PAAL: A Framework based on Authentication, Aggregation and Local Differential Privacy for Internet of Multimedia Things

Muhammad Usman, *Member, IEEE,* Mian Ahmad Jan*, *Member, IEEE,* and Deepak Puthal, *Member, IEEE*

***Abstract*—Internet of Multimedia Things (IoMT) applications generate huge volumes of multimedia data that are uploaded to cloud servers for storage and processing. During the uploading process, the IoMT applications face three major challenges, i.e., node management, privacy-preserving, and network protection. In this paper, we propose a multi-layer framework (PAAL) based on a multi-level edge computing architecture to manage end and edge devices, preserve the privacy of end-devices and data, and protect the underlying network from external attacks. The proposed framework has three layers. In the first layer, the underlying network is partitioned into multiple clusters to manage end-devices and Level-One Edge Devices (LOEDs). In the second layer, the LOEDs apply an efficient aggregation technique to reduce the volumes of generated data and preserve the privacy of end-devices. The privacy of sensitive information in aggregated data is protected through a local differential privacy-based technique. In the last layer, the mobile sinks are registered with a level-two edge device via a handshaking mechanism to protect the underlying network from external threats. Experimental results show that the proposed framework performs better as compared to existing frameworks in terms of managing the nodes, preserving the privacy of end-devices and sensitive information, and protecting the underlying network.**

***Index Terms*—IoMT, multimedia, multi-level edge computing, aggregation, privacy.**

I. Introduction

INTERNET of Multimedia Things (IoMT) applications, e.g., healthcare, transportation management, and surveillance, generate big multimedia data that are processed using cloud computing platforms [1]. The cloud computing platforms offer computing resources but they cannot deal with end-user side issues. For example, redundancy in the generated data requires sufficient amount of bandwidth from sources to destinations. The processing and storing of redundant data misuse the computing and storage resources on cloud computing platforms. To deal with these challenges, a Multi-Level Edge Computing (MLEC) architecture can be utilized [2]. In this architecture, computationally complex tasks are offloaded to edge devices on different levels to minimize the cost of renting and managing computing and storage resources on cloud computing platforms.

Data generated by IoMT applications need to be processed and transmitted with minimum latency which can be controlled by minimizing the redundancy in the generated data. Furthermore, the privacy of data sources also needs to be protected. In the MLEC architecture, cluster-based communication can help in managing edge and end devices, and operates in two phases, i.e., setup and steady-state phases. During the setup phase, Level-One Edge Devices (LOEDs) are elected, clusters are formed, and Time Division Multiple Access (TDMA) slots are allocated to end-devices. During the steady-state phase, end-devices forward data to a Level-Two Edge Device (LTED) via their corresponding LOEDs. The LOEDs and LTED use different techniques to minimize the redundancy and preserve the privacy of end-devices [3]. These techniques are efficient, however, the LOEDs and LTED may not be powerful enough to execute complex techniques. Therefore, there is a need for a lightweight technique to minimize the redundancy in the generated data and preserve the privacy of data sources.

Data generated by IoMT applications may contain sensitive information that need to be protected in end-to-end communication. Although privacy-preserving techniques help in protecting the privacy of data sources by hiding their identities, they cannot hide visual information present in multimedia data. To protect the privacy of data during transmission, the concept of Local Differential Privacy (LDP) was proposed [4]. This concept is based on the concept of differential privacy, i.e., adding some noise to the data to protect the privacy. However, it cannot directly be applied to multimedia data. Therefore, there is a need for an LDP-based technique that can add noise to multimedia data. Furthermore, it needs to allow users to adjust the privacy parameters and control the amount of noise.

In the MLEC architecture, it is possible that the Internet connectivity between the LTED and cloud servers may temporarily be down due to some technical faults and the LTED may not be able to forward data to the cloud servers. Moreover, the LTED is usually dealing with hundreds of LOEDs, and may not have enough storage space to hold the data for a longer duration. In this situation, mobile sinks can be used to collect data from LOEDs and forward to the cloud servers. However, the credibility of mobile sinks becomes a challenging task [5]. To deal with this challenge, there is a need for a technique to register mobile sinks with the LTED in an offline mode to stop intruders from infiltrating the network and compromising the performance of IoMT applications.

In this paper, we propose a multi-layer framework to provide Privacy protection using Authentication, Aggregation and LDP

A * indicates the corresponding author.

Muhammad Usman is with the School of Science, Engineering and Information Technology, Federation University, Australia. (E-mail: muhammad.usmanskk@gmail.com)

Mian Ahmad Jan is with the Department of Computer Science, Abdul Wali Khan University Mardan, Pakistan (E-mail: mianjan@awkum.edu.pk)

Deepak Puthal is with the School of Computing, Newcastle University, United Kingdom. (E-mail: dputhal88@gmail.com)

(PAAL) for IoMT applications. The proposed framework is based on the MLEC architecture. In this framework, edge devices are used to manage end-devices and provide security and privacy services. The main contributions of this work are summarized below.

- To the best of our knowledge, this is the first framework that provides security and privacy services. The security is provided through a handshaking mechanism while the privacy is achieved through a combination of aggregation and LDP-based techniques.
- An efficient and lightweight four-way handshaking mechanism is used to select LOEDs in the underlying network. A new set of LOEDs is formed in each round of simulation, based on their energy levels. A similar handshaking mechanism is used to register mobile sinks with the LTED. After the registration is completed, the mobile sinks are allowed to collect data from LOEDs.
- A lightweight aggregation technique based on frame matching is proposed to reduce the size of multimedia data. It also helps in protecting the privacy of end-devices by hiding their identities and location coordinates information. The privacy of visual contents is protected through an efficient LDP-based technique. This technique preserves the privacy by dynamically adding the noise to the aggregated data, based on the computed privacy budget.

The rest of this paper is structured as follows. An overview of recent literature is provided in Section II. The proposed framework is explained in Section III. Experimental setup and simulation results are discussed in Section IV. Finally, the paper is concluded in Section V.

II. Literature Review

In this section, we provide an overview of recent efforts made in the edge computing, data aggregation, and differential privacy domains.

A. Edge Computing

A survey on mobile edge computing was presented in [6]. This survey highlights and discusses various research challenges for delay-sensitive applications operating in the edge computing environment. However, a discussion on multimedia data management and processing is missing in this survey. An edge computing framework for multimedia processing was proposed in [7]. This framework is basically designed to support cooperative processing of video data in the Internet of things architecture. A quality of experience driven framework for mobile edge computing was proposed in [8]. This framework is designed to support dynamic and adaptive video streaming to minimize the cost of downloading from edge devices. Similarly, a caching framework for edge computing architecture was proposed in [9]. This framework is designed to support video streaming demands of mobile users. The frameworks presented in [7]–[9] support the processing of video streams in the edge computing environments. However, these frameworks lack support for real-time processing of video data.

B. Data Aggregation

A multi-functional aggregation scheme for Wireless Sensor Networks (WSNs) was proposed in [10]. This scheme uses homomorphic encryption to secure data in end-to-end communication. An aggregation scheme based on access control and node authentication for WSNs was proposed in [11]. This scheme targets sinkhole and Sybil attacks during end-to-end communication. A compressive aggregation scheme for WSNs was proposed in [12]. This scheme is based on compressed sensing and helps in reducing the energy consumption during the aggregation process. An LOEDs-based aggregation algorithm to efficiently utilize the available bandwidth in WSNs was proposed in [13]. This algorithm uses mobile sinks to perform the aggregation task and supports intra and inter-cluster aggregations. The schemes presented in [10]–[13] are designed to efficiently perform the aggregation task. However, these schemes are basically designed to deal with non-multimedia data, and may not be feasible for IoMT applications.

C. Differential Privacy

A framework to provide differential privacy during online distributed learning was proposed in [14]. In this framework, the video-selection strategy of end-users is used to control the performance loss and protect the privacy. A framework to protect graphical data on cloud platforms was proposed in [15]. This framework adds noise patterns to data to protect the privacy of data sources. A framework to derive the amount of required noise was proposed in [16]. This framework uses Gaussian distribution to select the right amount of noise. A framework for privacy protection in a crowdsensing scenario was proposed in [17]. This framework analyzes the privacy risks and makes recommendations. A framework to protect the privacy of individuals was proposed in [18]. This framework uses an anonymity model and allows participating individuals to decide the privacy constraints. The frameworks presented in [14]–[18] help in preserving the privacy of data sources. However, these frameworks are basically designed for cloud-based systems, and may not be feasible for IoMT applications.

III. Privacy Protection using Authentication, Aggregation, and Local Differential Privacy

In this section, we explain our proposed PAAL framework. It is a multi-layer framework based on the MLEC architecture and provides security and privacy services for IoMT applications. This framework operates in three layers. In the first layer, clusters are formed in the underlying Wireless Multimedia Sensor Network (WMSN) as shown in Fig. 1. Each cluster is represented by an LOED. The LOEDs are elected by an LTED using a four-way handshaking mechanism. The LOEDs are responsible to manage end-devices and collect data from them. In the second layer, the LOEDs apply a lightweight aggregation technique to reduce the size of generated data and protect the privacy of end-devices. To protect the privacy of visual contents during transmission, an efficient LDP-based technique is applied to the data. In the final layer, the mobile sinks are registered with the LTED using a similar handshaking mechanism used to to elect LOEDs. After the registration is completed, the mobile sinks are allowed to collect data from LOEDs. In the following subsections, we explain these layers in detail.

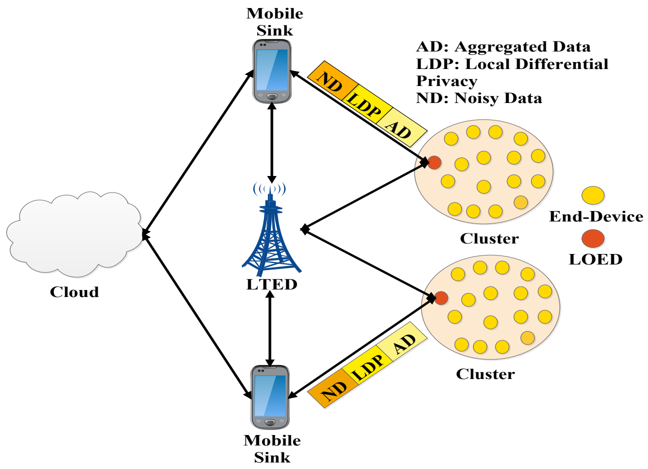


Fig. 1: PAAL Framework

A. Node Management Layer

This layer is based on our previous work published in [19]. Symbols used in this layer are summarized in Table I.

Notation	Description
E	Average Energy
e	Current Energy Level
N	Total Number of End-Devices
d	Shortest Distance
(x, y)	Coordinates of LOEDs
(x_n, y_n)	Coordinates of End-Devices

TABLE I: Symbols of Node Management Layer

In this layer, the underlying WMSN is partitioned into small clusters. Each cluster is represented by a head node known as an LOED. The elected LOEDs are responsible to collect data from member devices and coordinate with a nearby LTED where the LTED controls the entire network. It also performs computationally complex tasks and communicates with cloud servers. The selection of LOEDs is based on a four-way handshaking process, i.e., LOEDs advertisement, join-request, LOEDs acknowledge, and neighboring devices' responses. In the beginning, all end-devices share their Identities (IDs), energy levels, and location coordinates information with the LTED to register themselves and compete for the roles of LOEDs. After receiving the information, the LTED performs a comparison based on the received energy levels, and computes the average energy threshold, i.e., E, using the following equation.

$$E = \sum_{n=1}^{N} \frac{e_n}{N}, \tag{1}$$

where e and N represent the current energy level of an end-device and the total number of end-devices in the WMSN, respectively.

The LTED compares the energies of all end-devices with the computed threshold. End-devices with energies equal to or greater than the average energy threshold are shortlisted for the roles of LOEDs. During the comparison, it is possible that more than one end-device may have the same energy level. In this situation, the selection is based on very minor differences in the energy values of end-devices and can go up to four decimal places. Moreover, there may be a situation when the energy levels may exactly be the same. In this situation, end-devices selected as LOEDs in the past rounds of simulations are excluded.

Once the LOEDs are shortlisted, the LTED broadcasts a message containing the IDs and location coordinates information of shortlisted LOEDs. Upon receiving the message, the nominated LOEDs perform the following operations.

1) Retrieving the IDs and location coordinates information of their neighboring end-devices.
2) Storing the retrieved information in their buffers.
3) Sending an acknowledgment message back to the LTED to confirm the successful arrival of forwarded information.
4) Advertising themselves in their neighborhood by sharing their location coordinates information.

It is possible that end-devices may receive invitations from multiple LOEDs simultaneously. In this situation, the end-devices shortlist LOEDs by comparing their current energy levels. Once the LOEDs are shortlisted, the end-devices calculate the distance to these LOEDs using the following Euclidean distance formula.

$$d = \sqrt{(x - x_n)^2 - (y - y_n)^2}, \tag{2}$$

where d represents the shortest distance, and (x, y) and (x_n, y_n) are the coordinates of prospective LOEDs and neighboring end-devices, respectively.

It is possible that multiple LOEDs may exactly be at the same distance from an end-device. In this situation, a join-request is sent to an LOED with the strongest signal strength. In the case of a tie, the end-device randomly selects an LOED and associates itself with it. Alternatively, the end-device may send the join-request to all LOEDs and associates itself with the one who replies first. The latter case is beneficial as it is based on the availability of the TDMA slot with the particular LOED.

B. Privacy Protection Layer

Traditionally, the LOEDs collect data from end-devices and forward to the LTED for further processing. Later, the LTED is responsible to forward the processed data to the cloud servers. In our proposed framework, we assume that the Internet connectivity between the LTED and the cloud servers is temporarily down due to some technical fault and the LTED is unable to forward the data to the cloud servers. In this situation, there is a need for external entities, e.g., mobile sinks, to collect data from LOEDs and upload data to the cloud servers. However, the generated data need to be processed first before sharing with mobile sinks. The privacy protection layer is subdivided into two layers, i.e., data aggregation and noise addition layers. In the data aggregation layer, the LOEDs aggregate collected data to minimize the redundancy and protect the privacy of end-devices. In the noise addition layer, some noise is added to the aggregated data to protect the privacy of visual contents. In the following subsections, the data aggregation and noise addition layers are explained in detail. Symbols used in this layer are summarized in Table II.

Notation	Description
S	Similarity Index
H_s, H_r	Histograms
L_1, L_2	Levels
r_1, r_2	Sizes of Regions
ε	Privacy Budget
Λ, Q_1, Q_2	Constants
P	Pixels
λ	Noise Count
θ	Two-Dimensional Standard Deviation
(x_p, y_p)	Coordinates of a Pixel
$\hat{x}, \hat{y}$	Mean Value of Coordinates
ϑ	Privacy Need

TABLE II: Symbols of Privacy Protection Layer

1) Data Aggregation Layer: In a cluster, the end-devices usually focus on the same event but at different angles. As a result, data coming from a cluster may contain similar information captured at different angular positions and cause redundancy. In this situation, the aggregation techniques can be used to reduce the size of generated data. Furthermore, the aggregation techniques also help in protecting the privacy of data sources.

In this layer, we use a similarity-based aggregation technique to minimize the redundancy in the captured data. In this technique, video data coming from a cluster are distributed into two sets, i.e., standard and redundant sets. Videos captured by LOEDs are placed in the standard set while the videos forwarded by end-devices are placed in the redundant set. For a comparison, videos from both sets are distributed into multiple groups of frames, and are compared on a frame-by-frame basis as shown in Fig. 2.

In this comparison, a video frame is divided into multiple blocks of equal size, i.e., 64×64. Blocks of a video frame from redundant set are compared with the blocks of the corresponding video frame from standard set using a histogram normalization technique to compute the similarity index (i.e., S). This comparison is performed using the following equation.

$$S = H_s \times \log_2 \left(\frac{H_s}{H_r} \right), \tag{3}$$

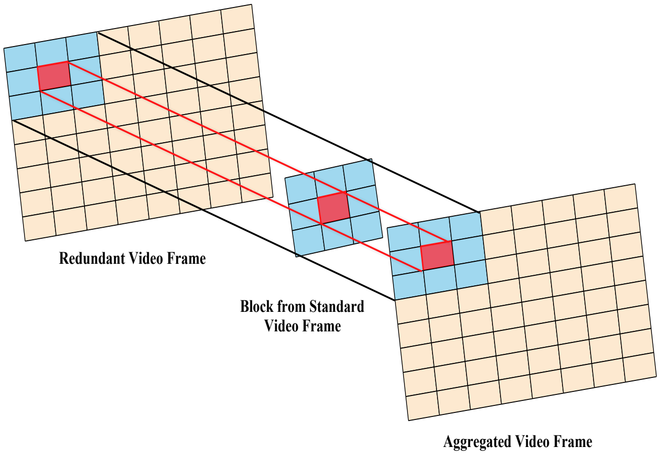


Fig. 2: Data Aggregation

where H_s and H_r represent the histograms of standard and redundant video frames, respectively.

The computed similarity index values of all blocks from redundant video frames are compared against a predefined threshold (i.e., $\bar{S}$) where $\overline{S} \in (S_{min}, S_{max})$. If the computed similarity index values are less than the predefined threshold, then the blocks are considered dissimilar. A video frame from redundant set is considered different if it contains at least 25% dissimilar blocks. The LOEDs store all dissimilar frames and only one sample video from standard set and discard the remaining videos from redundant set. The dissimilar frames are combined together to create aggregated videos. Furthermore, the location coordinates information of end-devices in a cluster is also aggregated by computing the mean value. The aggregated location coordinates information is associated with the aggregated videos to complete the aggregation process. The aggregation of data and location coordinates information help in reducing the size of data and preserving the privacy of end-devices.

2) Noise Addition Layer: Adding noise to videos is not only a time-consuming process, but may also corrupt the videos. Furthermore, in a video frame, there may be multiple background or flat regions. These regions usually do not contain any sensitive information and adding noise to these regions misuses the available privacy budget. To address these challenges, video frames need to be partitioned into multiple levels to add a suitable amount of noise to different regions. Each level can further be subdivided into multiple non-overlapping regions to meet the privacy needs. Our noise addition technique is based on the standard deviation theorem. This technique dynamically adds noise to different regions of a video frame. The noise addition technique helps in understanding the privacy-preserving need based on the divergence of data points in a specific region of a video frame. Furthermore, the noise addition technique satisfies the ε-differential privacy and operates in two phases, i.e., distribution of video frames and addition of noise, as shown in Fig. 3. In the following subsections, we explain these two phases in detail.

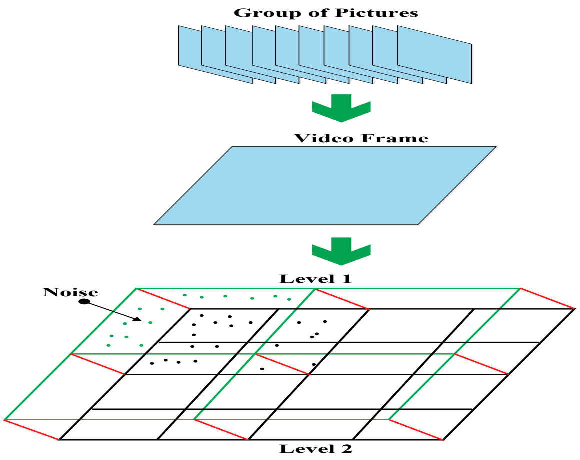


Fig. 3: Multi-Level Noise Addition

a) Distribution of Frames: In this phase, aggregated videos are distributed into multiple Groups of Pictures (GoP). In our proposed framework, each GoP consists of 10 video frames. A video frame from a GoP is divided into two levels, i.e., L_1 and L_2. Each level is further distributed into a grid of non-overlapping square-shaped regions. Each region at level L_1 is of size $r_1 \times r_1$. Here, the value of r_1 is computed using the following equation.

$$\varepsilon_1 = \varepsilon \times \Lambda, \tag{4a}$$

$$r_1 = max\left[10, \left(\frac{1}{4}\sqrt{\frac{P \times \varepsilon_1}{Q_1}}\right)\right], \tag{4b}$$

where ε represents the total privacy budget and Λ is a computing constant where $0 < \Lambda < 1$. It helps in distributing the privacy budget between two levels. Symbol P represents the total number of pixels in a region, symbol ε_1 is the total privacy budget set for regions at L_1, and symbol Q_1 is the experimental constant for L_1 and is set to 10 as recommended by [20].

Each region at level L_1 is subdivided into smaller regions of size $r_2 \times r_2$ to make a grid at level L_2. The value of r_2 is computed using the following equation.

$$\varepsilon_2 = \varepsilon - \varepsilon_1, \tag{5a}$$

$$r_2 = \sqrt{\frac{\lambda \times \varepsilon_2}{Q_2}}, \tag{5b}$$

where λ is noise count in the corresponding region at level L_1 and Q_2 is the experimental constant for level L_2 and is set to $\sqrt{2}$ as recommended by [20].

b) Addition of Noise: Traditionally, the same amount of noise is added to whole data based on the allocated privacy budget. However, this is not feasible for video data due to the presence of flat regions. These regions do not contain any sensitive information and addition of noise to these regions not only misuses the available privacy budget but also increases the computational cost. To deal with this challenge, our proposed framework dynamically adds noise to regions on both levels based on the privacy-preserving requirements. The noise addition is a four-step process. In the first step, a two-dimensional standard deviation is computed for each level to find the convergence of data points with respect to the mean central value. The two-dimensional standard deviations of individual regions (i.e., θ_i) and all regions (i.e., θ) on a certain level are computed using the following equation.

$$\begin{aligned} \theta_i &= \sqrt{\sum_{p=1}^{P} \frac{(x_p - \hat{x})^2 + (y_p - \hat{y})^2}{P-2}}, \\ \theta &= \sum_{i=1}^{I} \theta_i, \end{aligned} \tag{6}$$

where (x_p, y_p) represents the coordinates of a pixel p and $(\hat{x}, \hat{y})$ represent the mean value of coordinates of a particular region.

In the second step, privacy-preserving need of a particular region (i.e., ϑ_i) on a certain level is computed by the following equation.

$$\vartheta_i = \frac{\theta_i}{\theta}. \tag{7}$$

In the third step, privacy budget of a specific region (i.e., ε_i) on a certain level is computed by the following equation.

$$\varepsilon_i = \vartheta_i \times \varepsilon_j, \tag{8}$$

where ε_j represents the total privacy budget of a certain level j, i.e., L_1 or L_2.

In the last step, the Laplace noise is added to individual regions after computing the privacy-preserving need and budget parameters.

C. Sink Registration Layer

In our proposed framework, we use mobile sinks to upload data to cloud servers. However, the mobile sinks cannot be fully trusted and need to be registered with the LTED prior to data collection. In our proposed framework, the registration process is based on a mutual handshaking mechanism. The registration process with the LTED is accomplished in four steps as shown in Fig. 4. Symbols used in this layer are summarized in Table III.

Notation	Description
M	Mobile Sinks
η	Session Key
a, b, c	Random Integers
$\mathcal{F}$	Mapping Function
R	Registration Request
C	Challenge
δ	Time-stamp
Δ	Authentication Value
$\overline{R}$	Response to Challenge

TABLE III: Symbols of Sink Registration Layer

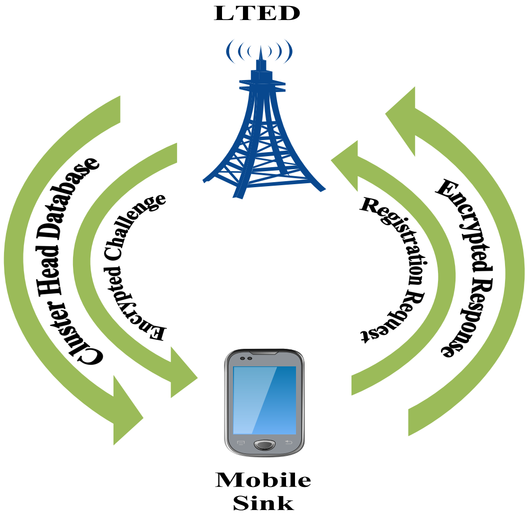


Fig. 4: Mobile Sink Registration

Before sending the registration request, a mobile sink (i.e., m where $m = \{1, 2, \cdots, M\}$) generates a session key (i.e., μ_m) and two random integers (i.e., a and b, where $a \in Z$ and $b \in Z$). Later, an encrypted registration request (i.e., R_m) is generated by using the following equation and forwarded to the LTED.

$$R_m = AES\{m, (\mathcal{F}(a \oplus \mu_m))\}, \tag{9}$$

where $\mathcal{F}$ is a mapping function used to perform a one-way secure hashing with a value within the range $[0 \quad 1]$ and $\oplus$ is an XOR operator [21]. For encryption, we use AES-256 bits.

After receiving the encrypted registration request, the LTED retrieves the embedded values, creates an encrypted challenge (i.e., C) by using the following equation, and forwards it back to the mobile sink.

$$\delta_m = \mathcal{F}(m \mid\mid c), \tag{10a}$$

$$\Delta_m = \mathcal{F}(m \mid\mid a \mid\mid c), \tag{10b}$$

$$C = AES\{\delta_m, \Delta_m\}, \tag{10c}$$

where δ_m represents a time-stamp, Δ_m represents an authentication value, and c is a random integer where $c \in Z$. The allocated time-stamp allows the mobile sink to collect data from LOEDs in the underlying network for a specific period of time and the authentication value is used to create a response to the challenge.

After receiving the encrypted challenge, the mobile sink retrieves the embedded values and creates an encrypted response (i.e., $\overline{R_m}$) by using the following equation.

$$O_m = \mathcal{F}(\Delta_m \mid\mid \mathcal{F}(a \oplus \mu_m) \mid\mid b), \tag{11a}$$

$$\overline{R_m} = AES\{O_m, b\}. \tag{11b}$$

After receiving the encrypted response, the LTED verifies the identity of the mobile sink using Eq. 11a. Once the entire verification process is completed, the status of mobile sink changes to registered. The LTED broadcasts the generated authentication value to all LOEDs and shares the information of nominated LOEDs with registered mobile sink.

IV. Experimental Setup

In this section, we evaluate the performance of our proposed framework by comparing it with existing privacy-preserving frameworks, i.e., SLICER with Transfer on Meet Up (SLICER-TMU), SLICER with Minimum Cost Transfer (SLICER-MCT), and Simple Exchanging (SE), [22], [23]. The SLICER-TMU and SLICER-MCT are k anonymity-based frameworks to preserve the privacy of participating entities. These frameworks integrate data coding techniques and message transfer strategies to preserve the privacy. Similar to our proposed framework, the SLICER-TMU and SLICER-MCT frameworks also use multiple algorithms to achieve the privacy-preserving goal. The SE framework, on the other hand, uses a decentralized mechanism to preserve the privacy of location coordinates information. Due to the geo-tagging feature, the sensed data are exchange between users in order to jumble the paths followed by the users. The SLICER-TMU, SLICER-MCT, and SE frameworks are designed to provide privacy-preserving services, however, they have certain limitations. Firstly, they are designed to preserve the privacy of location coordinates information of data sources. Secondly, their scope is limited and they cannot deal with large scale networks. Thirdly, they lack a support for privacy-preserving in end-to-end communication. On the other hand, our proposed framework covers all these limitations as shown in simulation results in this section. To compare the performances, we consider three different metrics, i.e., computational overhead, data load, and relative error. For simulations, we build a WMSN consisting of 500 randomly distributed Multimedia Sensor Nodes (MSNs). In each round of simulation, only 5% MSNs are selected as LOEDs. The simulations are performed in Matlab 2018a. The mobile sinks are always on the move with a constant speed and their movement is based on a random way-point mobility model [24].

In the traditional scenario, the LOEDs are responsible for three main tasks, i.e., manage member nodes, collect data from them, and coordinate with a nearby LTED. In our proposed framework, we introduce two extra tasks for LOEDs, i.e., aggregation of multimedia data and location coordinates information, and addition of noise to the aggregated data. Unlike the aggregation of location coordinates information, the aggregation of multimedia data and the addition of noise are computationally complex tasks. Videos are considered as a source of data in our proposed framework, therefore, the computational efficiency and extra storage space at LOEDs are the basic requirements. The computational overhead metric is used to approximate the total amount of time required to aggregate the video data and add noise along with performing other tasks. On average, our proposed framework shows better performance as compared to the SLICER-TMU and SLICER-MCT frameworks as shown in Fig. 5. The SE framework adds noise to the raw data directly and does not aggregate the data.

As a result, less computational overhead is observed in the case of SE framework. Initially, all frameworks show similar performances. However, the computational overhead increases in the case of SLICER-TMU and SLICER-MCT frameworks with an increase in the total number of transmitting MSNs. Furthermore, the targeted frameworks do not include a mechanism to manage and verify the identities of member MSNs. As a result, malicious nodes can easily flood the underlying WMSN and overload the LOEDs by injecting false data. Node management technique in our proposed framework not only helps in reducing the network traffic, but also helps in minimizing the computational load on LOEDs.

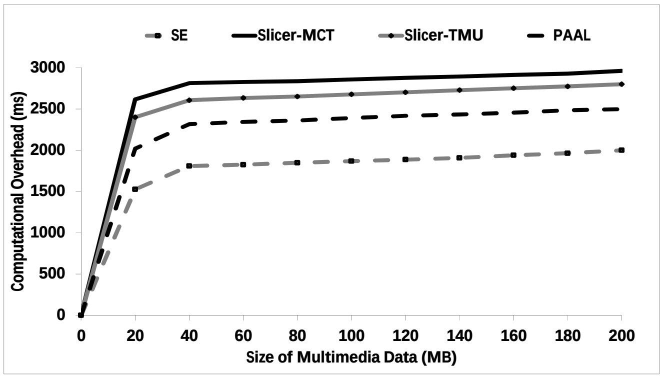


Fig. 5: Computational Overhead

The next comparison is based on data load. The data load metric represents the total amount of data transmitted from LOEDs to the mobile sinks. As shown in Fig. 6, our proposed framework transmits less amount of data as compared to the existing privacy-preserving frameworks.

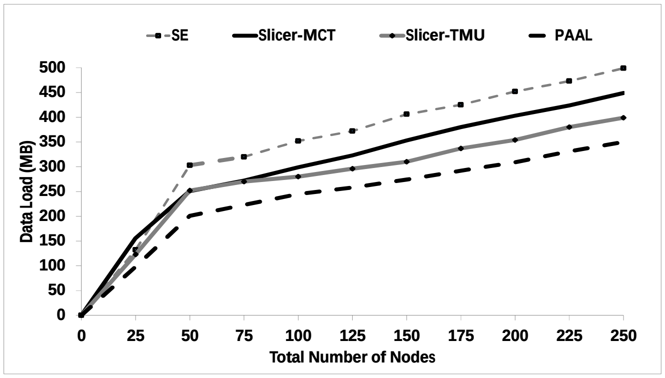


Fig. 6: Data Load

Our proposed framework uses an efficient aggregation technique to minimize the redundancy in the collected data. Without applying an efficient aggregation technique, large size video data need to be transmitted from sources to destinations which ultimately requires the availability of large amount of bandwidth. Furthermore, the video data captured in specific applications, e.g., surveillance and transportation management, usually contain repetitive information. As a result, the processing of redundant/repetitive data requires extra computational and storage resources which is not feasible for resource-constrained devices in the IoMT architecture.

The relative error metric represents a ratio between absolute error and the original data. The absolute error is computed by subtracting the data with noise from original data. A small amount of relative error means the noise amount is smaller in the generated results. As a result, the freshness of data, its accuracy, and privacy-preserving effects are higher. Fig. 7-8 show a comparison based on relative error. In Fig. 7, the block size is kept fixed with constant and variable amounts of noises while in Fig. 8, the amount of noise is kept constant with fixed and variable block sizes. Fig. 7(a) shows a comparison of relative error when the amount of noise is constant in all regions of a video frame while Fig. 7(b) shows a comparison of relative error when the amount of noise changes from one region to another in a video frame. In either case, our proposed framework shows a better performance. However, an improved performance is observed in the case of a variable amount of noise. Similarly, the block sizes are kept fixed and variable with constant amount of noise in Fig. 8(a) and Fig. 8(b), respectively. Again, the overall performance of our proposed framework is better as compared to the existing privacy-preserving frameworks. In both Fig. 7-8, it can clearly be seen that our proposed framework performs better, however, an improved performance is observed when block size and the amount of noise are variable. A lower ratio of relative error can assure the accuracy and the freshness of data without compromising the privacy factor.

V. Conclusion

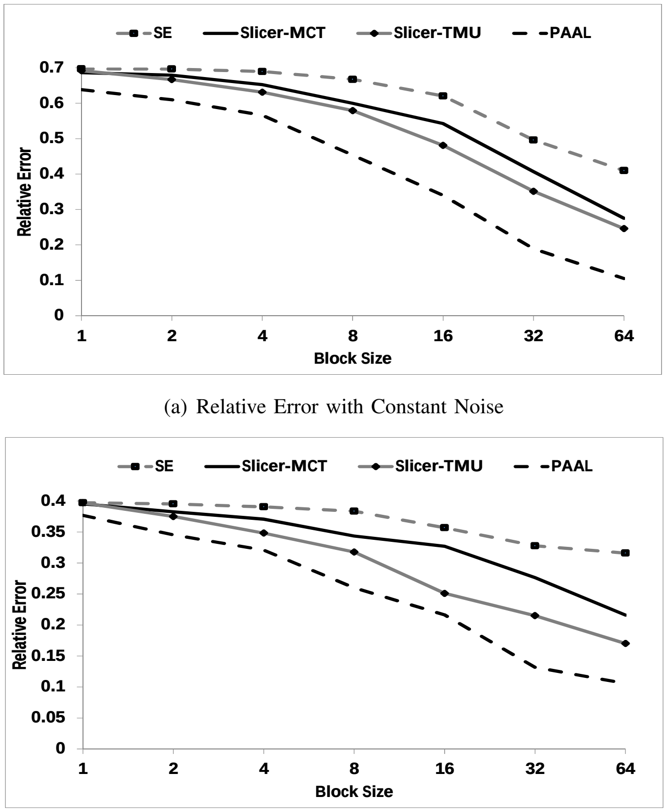
In this paper, we have proposed a multi-layer privacy-preserving framework based on the MLEC architecture for IoMT applications. This framework operates in three layers. In the first layer, the underlying network is partitioned into multiple clusters, each represented by an LOED. These LOEDs are responsible to manage member nodes and collect data from them. In the second layer, the LOEDs aggregate the collected data and location coordinates information to preserve the privacy of data sources. They add noise to the aggregated data to preserve the privacy of visual contents in end-to-end communication. In the last layer, prior to data collection from LOEDs, the mobile sinks are registered with the LTED using a mutual handshaking mechanism. In experiments, our proposed framework has shown a better performance as compared to the existing privacy-preserving frameworks. In future, we aim to further improve the performance of our proposed framework by incorporating the use of machine learning algorithms to speed up data processing and tune the performance of privacy-preserving algorithms.

References

[1] S. A. Alvi, B. Afzal, G. A. Shah, L. Atzori, and W. Mahmood, “Internet of multimedia things: Vision and challenges,” *Ad Hoc Networks*, vol. 33, pp. 87–111, 2015.

[2] M. Satyanarayanan, “The emergence of edge computing,” *Computer*, vol. 50, no. 1, pp. 30–39, 2017.

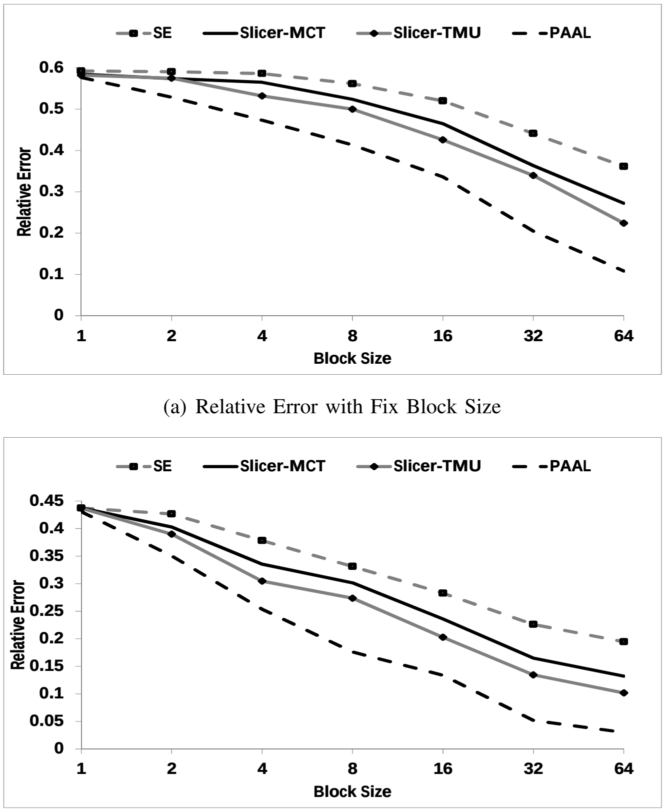
[3] P. Jesus, C. Baquero, and P. S. Almeida, “A survey of distributed data aggregation algorithms,” *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 381–404, 2015.



(a) Relative Error with Constant Noise

(b) Relative Error with Variable Noise

Fig. 7: Relative Error with Constant and Variable Noises



(a) Relative Error with Fix Block Size

(b) Relative Error with Variable Block Size

Fig. 8: Relative Error with Constant and Variable Block Sizes

[4] P. Kairouz, S. Oh, and P. Viswanath, “Extremal mechanisms for local differential privacy,” in *Advances in neural information processing systems*, 2014, pp. 2879–2887.

[5] M. Usman, M. A. Jan, X. He, and J. Chen, “A mobile multimedia data collection scheme for secured wireless multimedia sensor networks,” *IEEE Transactions on Network Science and Engineering*, 2018.

[6] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, “A survey on mobile edge computing: The communication perspective,” *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2322–2358, 2017.

[7] C. Long, Y. Cao, T. Jiang, and Q. Zhang, “Edge computing framework for cooperative video processing in multimedia iot systems,” *IEEE Transactions on Multimedia*, vol. 20, no. 5, pp. 1126–1139, 2018.

[8] C. Li, L. Toni, J. Zou, H. Xiong, and P. Frossard, “Qoe-driven mobile edge caching placement for adaptive video streaming,” *IEEE Transactions in Multimedia*, vol. 20, no. ARTICLE, pp. 965–984, 2018.

[9] Z. Su, Q. Xu, F. Hou, Q. Yang, and Q. Qi, “Edge caching for layered video contents in mobile social networks,” *IEEE Transactions on Multimedia*, vol. 19, no. 10, pp. 2210–2221, 2017.

[10] P. Zhang, J. Wang, K. Guo, F. Wu, and G. Min, “Multi-functional secure data aggregation schemes for wsns,” *Ad Hoc Networks*, vol. 69, pp. 86–99, 2018.

[11] A. Razaque and S. S. Rizvi, “Secure data aggregation using access control and authentication for wireless sensor networks,” *Computers & Security*, vol. 70, pp. 532–545, 2017.

[12] S. Abbasi-Daresari and J. Abouei, “Toward cluster-based weighted compressive data aggregation in wireless sensor networks,” *Ad Hoc Networks*, vol. 36, pp. 368–385, 2016.

[13] D. S. Mantri, N. R. Prasad, and R. Prasad, “Bandwidth efficient cluster-based data aggregation for wireless sensor network,” *Computers & Electrical Engineering*, vol. 41, pp. 256–264, 2015.

[14] P. Zhou, Y. Zhou, D. Wu, and H. Jin, “Differentially private online learning for cloud-based video recommendation with multimedia big data in social networks,” *IEEE transactions on multimedia*, vol. 18, no. 6, pp. 1217–1229, 2016.

[15] Z. Qin, K. Ren, T. Yu, and J. Weng, “Dpcode: privacy-preserving frequent visual patterns publication on cloud,” *IEEE Transactions on Multimedia*, vol. 18, no. 5, pp. 929–939, 2016.

[16] H. Zhang, N. Yu, Y. Wen, and W. Zhang, “Towards optimal noise distribution for privacy preserving in data aggregation,” *computers & security*, vol. 45, pp. 210–230, 2014.

[17] R. Liu, J. Liang, W. Gao, and R. Yu, “Privacy-based recommendation mechanism in mobile participatory sensing systems using crowdsourced users’ preferences,” *Future Generation Computer Systems*, vol. 80, pp. 76–88, 2018.

[18] J. Le, X. Liao, and B. Yang, “Full autonomy: A novel individualized anonymity model for privacy preserving,” *Computers & Security*, vol. 66, pp. 204–217, 2017.

[19] M. Usman, N. Yang, M. A. Jan, X. He, M. Xu, and K.-M. Lam, “A joint framework for qos and qoe for video transmission over wireless multimedia sensor networks,” *IEEE Transactions on Mobile Computing*, vol. 17, no. 4, pp. 746–759, 2018.

[20] W. Qardaji, W. Yang, and N. Li, “Differentially private grids for geospatial data,” in *Data Engineering (ICDE), 2013 IEEE 29th International Conference on*. IEEE, 2013, pp. 757–768.

[21] F. Zhang, R. Safavi-Naini, and W. Susilo, “An efficient signature scheme from bilinear pairings and its applications,” in *International Workshop on Public Key Cryptography*. Springer, 2004, pp. 277–290.

[22] F. Qiu, F. Wu, and G. Chen, “Privacy and quality preserving multimedia data aggregation for participatory sensing systems,” *IEEE Transactions on Mobile Computing*, vol. 14, no. 6, pp. 1287–1300, 2015.

[23] D. Christin, J. Guillemet, A. Reinhardt, M. Hollick, and S. S. Kanhere, “Privacy-preserving collaborative path hiding for participatory sensing applications,” in *Mobile Adhoc and Sensor Systems (MASS), 2011 IEEE 8th International Conference on*. IEEE, 2011, pp. 341–350.

[24] C. Bettstetter, G. Resta, and P. Santi, “The node distribution of the random waypoint mobility model for wireless ad hoc networks,” *IEEE Transactions on mobile computing*, vol. 2, no. 3, pp. 257–269, 2003.