Chen *et al. Journal of Cloud Computing* (2022) 11:20
https://doi.org/10.1186/s13677-022-00295-5

Journal of Cloud Computing:
Advances, Systems and Applications

RESEARCH

Open Access

ES-PPDA: an efficient and secure privacy-protected data aggregation scheme in the IoT with an edge-based XaaS architecture

Quan Chen[1], Liangshun Wu[2] and Congshi Jiang[1]*

Abstract

In an Internet of Things (IoT) system based on an anything as a service (XaaS) architecture, data are uploaded from heterogeneous nodes in a nonstandardized format and aggregated on the server side. Within this data-intensive architecture, privacy preservation is one of the most important issues. In response to this concern, there are numerous privacy-protection data aggregation (PPDA) solutions available for various IoT applications. Because of the limited resources of intelligent IoT devices, traditional PPDA cannot meet practical privacy and performance needs. To tackle this challenge, we provide a more efficient and secure PPDA solution that guarantees data security and integrity through Paillier homomorphic encryption and online/offline signing technology. Detailed security analysis shows that our system is unpredictable under a chosen message attack, and the data integrity may be guaranteed under the assumption of q-strong Diffie-Hellman (Q-SDH). We choose an M/G/1 priority queue model to maximize system performance. M/G/1 enhances queuing efficiency and accelerates channel access, thus reducing waiting time and increasing reliability. The experimental results show that our data aggregation scheme is reliable with low latency.

Keywords: XaaS, Data aggregation, The internet of things, Security

Introduction

The Internet of Things (IoT) brings together "things" in the physical world. The traditional IoT model is cloud-based, with end devices solely responsible for data collection. Cloud-based subscription models are evolving to offer greater flexibility and efficiency to meet strong business demand. Anything as a service or everything as a service, also known as X as a service (XaaS), is a platform for SaaS, PaaS, and IaaS. XaaS provides built-in services such as software, networks, platforms, security services, and applications. XaaS also offers high cloud storage capacity and dynamic resource allocation, which reduces enterprise burden. XaaS maximizes performance while minimizing latency at peak loads. New XaaS models allow users to lease cloud resources on demand without having to purchase ownership and can be flexibly expanded upon request at a later stage. Cloud-based XaaS can guarantee a uniform format of different data sources after downloading, which is conducive to setting industry standards. XaaS holds great potential for smart services such as smart grids [1], smart healthcare [2], smart cities [3], and vehicle detection. However, due to bandwidth limitations and limited hardware resources, a traditional cloud-based paradigm can be difficult to accommodate, particularly for real-time services. Hence, edge-based XaaS has been increasing [4]. As shown in Fig. 1, data are collected via sensors and routed to edge servers for local aggregation, sharing, and extraction. The data are then forwarded to the cloud for final processing. Edge servers can be considered preprocessing units that deliver efficient local services using a combination of cloud servers. Thus, the consumption of physical resources can be significantly reduced.

---

*Correspondence: 00200286@WHU.edu.cn

[1] School of Remote Sensing Information Engineering, Wuhan University, Wuhan 430072, China
Full list of author information is available at the end of the article

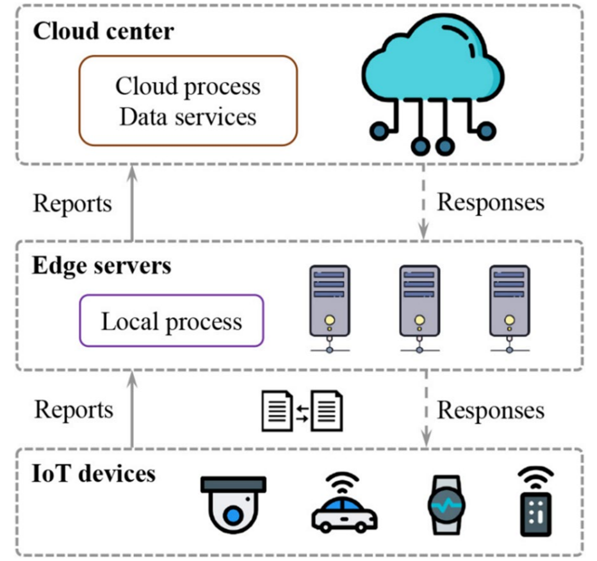

Fig. 1 Data aggregation in an edge-based XaaS architecture

Physical objects connected to an XaaS IoT based on the periphery operate on devices with limited resources and require effective communication protocols to enhance energy efficiency. Traditional internet status transfer protocols, such as REST, utilize event-based frameworks to minimize the number of messages sent. In an edge-based XaaS IoT, heterogeneous applications use these standards and protocols for aggregating edge-side data. Organizations such as OpenIoT, the AllSeen Alliance, and the IPSO Alliance are working to standardize communication protocols to ensure interoperability among vendor islands. The IoT in smart cities focuses on using light protocols, such as CoAP and XMPP, to connect sensor interfaces to physical supports. Organizations such as the IETF and the XMPP are working on expanding CoAP and XMPP. However, these efforts seek to improve protocols rather than provide integrated solutions.

Edge-based XaaS devices are vulnerable due to their distributed nature, which allows more attack paths for both internal and external attackers. Some devices are not completely reliable, revealing user privacy. Put another way, edge-based XaaS undermines the confidentiality, accuracy, and robustness of data aggregation protocols. External attackers can spy on communication channels between the entities involved, alter messages on the network, falsify signatures, or even launch rebroadcasting attacks.

In response to this concern, many privacy protection data aggregation (PPDA) systems have been suggested. Most use homeomorph cryptosystems to implement specific functions, such as summative computation, to ensure the privacy of data in transit. Other measures have been considered to improve the safety of the PPDA, such as dimensional reduction, data integrity verification, and random noise techniques.

However, existing PPDA regimes present practical difficulties.

- Frequent data transfer is normal for an XaaS system based on the edge [4]. When performing live data processing tasks, high communication latency is not acceptable.
- Authentication and validation of data sources are essential to prevent attackers from falsifying, altering, and replaying messages and signatures. However, authentication and validation require the support of edge devices, making it difficult to implement them in resource-constrained IoT devices [5].

In what follows, to address the above challenges, efficient and secure privacy-protection data aggregation (ES-PPDA) is suggested. ES-PPDA considers both security and efficiency. In ES-PPDA, the heavy computational costs associated with data integrity operations can be drastically reduced through online/offline signing mechanisms. The main contributions are summarized as follows:

- Paillier homomorph encryption and online/offline signing are used to ensure data security and integrity.
- A detailed security analysis is presented.
- An M/G/1 priority queue model is utilized to optimize system performance. M/G/1 improves queuing efficiency and speeds up channel access, thus reducing wait time and increasing reliability. Experimental results show that ES-PPDA is reliable with low latency

The remaining content is organized as follows: Related work section reviews related work; Preliminaries section introduces background information; Proposed scheme section proposes the ES-PPDA scheme; Security analysis section carries out security analysis; Performance optimization section optimizes the scheme; Experiments section conducts performance testing; and Conclusion section concludes.

Related work

Recently, the aggregation of privacy data (PPDA) has received growing attention in areas such as smart grids [6] and vehicle detection systems [7, 8]. Some earlier work has considered uploading using homomorph cryptosystems [9, 10]. Subsequently, to further enhance the ability to protect privacy, it is necessary to add blind factors in the encryption steps, thus making them resistant to internal attacks [11, 12]. To prevent malicious aggregators, hatch hash [13] and random noise technology are introduced into PPDA schemes to ensure data integrity in the encrypted message forwarding process [14]. However, these systems do not consider the cost of designing a cryptosystem [15].

Recently, researchers have focused on reducing the cost of computing cryptographic operations in conventional PPDA systems [16]. By anticipating the demand for electricity in a smart grid system, a light and secure system was proposed. It can effectively meet safety and privacy requirements and further reduce the indirect costs of communications. An improved version for IoT fog computing systems allows multidimensional data to be compressed into composite dummy data and the early injection at the fog nodes can be filtered. Researchers have proposed a system for classifying and aggregating privacy data for vehicle sensor systems that resist data link attacks [17]. A PPDA scheme presuming that several locally authenticating accreditation bodies can be anonymous, a dual trap chameleon hash-based online/offline signing and verification method, was proposed [18]. The scheme proposed in [19] protects data privacy by hiding data transfers between the cloud and edge servers. The scheme proposed in [20] used RSA and RC4 to generate a key, which achieves higher security in the image. Recently, there have been data aggregation optimizations, for example, using a hybrid metaheuristic algorithm, i.e., a whale optimization algorithm (WOA) and simulated annealing (SA) algorithm, to select the optimal CH in an loT network cluster [21]; optimizations based on LSTM models [22]; and reducing network latency by increasing time slots and reducing the power consumption through weighted load balancing [23] (Table 1).

Table 1 Comparison of data aggregation schemes

Ref.	Network Latency	Power Consumption	Data Integrity	Privacy Protection
[6]	√	×	×	×
[7]	√	×	√	×
[8]	×	√	×	×
[9]	×	√	×	×
[10]	√	×	√	
[12]	×	×	√	√
[11]	×	×	√	×
[13]	√	√	√	×
[14]		√	×	√
[15]	√	√	×	√
[16]	√	×	√	√
[17]	×	√	×	√
[18]	×	√	×	×
[19]	×	√	×	√
[20]	√	×	×	√
[21]	×	√	×	√
[22]	√	×	√	
[23]	√	√	√	√

√: item is considered, ×: item is not considered

Preliminaries

This section presents several definitions and notations used in ES-PPDA schemes, including bilinear pairings, Paillier homomorph cryptosystems [24], online/offline signing, and security definitions.

Bilinear pairings

G and G_T are two cyclic groups, g is the generator of group G, and p is the prime order. $e: G \times G \to G_T$, satisfying [25]:

- $e(u^a, v^b) = e(u,v)^{ab}$, where $u, v \in G$, $a, b \in Z_p{}^*$;
- $e\left(g,g\right) \neq 1_{G_T}$,
- $e(u,v)$ is computable, where $u, v \in G$.

Definition 1*(q-strong Diffie-Hellman problem (q-SDH)). x is a random element in* $Z_p{}^*$. *For* $\left(g, g^x, g^{(x^2)}, \dots, g^{(x^q)}\right)$, *and pair* $(m, \sum_x)$, *where* $m \in Z_p{}^*$. *The q-SDH is defined as an* (q, t, ε) *problem:*

$$\Pr\left[A\left(g, g^x, \dots, g^{(x^q)}\right) = \left(m, \sum\nolimits_x\right), m \in Z_p^*\right] < \varepsilon \tag{1}$$

Paillier homomorph cryptosystem

A Paillier homomorph cryptosystem satisfies addition and multiplication homomorphism.

(a) Key generation algorithm

Step 1 Pick two large prime numbers at random that satisfy $gcd(pq,(p-1)(q-1)) = 1$.
Step 2 Compute $n = pq$ and $\lambda = lcm\,(p-1, q-1)$.
Step 3 Define $L(x) = \frac{x-1}{n}$
Step 4 Randomly select an integer g less than positive n^2, and there exists$\mu = (L(g^{\lambda} \,mod\, n^2))^{-1} \,mod\, n$.
Step 5 The public key is (n,g), and the private key is (λ, μ).

(b) Encryption

Assume m is plaintext, $0 < m < n$

Step 1 Select a random number r, $0<r<n$, and $gcd(r,n)=1$.

Step 2 Encrypt: $c=g^m \cdot r^n \, mod \, n^2$

(c) Decryption

Compute $m=L(c^\lambda \, mod \, n^2)\cdot\mu \, mod \, n$

Online/offline signing

The double trapdoor chameleon hash (DTCH) function is usually used for implementing online/offline signing. For the generator g_1 of a prime number p_1 and G_{p_1}, select two trapdoor keys $y,z \in Z_{p_1}^*$. Then, compute hash: $H_{ch}(r,s,u) = g_1^r g_2^r g_3^r$, where $g_2 = g_1^y$, $g_3 = g_1^y$. The DTCH function has the following properties [26]:

- Computability: For $pk \in G$ and the triad $(r,o,o)\in Z_p$, $H_{ch}(r,s,u)$ can be computed in polynomial time.
- Anti-collision: If a key is missing, two hash pairs (r_1,s_1,u_1), (r_2,s_2,u_2) cannot be found such that $r_1 \neq r_2$ and $H_{ch}(r_1,s_1,u_1)=H_{ch}(r_2,s_2,u_2)$.
- Valve collision: Given H_{ch} and (pk,sk), hash pair (r_1,s_1,u_1) and additional message $r_2 \in Z_p$, such that $H_{ch}(r_1,s_1,u_1)=H_{ch}(r_2,s_2,u_2)$. First, a random u_2 (or s_2) is selected, and the value of u_2 (or s_2) can be calculated in polynomial time by $s_2=((r_1-r_2)+(u_1-u_2)y+s_1z)z^{-1}$ or $u_2=((r_1-r_2)+(s_1-s_2)y+u_1z)z^{-1}$.

Based on the above properties of the DTCH function, online/offline signing can be built using five algorithms:

- *Setup*: When the security parameter 1^λ is entered, it returns a public key Ver_{pk} and a private key Sig_{sk}.
- *Sign.off*: The offline signing algorithm turns off $\sum_{off}$ and St on the input signing key Sig_{sk}.
- *Ver.off*: If Ver_{pk} and $\sum_{off}$ are entered, it returns a valid $\sum_{off}$ value. Otherwise, reject is output.
- *Sign.on*: Returns the online signing token $\sum_{on}$ input Sig_{sk}, status information St, and message m.
- *Ver.on*: When Ver_{pk} is entered, message M is displayed and returns accept. Otherwise, it returns reject. The signature of m is defined as $\sum = (\sum_{on}, \sum_{off})$.

Security rule

Definition 2 *(Unforgeable). The mechanisms are unforgeable if under a chosen message attack, which can be considered an attacker challenge game. Suppose an opponent A can query trailers $(sig^{on}(sk,St_i,m_i),sig^{off}(sk))$ multiple times, where $St_i$is the status information of the signer* [27]*:*

- Initiation: Challenger C generates public/private keys (pk,sk) from 1^k. Then, pk is given to A.
- Sign.off query: The opponent requests and the challenger C replies with $\sum_i^{off}$ to the opponent, while the status information St_i is stored by itself. Assume that the opponent can make up to q_1 queries at this stage.
- Sign.on query: The opponent requests and the challenger C uses the St_i to calculate the online signature and then returns $\sum_i^{on}$ to the opponent. Assume that the opponent can make a maximum of q_2 queries at this stage.
- Forgery: Opponent A generate $(m^*,\sum^*)$ and forward to C. The challenger C checks by $Ver_{on}(pk,m^*,\sum^*)$. If the signature is valid, output 1 (success); otherwise, output 0 (failure).

The existing advantages of forging opponent A's signature are as follows:

$$Adv_A = \Pr\left[\begin{array}{l} Ver_{on}\left(pk,m^*,\sum^*\right) = 1 : (pk,sk) \leftarrow \\ KeyGen\left(1^k\right), \left(m^*,\sum^*\right) \leftarrow A^{\left(\sum^{off},\sum^{on}\right)} \end{array}\right] \tag{2}$$

Proposed scheme

The notations used in this section are listed in Table 2.

System model

The ES-PPDA system model is shown in Fig. 2 and consists of four components: cloud servers, edge servers

Table 2 Notations and symbols

Notation	Description
(k,k_1)	Security parameters
SP_{pub}	Public key of SP
msk	Master key
X_i	Random number
k_i	Blind factor
α_i, β_i	Registration knowledge
(y,z)	Random numbers
St	Status information of signer
$\sum_i^{off}$	Offline signature
$\sum_i^{on}$	Online signature
Ver_{on}	Online signature verification
Ver_{pk}	Verify with public key
c_i	Encrypted report
v_i	An integer
m_i	Message
Y_i	Public key
$\sum_{Agg}$	Aggregate signature
g_1	DTCH generator
H_i	Hashing function

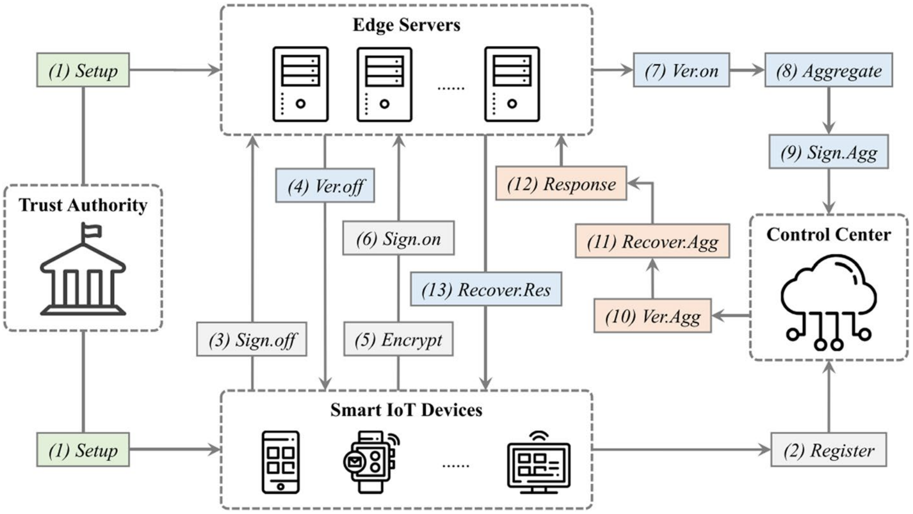

Fig. 2 ES-PPDA system model

(ESs), smart IoT devices (SDs), trust agencies (TAs) (or other control centers, (CCs)) [28].

A TA boots the entire system and distributes critical information and system parameters. When the configuration is finished, the TA disconnects.

A CC collects data packets from the edge. It then sends responses to the edge server (see steps 10, 11, and 12 in Fig. 2). The CC also offers registration services for the XaaS IoT.

An ES acts as an aggregator that processes encrypted data from an SD, and forwards and communicates between the CC and SD (see steps 8, 9, and 13 in Fig. 2). The ES also performs integrity verification (see steps 4 and 7 in Fig. 2).

An SD collects private data generated by sensors and transmits it to the CC in encrypted form via an ES (see steps 2, 3, 5, and 6 in Fig. 2).

Note: since an SD is typically a resource-constrained device, it cannot effectively carry out computationally complex privacy-protected data aggregation processes, particularly cryptography operations involved in data integrity mechanisms. This has led to the exploration of a lightweight/efficient PPDA optimization that supports edge-based XaaS architectures.

Workflow

The proposed ES-PPDA scheme proceeds as follows:

a) **Initialization**

- $Setup(k, k_1) \rightarrow (SP_{pub}, msk)$: When two security parameters (k, k_1) are input, it outputs SP_{pub} and the master key msk.

b) **Registration**

- $\text{Register}(X_i, k_i) \rightarrow (\alpha_i, \beta_i)$: When random X_i and blind factor k_i are set, the output verifies public key Y_i and registration knowledge (α_i, β_i).
- $Sign.off\,(y, z, s_i, u_i) \rightarrow \left(St, H_{ch_i}, \sum_i^{off}, Ver_{on}\right)$: When inputting random (y, z) and (s_i, u_i), it outputs St, $\sum_i^{off}$ and Ver_{on}.

c) **Report Generation**

- $Ver.off\left(Ver_{pk}, \sum_i^{off}\right) \rightarrow b_1$: When the input is Ver_{pk} and $\sum_i^{off}$, it outputs $b_1 \in \{0,1\}$, where $b_1 = 1$ indicates that the offline verification result is accepted and $b_1 = 0$ indicates that it is rejected.
- $Encrypt(PK_p, m_i, v_i) \rightarrow c_i$: When PK_p, the message m_i, and the integer v_i are input, the output is the encrypted report c_i.
- $Sign.on\left(c_i, St, s_i^{'}\right) \rightarrow \sum_i^{on}$: For input $(c_i, St, s_i^{'})$, the output is an online signature $\sum_i^{on}$.

d) **Report Summary**

- $Ver.on(\sum_i^{on}, Ver_{on}) \rightarrow b_2$: For input Ver_{on} and $\sum_i^{on}$, the output is $b_2 \in \{0,1\}$, where $b_2 = 1$ indicates that the online verification result is accepted, and $b_2 = 0$ indicates that the online verification result is rejected.

- *Aggregate*$(c_i) \rightarrow c$: Output the aggregation result.
- *Sign. Agg*$(X_i, c) \rightarrow (Y_i, \sum_{Agg})$: For input c and X_i, it outputs the aggregate signature public key Y_i and the aggregate signature $\sum_{Agg}$.

e) Report Reading

- *Ver. Agg*$(Y_i, \sum_{Agg}) \rightarrow b_3$: For input Y_i and $\sum_{Agg}$, it outputs $b_3 \in \{0, 1\}$, where $b_3 = 1$ indicates that the verification result of aggregation is accepted, and $b_3 = 0$ indicates that the verification result of aggregation is rejected.
- Recover. *Agg*$(c) \rightarrow m$: Output m.

f) Response

- Reponse$\left(e(g_1, g_1)^{\widetilde{\alpha}}, \widetilde{\beta}, Q, Y, M_R\right) \rightarrow \left(\widetilde{C_1}, \widetilde{C_2}, \widetilde{C_3}:\right)$ For $\widetilde{\beta}, Q$, public key $\left(e(g_1, g_1)^{\widetilde{\alpha}}, Y\right)$ and response message M_R, the output is response ciphertext $\left(\widetilde{C_1}, \widetilde{C_2}, \widetilde{C_3}\right)$.
- Recover.Res$\left(ak_i, \widetilde{C_1}, \widetilde{C_2}, \widetilde{C_3}\right) \rightarrow M_R$: For input response ciphertext $\left(\widetilde{C_1}, \widetilde{C_2}, \widetilde{C_3}\right)$ and authorization key ak_i, the output is response message M_R.

Security claims

We assume that both the TA and CC are completely trustworthy. The ES may be partly trustworthy, that is, it will not manipulate the sensitive user data, but may reveal personal information in the grouping process. Furthermore, the external opponent A threatens the data integrity and carries out attacks. It can spy on the transmitted data or invade the server in the ES and CC to steal the processed data. The opponent can actively falsify the signature of the data report and further damage the integrity of the data.

Security analysis

In what follows, we examine system security in terms of authentication, confidentiality, and privacy protection.

Authentication

In ES-PPDA, we integrated Schnorr's extended signature method into the recording step, which turned out to be safe under the discrete logarithmic hypothesis. The explanation is as follows:

$$g_1^{\beta_i} Y_i^{H_2(\alpha_i)} = g_1^{(r_i - X_i H_2(\alpha_i))} \cdot g_1^{X_i H_2(\alpha_i)} = g_1^{r_i} = \alpha_i \quad (3)$$

An attacker cannot tamper with recording without knowing the true identifier of SD_i, the ID_i, because the ID_i is obtained using hashing (H_i) and is a secret. Furthermore, although an attacker can steal the true identifier ID_i of SD_i, it still cannot obtain r_i because r_i is further hidden by a randomly selected blind factor k_i, thus assuring that X_i is secure. Therefore, our scheme proves that the SD-CC authentication is secure.

Confidentiality

We use the Paillier cryptosystem to encrypt all sensory data and aggregate encrypted text based on additive homomorphism. Confidentiality is assured based on the following three points.

First, SD_i 's private data m_i are encrypted as $c_i = g^{m_i} \cdot {v_i}^n \bmod n^2$. The Paillier cryptographic system is semantically secure under CPA based on q-SDH and does not reveal sensitive information.

Second, when aggregating reports, the ES cannot retrieve every individual's full text, and the ciphertext received is aggregated as $c = g^{\sum_{i=1}^{\omega} m_i} \cdot \left(\prod_{i=1}^{\omega} v_i\right)^n \bmod n^2$. Therefore, the confidentiality and privacy of user data can be guaranteed even when the ES is not trusted.

Finally, let us assume that an outside attacker could spy on the entire communication channel SD_i to the CC and simultaneously obtain a single ciphertext c_i, aggregated ciphertext c, and plaintext m but still cannot recover a single plaintext m_i. All plaintext is compressed through the process of aggregating reports. In summary, the confidentiality and privacy of each SD_i 's private data can be well protected.

Integrity and unforgeability

In the proposed scheme, we develop a signing method that reduces the cost of calculation while ensuring data integrity. Here, we show that our system is robust under a chosen message attack. Because of definition 2, without asking for an Oracle token EZ online signature, an opponent may not counterfeit any probabilistic polynomial-time pair $(m^*, \sum *)$.

Performance optimization

Suppose we have N heterogeneous sensor nodes within an $l \times l$ region (i.e., rectangular industrial subunit). The data captured by the sensors fall into two categories: normal data (ND) and event data (ED). Low priority P_1 nodes generate ND packets, and high priority P_h nodes generate ED packets when the value exceeds its threshold. Suppose that each node only supports one type of data, i.e., ND or ED. Similarly, M of N nodes send high-priority packets, namely, P_h packets, while the rest send only low-priority packets, namely, P_1 packets. Network topologies are considered static for a certain period. The gateway and cloud center are expected to be connected via broadband wireless links, and latency and packet loss are negligible.

Sensor nodes are connected to the channel (CH) aggregator. Nodes, including CH and gateway, have child–parent relationships. All sensor nodes within a single CH compete for the respective parent node access channel for link resources. The data generated from the end node are aggregated to the CH and then forwarded to the gateway. Gateways and CHs are located in specific areas and generally have greater electrical power than sensor nodes. A CH can retrieve application-specific information, including priority and location. The waiting time for each priority depends upon the scheduling policy that the CH has adopted.

The M/G/1 queue method accommodates the randomness of devices for measuring network performance, including throughput, waiting time, packet loss rate, and resource consumption [29]. The M/G/1 queue system with priority may be divided into nonrepetitive and preemptive queue models. For nonrepetitive package planning, when the lower priority package starts to run, the ongoing task continues even though the top priority packet hits the queue. Additionally, the package should wait in the queue until the task for the package is complete. However, in scheduling priority packages, higher priority packages are handled first, and lower priority packages may be preemptive by backing up their context if the task has already been run. We propose to use an M/G/1 to CH priority queue model.

Priority data partitioning is built by the application layer taking into account the parameters of the MAC layer depending on industrial requirements and network conditions. IEEE 802.15.4 uses the carrier-sense multiple access with collision avoidance (CSMA/CA) conveyor to access wireless channels. However, it does not suit delayed industrial applications because it does not have priority characteristics and delayed intervention [30]. In industrial IoT systems, flow control, process monitoring, and fault detection subsystems must have media access mechanisms that are sensitive to delays and priorities.

Figure 3 shows a sequence diagram of various nodes in competition for channel access depending on the priority of the nodes. All packets in the lower priority queue need not be processed until the higher priority queue is blank. The P_h node still has a short, fixed withdrawal period, more frequent channel access detection, and many retreats. However, P_l nodes use longer, random withdrawal times, fewer detection frequencies, and shorter withdrawal times. Moreover, the clear channel assessment (CCA) detection time of a P_l node is more continuous than the CCA and P_h node removal period.

CSMA/CA behavior is influenced by various MAC settings, such as minimum and maximum withdrawal indexes (*macMinBE*), the maximum withdrawal indexes (*macMaxBE*), the initial values of competing windows (*CW*), and the maximum backoffs (*macMaxCSMABackoffs*). The different values of these MAC settings significantly affect the performance of an IoT network. Instead of having to configure the same CSMA/CA parameter values (i.e., low priority and high priority) for both traffic types, each category may have its attributes assigned

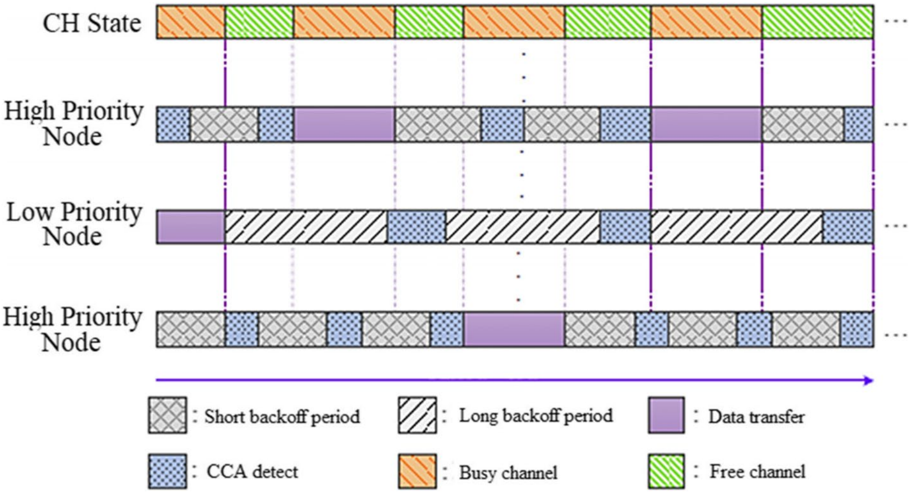

Fig. 3 Channel access priority

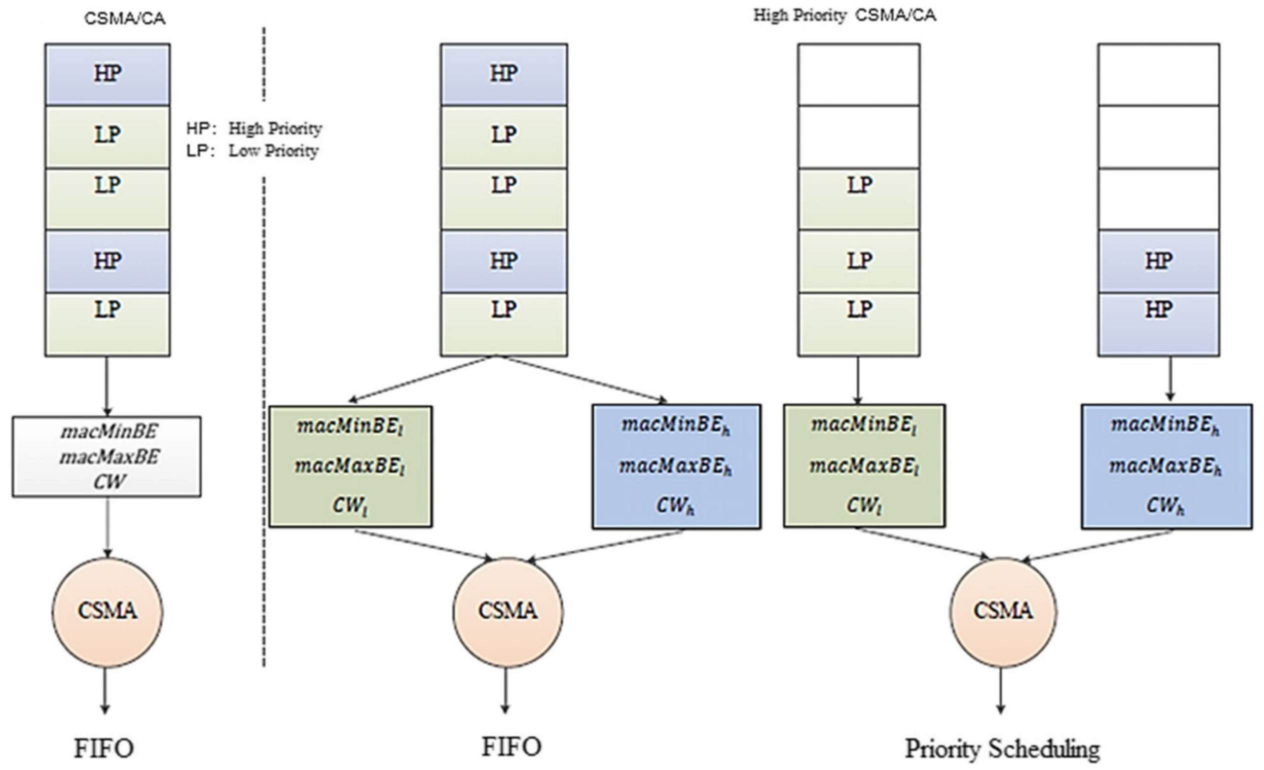

Fig. 4 CSMA/CA FIFO scheduling and priority scheduling

to it. This study defines $[macMinBE_h, macMaxBE_h]$ and CW_h as the high-priority nodes backoff window value interval and competition, and $[macMinBE_l, macMaxBE_l]$ and CW_l are defined as the values of low-priority nodes. In addition, by specifying different CSCM/CA parameters, it is possible to implement prioritized scheduling to reduce channel access times for high-priority packets, as shown in Fig. 4.

For data aggregation with priority, the M/G/1 queue model with priority maintains the data priority category. Packets with priority i have arrival rates of λ_i, $\lambda_i \in \{1, 2, \ldots, \mathrm{P}\}$, and follow a Poisson distribution. A lower value of i indicates a higher priority packet type. Within the system model, the priority rule is implemented. This means that the arrival of the i-priority packet immediately precedes the lower priority data and obtains service access.

The wait time for i priority packets W_i is the queue time before the CH. The average remaining service time of existing service packets and the CH service time are represented by R_i and S_i, respectively. The total system delay is given by the sum of the packet's wait time and serve time. Little's law states that the expected wait time for the i-th priority packet is:

$$E[W_i] = \frac{\sum_{j=1}^{i} \rho_j E\left[R_j\right]}{(1-(\rho_1+\cdots+\rho_i))(1-(\rho_1+\cdots+\rho_{i-1}))} \tag{4}$$

Where $\rho_i = \lambda_i E[S_i]$, $E[S_i]$ is the expected service time, and $E[R_i]$ is the expected time remaining. $E[S_i]$ and $E\left[D_i^{sys}\right]$ are the expected service time of the i-th priority packet and the expected system delay in the i-th priority queue, respectively, which are calculated by the following formula:

$$E\left[\hat{s}_i\right] = \frac{E[S_i]}{(1-(\rho_1+\cdots+\rho_{i-1}))} \tag{5}$$

$$E\left[D_i^{\hat{sys}}\right] = E\left[\hat{s}_i\right] + E[W_i] \tag{6}$$

Furthermore, the second moment can be expressed in the following manner:

$$E[R_i] = \frac{2}{3}\lambda_i E[S_i]^2 \tag{7}$$

$$E\left[S_i^2\right] = \frac{4}{3}E[S_i]^2 \tag{8}$$

Table 3 Simulation parameters

Parameter	Value
Max backoff	5
Min backoff	3
Max CSMA backoff	4
Queue size	51 frames
MAC frame payload	800 bits
ACK size	88 bits
Data rate	19.2 kbps
MAC overhead	48 bits

Experiments

ES-PPDA's performance is assessed based on the anticipated latency and reliability of the system, which is implemented in MATLAB. The simulation parameters are identified in Table 3.

System latency

Figure 5 shows the package latency having different prioritization and the quantity of nodes. The latency for high/low-priority packets increases with the quantity of nodes, as aggregating more packets leads to longer service durations. The latency of low-priority packets is longer than that of high-priority packets because we have to take into account the disruption of all high-priority packets.

Additionally, Fig. 6 compares the performance of the proposed priority scenarios against the nonpriority scenarios. Nonpriority regimes show similar curves, but the latency exceeds priority methods. In addition, because of the preferential channel access and the preemptive priority rule, a high-priority packet is free from interference from a lower packet, thereby reducing the expected system time.

System reliability

Our scheme is modeled as a K-size M/G/1 priority queue. Each queue receives packet data frames per second using Poisson's arrival process of λ. The probability of packages being in the queue is:

$$p_i = \frac{\rho^i}{\sum_{j=0}^{K} \rho^j} \tag{9}$$

A sensor node may not be capable of sending packets to the CH, including (a) if the buffer is full, (b) if the node cannot find a free channel, or (c) if the packet is thrown past the retry limit. Considering these aspects, the reliability of the η system can be calculated as follows:

$$\eta = (1 - p_k)\left(1 - p_{cf}\right)(1 - p_{cr}), \tag{10}$$

where p_k is the probability of the entire buffer with K frames, provided by Eq. (9), p_{cf} is the packet loss resulting from channel access failure, and p_{cr} is the packet collapse resulting from retry.

Figure 7 illustrates the relationship between reliability and the node number of the entire system that

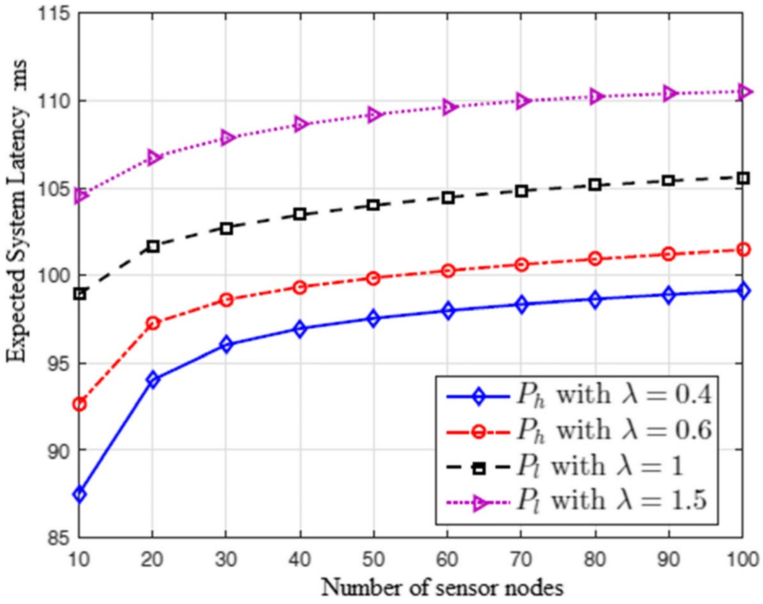

Fig. 5 Expected system latency

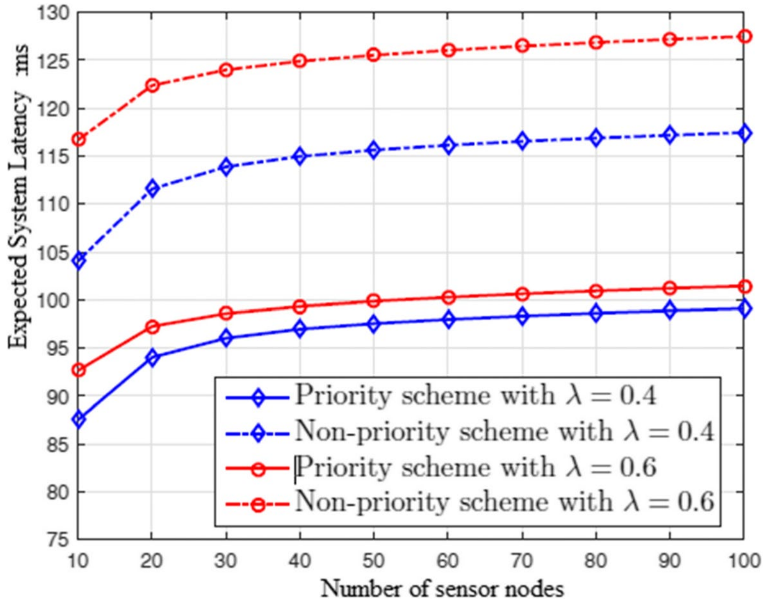

Fig. 6 Performance comparison of the proposed priority scheduling to nonpriority scheduling

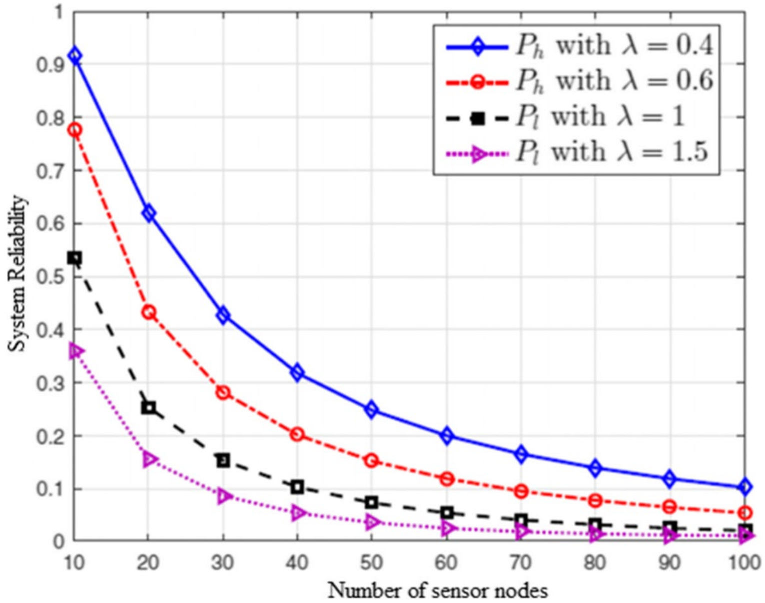

Fig. 7 Performance evaluation in terms of reliability

is observable and that the reliability of the network increases and diminishes the number of nodes. Due to the node number, each node in the queue congestion problems are conflicts become more frequent, more frequent and packet retransmission. Then, as the queue becomes busier and delayed longer, the possibility of frame loss is also increased due to conflicts, retry constraints, and link constraints. It should be noted that high-priority nodes have greater network reliability than low-priority nodes because of the use of the priority channel planning mechanism and the queue strategy.

An IoT network typically involves many sensors for detection. In a high-density IoT network, resource-constrained end devices may be limited by packet delay and data conflict. The end devices typically contain various data flows and face various reliability requirements. This paper proposes a cloud-based delay reduction plan using preferential channel access and data aggregation at CHs. Furthermore, the combined effects of packet planning and aggregation are considered using a preemptive M/G/1 queue model. Experimental results have shown that the priority system has significantly decreased the wait time and increased the reliability of the nonpriority system. Then, the network emulator tool was used to analyze system performance in real IoT applications such as e-health and industrial automation.

A future IoT network is expected to support a wide range of heterogeneous equipment/sensors in areas such as e-health and industrial control. In high-density deployment scenarios such as an industrial internet system, reliable communication links with low latency are difficult because of the latency of the system involved. Using the information offered by the application, the data from IoT nodes of two types, the high-priority nodes and low-priority nodes, allocate different MAC layer properties to provide priority channel access mechanisms for data processing with the heart of the cloud. Then, before sending the aggregated data to the cloud, using a separate low-priority, high-level queue, the m/G/1 preemptive queue model is adopted. The results show that, in comparison with the nonpriority regime, the basic method proposed in this paper can significantly improve the timing and reliability of an IoT system.

Conclusion

In this paper, we propose a secure and efficient PPDA solution for IoT systems based on an XaaS architecture. Our scheme greatly reduces the time of resource consumption. In addition, by taking advantage of edge computing, ES-PPDA can effectively transfer complex cryptographic operations to ES while minimizing the real-time cost. We select an M/G/1 queuing model to optimize the system performance. This optimization can be applied to an XaaS architecture IoT, for example, a smart grid. Experimental results show that the scheme is unassailable under the security model we defined. Performance evaluation experiments proved that the scheme is lightweight and highly efficient. However, our approach is somewhat vulnerable to malicious users such as ESs. Subsequently, we plan to make our security model robust.

Abbreviations

PPDA: Privacy-protection data aggregation; IoT: Internet of Things; XaaS: Anything as a service; ES: Edge server; SD: Smart IoT device; TA: Trust agency; CC: Control center; Q-SDH: q-strong Diffie-Hellman; DTCH: Double trapdoor chameleon hash; ES-PPDA: Efficient and secure PPDA; ND: Normal data; ED: Event data; CH: Channel; CSMA/CA: Carrier-sense multiple access with collision avoidance; CCA: Clear channel assessment.

Acknowledgments

The authors wish to acknowledge Prof. Yulin Wang for his help with the paper structure and methodology.

Authors' contributions

Quan Chen: data curation, formal analysis, investigation, and methodology. Liangshun Wu: software, validation, visualization, and writing. Congshi Jiang: conceptualization, ideas, funding acquisition, project administration, resource allocation, and supervision. All authors reviewed the manuscript. The author(s) read and approved the final manuscript.

Funding

The research received no funds.

Availability of data and materials

The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

Competing interests

The authors declare no conflicts of interest.

Author details

[1]School of Remote Sensing Information Engineering, Wuhan University, Wuhan 430072, China. [2]School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China.

Received: 13 June 2022 Accepted: 13 July 2022

References

1. Bonomi F, Addepalli R (2018) Fog computing and its role in the internet of things. ACM. https://dl.acm.org/doi/abs/10.1145/2342509.2342513
2. Fredj SB, Boussard M, Kofman D, Noirie L (2013) A scalable IoT service search based on clustering and aggregation. In: IEEE international conference on green computing & communications & IEEE internet of things & IEEE cyber. IEEE Computer Society. https://ieeexplore.ieee.org/abstract/document/6682100
3. Fan T, Chen Y (2010) A scheme of data management in the internet of things. In: IEEE international conference on network infrastructure & digital content. IEEE. https://ieeexplore.ieee.org/abstract/document/5657908
4. Shi W, Jie C, Quan Z, Li Y, Xu L (2016) Edge computing: vision and challenges. Internet Things J IEEE 3(5):637–646
5. Hattab G, Cabric D (2018) Performance analysis of uplink cellular IoT using different deployments of data aggregators. In: GLOBECOM 2018 - 2018 IEEE global communications conference. IEEE. https://ieeexplore.ieee.org/abstract/document/8647923
6. Darzi S, Akhbari B, Khodaiemehr H (2022) Lpm2da: a lattice-based privacy-preserving multi-functional and multi-dimensional data aggregation scheme for smart grid. Clust Comput 25(1):263–278
7. Yang Y, Zhang L, Zhao Y, Choo K, Zhang Y (2022) Privacy-preserving aggregation-authentication scheme for safety warning system in fog-cloud based vanet. IEEE Trans Inf Forensics Secur PP(99):1–1
8. Jastaniah K, Zhang N, Mustafa MA (2022) Privacy-friendly flexible iot health data processing with user-centric access control
9. Bohli JM, Skarmeta A, Moreno MV, Dan G, Langendorfer P (2015) SMARTIE project: secure IoT data management for smart cities. In: 2015 international conference on recent advances in internet of things (RIoT). IEEE. https://ieeexplore.ieee.org/abstract/document/7104906
10. Jin J, Gubbi J, Marusic S, Palaniswami M (2014) An information framework for creating a smart city through internet of things. IEEE Internet Things J 1(2):112–121
11. Khodadadi F, Calheiros RN, Buyya R (2015) A data-centric framework for development and deployment of internet of things applications in

clouds. In: IEEE tenth international conference on intelligent sensors. IEEE. https://ieeexplore.ieee.org/abstract/document/7106952
12. Fonseca J, Ferraz C, Gama K (2016) A policy-based coordination architecture for distributed complex event processing in the internet of things: doctoral symposium. In: The 10th ACM international conference. ACM. https://dl.acm.org/doi/abs/10.1145/2933267.2933431
13. Chen Q, Ayong YE, Zhang Q, Huang C (2022) A new edge perturbation mechanism for privacy-preserving data collection in iot. Chin J Electron 32:1–12. https://cje.ejournal.org.cn/en/article/doi/10.1049/cje.2021.00.411
14. Luo W, Bai G (2011) Ensuring the data integrity in cloud data storage. In: 2011 IEEE international conference on cloud computing and intelligence systems. IEEE. https://ieeexplore.ieee.org/abstract/document/6045067
15. Wu J, Sheng X, Li G, Yu K, Liu J (2022) An efficient and secure aggregation encryption scheme in edge computing. China Commun 19(3):13
16. Chen L, Fu S, Lin L (2022) Privacy-preserving swarm learning based on homomorphic encryption. In: International conference on algorithms and architectures for parallel processing. Springer, Cham. https://link.springer.com/chapter/10.1007/978-3-030-95391-1_32
17. Dhinakaran D, Joe Prathap PM (2022) Ensuring privacy of data and mined results of data possessor in collaborative ARM
18. Bowers KD, Juels A, Oprea A (2009) HAIL: a high-availability and integrity layer for cloud storage
19. Wang T, Yang Q, Shen X, Gadekallu TR, Wang W, Dev K (2022) A privacy-enhanced retrieval technology for the cloud-assisted internet of things. IEEE Trans Industr Inform 18(7):4981–4989
20. Vashishtha M, Chouksey P, Rajput D, Reddy S, Reddy P, Gadekallu T, Patel H (2021) Security and detection mechanism in IoT-based cloud computing using hybrid approach. Int J Internet Technol Secur Trans 11:436–451
21. Iwendi C, Maddikunta PKR, Gadekallu TR, Lakshmanna K, Bashir AK, Piran MJ (2021) A metaheuristic optimization approach for energy efficiency in the IoT networks. Softw Pract Exper 51:2558–2571
22. Iwendi C, Khan S, Anajemba JH, Bashir AK, Noor F (2020) Realizing an efficient IoMT-assisted patient diet recommendation system through machine learning model. IEEE Access 8:28462–28474
23. Ponnan S, Saravanan AK, Iwendi C, Ibeke E, Srivastava G (2021) An artificial intelligence-based quorum system for the improvement of the lifespan of sensor networks. IEEE Sensors J 21(15):17373–17385
24. Falk J, BJöRK S (2000) Privacy and information integrity in wearable computing and ubiquitous computing. Iso/iec Jtc1/sc29/wg11 Mpeg00/n3705, La Baule, New York, p 177. https://dl.acm.org/doi/abs/10.1145/633292.633390
25. Pietro RD, Mancini LV (2003) Security and privacy issues of handheld and wearable wireless devices. Commun ACM 46(9):74–79
26. Frikken KB, Joseph IV (2008) An efficient integrity-preserving scheme for hierarchical sensor aggregation. In: ACM conference on wireless network security. ACM. https://dl.acm.org/doi/abs/10.1145/1352533.1352546
27. Chen F, Liu AX (2012) Privacy- and integrity-preserving range queries in sensor networks. IEEE/ACM Trans Networking 20(6):1774–1787
28. Yang J, He S, Lin Y, Lv Z (2015) Multimedia cloud transmission and storage system based on internet of things. Multimed Tools Appl 76:1–16. https://link.springer.com/article/10.1007/s11042-015-2967-9
29. Xu LD, He W, Li S (2014) Internet of things in industries: a survey. IEEE Trans Industr Inform 10(4):2233–2243
30. Dastjerdi AV, Buyya R (2016) Fog computing: helping the internet of things realize its potential. Computer 49(8):112–116

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.