DOI: 10.23977/jeis.2023.080106
ISSN 2371-9524 Vol. 8 Num. 1

A Privacy Protection Scheme for Smart Meter Electricity Consumption Data Based on Certificateless Aggregate Signcryption with Public Verifiability

Baoyi Wang[a,*], Xindong Liang[b], Shaomin Zhang[c]

School of Control and Computer Engineering, North China Electric Power University, Baoding Hebei, 071003, China
[a]wangbaoyi@126.com, [b]lyonxxxd@163.com, [c]zhangshaomin@126.com
[*]Corresponding author

Keywords: Privacy protection, smart meters, smart home, aggregate signcryption

Abstract: Smart meters bring convenience to the reasonable use of electricity by smart homes and effective power supply and efficient charging of power companies. Sensitive information such as user power is leaked by fine-grained smart meter data, and existing solutions achieve privacy protection of electricity data through signing passwords, data aggregation, etc. but most of them use bilinear pair operations or exponential operations. Due to the limited storage resources of smart meters in smart homes, they are not suitable for algorithms with high computational power consumption. And verification one by one leads to a low level of computational efficiency. Based on this, this paper proposes a publicly verifiable aggregate signcryption scheme, which only uses elliptic curve point multiplication, which protects privacy while reducing computing overhead and communication overhead, and prevents the intelligent gateway from being decrypted when verifying data.

1. Introduction

Advanced metering infrastructure (AMI), also known as smart meter infrastructure (SMI), which is equipped with a large number of smart grid terminals to support high-speed dual-phase communication between smart meters and utility backends [1]. But this large-scale data collection will inevitably involve user security and privacy. Literature [2] classifies privacy into four types: information privacy, personal privacy, behavioral privacy, and personal communication privacy, so the privacy leakage in this article mainly violates the user's behavioral privacy and personal communication privacy, for example, malicious users through remote access to smart meter metering data, or capture smart meters and control center two-way communication process metering data, may obtain a user's energy consumption data access, in addition, semi-honest smart gateways may also sell the user's electricity data plaintext. These granular, high-frequency energy usage data can easily analyze consumers' lifestyles, such as meal times, bedtime times, holidays, working hours, and even more private lifestyle habits[3]. Recent research in the field of non-intrusive load monitors (NILM) shows that this technology can accurately analyze energy consumption data to obtain behavioral privacy and personal communication privacy of customers' living habits [4].

Literature [5] is based on certificateless ring signature and federated blockchain. It first generates the user's corresponding pseudo identity, and then performs ring signature on the data. This is very useful in smart meter and smart gateways with limited computing resources. When the gateway is validated, the data collected by the intelligent gateway for a long time may trigger correlation analysis attacks due to fixed pseudonyms [6]. Reference [7] proposed a privacy protection scheme for smart meter that can verify the reliability of terminals. It completes the credibility of the smart meter fact verification, and the intelligent gateway will not have efficiency bottlenecks in the verification process like [8]. However, the security is poor in identity based password systems.

Literature [9] proposes a group blind signature privacy protection scheme, he provides a conditional anonymity scheme, the scheme will carry out fault analysis, the control center can conditionally revoke the anonymity of the user, but the scheme is based on bilinear operation, the overhead is too large.

Literature [8] proposes an edge-block-assisted privacy protection scheme, which supports two-level aggregation to be more secure and efficient. Literature [9] proposes a data aggregation scheme combined with fog calculation, and it is fault-tolerant, which means that even if smart meters downtime, other smart meters can participate in the aggregation. However, so it is necessary to collect fine-grained data of users. In addition, the homomorphic encryption algorithm used when aggregating on the basis of ciphertext is expensive [6].

To sum up the above problems, we study a publicly identifiable aggregation sign-off scheme, which does not use bilinear operations with low computational efficiency, and combined with the aggregation idea, it can merge the signed ciphertexts from different signers into a single ciphertext, greatly reducing the total length and verification time of the ciphertext, and is more suitable for intelligent gateways and smart meters with limited computing resources and storage resources. More suitable for the scenario of this article.

The rest of the paper is as follows. The second part describes the technology. The third part describes the model for this scenario. The fourth part shows the performance analysis. The fifth part is a summary of the article.

2. Related technology

2.1. Aggregate signcryption

Signing is to realize the functions of public key encryption and digital signature in the same operation step, and ensure the confidentiality and authentication of messages. Compared with the way signing and encryption are completed in two steps, the computational cost and communication overhead of signing are greatly reduced, and the security factor and efficiency are higher [6].

2.2. Difficult questions

1) Computational Diffie-Hellman (computeDiffie-hellman) problem: Problem: It is known that G is an additive cyclic group on an elliptic curve, the order of G is a large prime q, the generator is P and the CDH problem refers to the given (P, aP, bP), where $a, b \in Z_4$ the last knowledge and solves the value of abP.

2) Discrete logarithm (discrete logarithm, DL) problem: It is known that G is an additive cyclic group on an elliptic curve, the order of G is a large prime q, the generator is P, and the DL problem refers to a given tuple (P, aP), where$a \in Z$ are unknown, and solve the value of a.

3. Scheme design

References [10], this paper designs a network framework consisting of smart meter, regional gateway BG, power control center CC, trusted third parties KGC, and power supply (ps). As shown in Figure 1:

Smart Meter: A smart meter is a solid-state programmable device with an embedded tamper-proof hardware unit that communicates bidirectionally with the control center to transmit periodic energy consumption readings.

Regional intelligent gateway BG: It mainly completes two functions, namely collection and trunking. The duty of the collection is to collect and check the validity of residential user electricity usage data, while the relay's duty is to help the communication flow between CC and residential users be forwarded in a secure manner.

Control Center CC: This article assumes that the control center is an honest entity responsible for the collection and processing of back-end data and then send the results of the analysis to the PS.

Power supplier PS: According to the analysis results sent by the control center, it provides users with value-added services, such as electricity bill metering

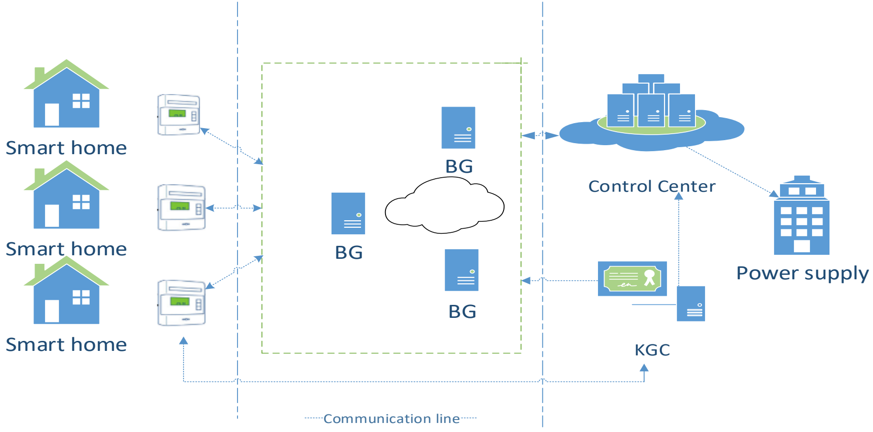

Figure 1: Smart home architecture diagram

3.1. Scheme implementation process

① KGC system parameter generation: input safety parameter k, generate two large prime numbers p、q, and satisfy $q \mid p-1$. Define 3 safe hash functions:

$H_1:\{0,1\}^* \times G \times G \rightarrow Z_a^*$,$H_2:\{0,1\}^* \rightarrow Z_a^*, H_3:\{0,1\}^* \times G \rightarrow Z_a^*$. KGC randomly selects

$x \in Z_a^*$ as the system master key and keeps it secret x, system public key $P_{pub} = xP$, and exposed system parameters params$= \{p, q, P, P_{pub}, G, H_1, H_2, H_3\}$.

②(1) The user goes to the designated key issuance center (KGC) to open an account and obtains an identity $ID_i = H_3(Device_i||address||time) \in \{0,1\}^*$, $Device_i$ is the device number, $address$ is the user address, and $time$ is the registration validity time.

(2) Partial key setting: the algorithm is executed by KGC, given the user's $ID_i, (0 \le i \le n)$, KGC randomly selects$r_{ID_i} \in Z_a^*$, calculate the user's partial public key $R_{ID_i} = r_{ID_i}P$, part of the private key $s_{ID_i} = r_{ID_i} + h_{ID_i}x$, where $h_{ID_i} = H_1(ID_i, R_{ID_i}, P_{pub})$, the user obtains the key (R_{ID_i}, s_{ID_i}).

(3) User key setting: The algorithm is executed by the user ID_i and first verifies whether the equation $s_{ID_i}P = R_{ID_i} + P_{pub}h_{ID_i}$ is valid, and if so, it means that part of the key issued by KGC is valid. Then randomly select $x_{ID_i} \in Z_a^*$ as the secret value, set its private key $SK_{ID_i} = (x_{ID_i}, s_{ID_i})$, calculate $P_{ID_i} = x_{ID_i}P$, set its public key $PK_{ID_i} = (R_{ID_i}, P_{ID_i})$, then the control center staff will install the smart meter in the user's home, and guide the user to inject the key information into the smart meter

tamper-proof module.

(4) Smart gateway, control center registration method is the same as above.

③ The sign-off algorithm is performed by the signer, and the smart meter ID_i, by the user performs the following actions on the message to be sent to the receiver ID_B the control center m_i:

(a) Randomly select $u_i \in Z_a^*$ to calculate $U_i = u_i P$

(b) Calculate $T_i = u_i(R_{ID_B} + P_{ID_B} + P_{pub} h_{ID_B}), W_i = H_3(ID_B, T_i), c_i = (ID_i \parallel m_i) \oplus W_i$

(c) Calculate $h_i = H_2(c_i, ID_i, U_i, R_{ID_i}, P_{ID_i}, P_{ID_B}, P_{ID_B}, P_{pub})$

(d) Calculate $s_i = u_i + h_i(x_{ID_i} + s_{ID_i})$

(e) The user sends a signed ciphertext ID the smart meter to the regional gateway $\sigma_i = (U_i, c_i, s_i, \text{timestamp})$

④ After receiving n signs, the regional intelligent gateway calculates $S = \sum_{i=1}^{n} s_i$,after receiving n signs, c, s, timestamp, and the aggregate signcode is $\sigma\langle\{U_i, c_i\}_{i=1}^{n}, S, \text{timestamp}\rangle$ and then calculates $h_i = H_2(c_i, ID_i, U_i, R_{ID_i}, P_{ID_i}, R_{ID_B}, P_{ID_B}, P_{pub})$, and then verify whether the equation $SP = \sum_{i=1}^{n} U_i + \sum_{i=1}^{n} h_i(P_{ID_i} + R_{ID_i} + P_{pub} h_{ID_i})$ is true, and if so, the aggregate signature is forwarded to the control center ID_B; Otherwise, the signing is considered invalid.

⑤ Control Center aggregation unsigning: The algorithm is executed by the control center ID_B, that receives the aggregate ciphertext $\sigma = \langle\{U_i, c_i\}_{i=1}^{n}, S, \text{timestamp}\rangle$ perform the following operations:

a) Calculate $T_i' = (x_{ID_B} + s_{ID_B})U_i, W_i' = H_3(ID_B, T_i')$.

b) Calculate $h_i = H_2(c_i, ID_i, U_i, R_{ID_i}, P_{ID_i}, R_{ID_B}, P_{ID_B}, P_{pub})$.

c) Secondary verification of whether $SP = \sum_{i=1}^{n} U_i + \sum_{i=1}^{n} h_i(P_{ID_i} + R_{ID_i} + P_{pub} h_{ID_i})$ is true, and if so, the output is m_i; Otherwise, the signing is considered invalid.

⑥ The control center analyzes the user's electricity consumption data and pushes the corresponding electricity fee and demand response signal to the power supplier, and the power supplier provides value-added services to the user.

4. Performance analysis

4.1. Calculation overhead comparison

In this topic, the calculation overhead is selected as the reference for the calculation efficiency in different scenarios. Table 1 lists the calculation amounts of the signing algorithm and de-signing algorithm of the proposed scheme and the other four aggregate sign-up schemes and whether they are publicly verifiable, and because the number of hash operations and XOR operations in each scheme is similar and the time of one operation is small enough (3 orders of magnitude difference), only the computational overhead of exponential operations, bilinear operations and scalar multiplication operations in each scheme is considered in the simulation comparison, as listed in Table 1. Where e represents exponential operations, p represents a bilinear operation, s represents a scalar multiplication operation, and represents the number of users. Through literature [6], it is pointed out that the operation time of bilinear logarithmic operation, exponential operation and point multiplication operation under the environment 2.40GHz CPU, 8GB ram and Windows 7 is 8.419ms, 0.996ms, and 0.392 ms, respectively. The calculation of these three scenarios is compared with the computational overhead of the aggregation scheme proposed in this paper. When the number of messages is n, the calculation cost of the proposed scheme is still less than that of the document [8] that uses complex operations, and the total number of operations in literature [9-10] is n-1 and n+2 times more than that of the proposed scheme, respectively, and the computational efficiency of this scheme is higher in both the signing stage and the unsigning stage. Experiments show that [6], the time of aggregate ciphertext decoding is about 1/2 of the unsigning time of n ciphertexts, which greatly improves the verification efficiency of ciphertext. Literature [9] In the stage of aggregate signcode verification, it is necessary to use the value of R_{s}: that is, it is calculated according to the recipient's private key, and any third party cannot calculate the value; Literature [8] In the verification stage of signing secret, although the user's private key is not required, plaintext messages need to be used, and if public verification is carried out, plaintext messages will be leaked, so the above schemes do not meet the public verification.

In summary, this solution has obvious advantages.

Table 1: Comparison of total signcode operation efficiency

scheme	Total number of signatures	Public verifiability
Ref.[8]	$7ns$	×
Ref.[9]	$(7n+3)s$	√
Ref.[10]	$ne+4p+2ns$	×
This scheme	$(5n+1)s$	√

4.2. Communication overhead

Based on literature [10], we analyze literature [10-11] and the scheme in this paper. Assuming that literature [10] has 4 first-level gateways (BG) and 3 second-level gateways (WG), shown in Table 2.

Table 2: Comparison of communication costs

Scheme	Communication overhead
Ref.[10]	$\lvert\sigma\rvert = (n+7)(\lvert c_i\rvert + 2\lvert ID\rvert + \lvert timestamp\rvert + \lvert S_i\rvert)$
Ref.[11]	$\lvert\sigma\rvert = (n+4)(\lvert c_i\rvert + 2\lvert ID\rvert + \lvert timestamp\rvert + 2\lvert S_i\rvert)$
This scheme	$\lvert\sigma\rvert = n(\lvert c_i\rvert + \lvert U_i\rvert + \lvert timestamp\rvert) + \lvert S\rvert$

Literature [11] gives that the ciphertext $|c_i|$ length is 1024 bit, $|ID|$, $|U_i|$, $|timestamp|$ are all 32bit, $|S_i|$ is 160 bits, and the communication overhead of the three schemes under different number of members is shown in Figure 2.

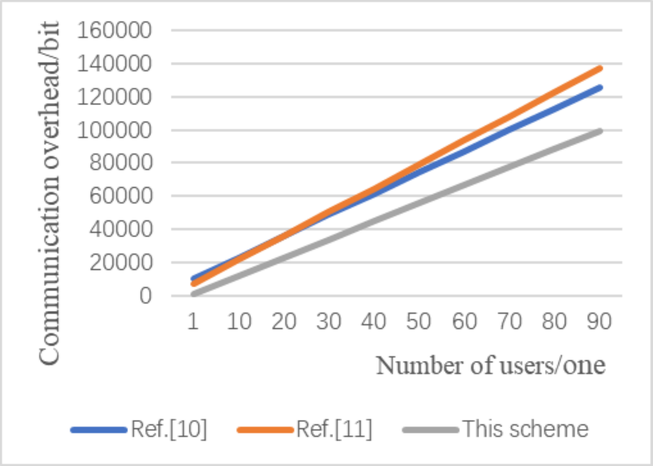

Figure 2: Communication overhead

Figure 2 shows that the communication overhead is less than in the literature [10], [11], which is applicable to the scenario in this article.

5. Conclusion

Aiming at the problem of privacy leakage of user electricity data in the advanced two-sided system, this paper studies a publicly verifiable aggregation ring signcrypt scheme, which can be verified without decryption, we prove that the proposed scheme has high computational efficiency, and can support the control center to perform statistical analysis on user fine-grained data. However, in the assumption of this article, the control center is a highly trusted entity, once it is attacked, it will produce a single point of failure, loss of all users of fine-grained power consumption data and corresponding analysis results, so how to further improve the security value of the control center we further research

References

[1] Zhou Y, Chen X, Chen M. Privacy-Preserving Multidimensional Data Aggregation Scheme for Smart Grid [J]. Security and Communication Networks, 2020, 2020(5):1-14.

[2] Desai S, Alhadad R, Chilamkurti N, et al. A survey of privacy preserving schemes in IoE enabled Smart Grid Advanced Metering Infrastructure [J]. Cluster Computing, 2019, 22(3).

[3] Grid N.S., Guidelines for smart grid cyber security: vol. 2, privacy and the smart grid. Guideline (2010)

[4] Dinesh C., Nettasinghe B.W., Godaliyadda R.I., Ekanayake M.P.B., Ekanayake J., Wijayakulasooriya J.V.: Residential appliance identification based on spectral information of low frequency smart meter measurements. IEEE Trans. Smart Grid 7(6), 2781–2792 (2016)

[5] Ambrosin M., Hosseini H., Mandal K., Conti M., Poovendran R.: Despicable me (ter): anonymous andfine-grained meteringdata reporting with dishonest meters. In: 2016 IEEE Conferenceon Communications and Network Security (CNS), pp. 163–171(2016)

[6] Liu Y, Zhong L, Qiu J, et al. Unsupervised Domain Adaptation for Non-Intrusive Load Monitoring Via Adversarial and Joint Adaptation Network [J]. IEEE Transactions on Industrial Informatics, 2021, PP (99):1-1.

[7] Abouelkheir E, El-sherbiny S. Pairing free identity based aggregate signcryption scheme [J]. IET Information Security, 2020, 14(6): 625-632.

[8] LI C, QI Z H. Efficient and secure certificateless aggregate signcryption scheme [J]. Computer Technology and Development, 2020, 30(10): 117-122.

[9] NIU S F, LI Z B, WANG C F. Anonymous heterogeneous aggregation signcryption scheme for vehicular network [J]. Computer Engineering & Science. 2019, 41(1): 80-87.

[10] Qu Y, Mu Q. An efficient certificateless aggregate signature without pairing [J]. International Journal of Electronic Security and Digital Forensics, 2018, 10(2): 188-203.

[11] H. Zhou, J. Chen, et al. A multidimensional data aggregation scheme in multilevel network in smart grid [J]. Cryptologic Res, VOL: 4, Pages: 114–132, Sep 2017.