Feng, R.; et. al. Intelligent Distribution Network Information Processing Based on Power Data Virtual Plane. *Appl. Sci.* **2020**, *10*, 736.
28. Introduction to NISTIR 7628 guidelines for smart grid cyber security. Available online: https://nvlpubs.nist.gov/nistpubs/ir/2010/NIST.IR.7628.pdf (accessed on 6 February 2020).
29. Pillitteri, V.Y.; Brewer, T.L. Guidelines for smart grid cyber security. *Nat. Inst. Stand. Technol.* **2014,** doi:10.6028/NIST.IR.7628r1.
30. Zhu, W.; Wang, Z.; Zhang, Z. Renovation of automation system based on industrial internet of things: A Case Study of a Sewage Treatment Plant. *Sensors* **2020**, *20*, 2175.
31. Asghar, M.R.; Dán, G.; Miorandi, D.; Chlamtac, I. Smart meter data privacy: A Survey. *IEEE Commun. Surv. Tutorials* **2017**, *19*, 2820–2835.
32. Ni, J.; Zhang, K.; Lin, X.; Shen, X.S. Balancing security and efficiency for smart metering against misbehaving collectors. *IEEE Trans. Smart Grid* **2019**, *10*, 1225–1236.
33. Xiao, Y.; Zhang, L.; Li, X. EU-27 AMI rollout and inspiration of the next generation IR46 smart meter planning. *Electr. Meas. Instrum.* **2019**, *56*, 146–152.
34. Li, Z.; Tong, W.; Jin, X. Construction of cyber security defense hierarchy and cyber security testing system of smart grid: thinking and enlightenment for network attack events to national power grid of Ukraine and Israel. *Autom. Electr. Power Syst.* **2016**, *40*, 147–151.
35. Tian, M.; Li, S. Application of the ESAM in the smart meters. *Res. Discuss.* **2018**, *4*, 29–31.
36. Zhou, W.; Li, P.; Wang, Q.; Nabipour, N. Research on data transmission of wireless sensor networks based on symmetric key algorithm. *Measurement* **2020**, *153*, 107454.
37. Lavanya R; Karpagam M. Enhancing the security of AES through small scale confusion operations for data communication. *Microprocess. Microsyst.* **2020**, *75*, doi:10.1016/j.micpro.2020.103041.
38. Pang, L.; Wei, M.; Li, H. Efficient and anonymous certificateless multi-message and multi-receiver signcryption scheme based on ECC. *IEEE Access* **2019**, *7*, 24511–24526.
39. Wang, F.; Xu, L.; Choo, K.-K. R.; Zhang, Y.; Wang, H.; Li, J. Lightweight certificate-based public/private auditing scheme based on bilinear pairing for cloud storage. *IEEE Access* **2019**, *8*, 2258–2271.
40. Feng, R.; Wang, Z.; Li, Z. Design and experiment on fertile solution pH value and electrical conductivity measuring instrument. *J. Irrig. Drain. Eng*. **2020**, *38*, 643–648.