Dual Scheme Privacy-Preserving Approach for Location-Aware Application in Edge Computing

Bruce Gu[1(✉)], Youyang Qu[2], Khandakar Ahmed[1], Wenjie Ye[1], Chenchen Tan[2], and Yuan Miao[1]

[1] Intelligent Technology Innovation Lab, Victoria University, Footscray, Australia
{bruce.gu,khandakar.ahmed,wenjie.ye,yuan.miao}@vu.edu.au
[2] Deakin Blockchain Innovation Lab, Deakin University, Burwood, Australia
{y.qu,tanchen}@deakin.edu.au

Abstract. The location awareness capabilities of edge computing (EC) contains large quantity of the physical devices with short coverage range. The possibilities of the potential private data attacks from adversaries increases dramatically through easily accessible location information. The existing research on privacy-preserving schemes cannot meet various privacy-preserving expectations in practice for EC variants. In this paper, we proposed a dual scheme customizable ϵ-differential privacy preservation to provide comprehensive protection. We establish the first scheme by clustering Edge Nodes (ENs) with SDN-enabled EC where SDN enables the capabilities of the programmability. In addition, we customize the ϵ-differential privacy preservation scheme for variant EC services with the employment of modified Laplacian mechanism to generate noise, where the optimal tradeoff been found. The extensive experiments results demonstrate the significance of the proposed model in terms of privacy protection level and data utility, respectively.

Keywords: Edge computing · Privacy-preserving · Software defined network · Differential privacy · Location-aware application

1 Introduction

The Internet of Things (IoTs) is advancing at a breakneck pace, and connection between things is becoming more pervasive [3,7]. Increasingly more objects are becoming intelligent as they are capable of seeing their surroundings, connecting to the internet, and receiving instructions remotely. These intelligence of the objects are derived from data, analysis, and input from a variety of systems or servers connected to a variety of mobile devices [4].

The wireless and decentralized properties of ENs are vulnerable to adversaries in the actual application process of edge computing (EC), resulting in

Supported by Intelligent Technology Innovation Lab (ITIL), Victoria University.

W. Bao et al. (Eds.): ADHOCNETS 2021/TridentCom 2021, LNICST 428, pp. 301–316, 2022.
https://doi.org/10.1007/978-3-030-98005-4_22

major location-aware privacy disclosure vulnerabilities [14]. EC enables computation to take place closer to the user [17]. It evaluates data locally and makes decisions based on it. It decreases the danger of privacy leakage by avoiding long-distance data transfer in the network to certain extent [27]. However, a substantial amount of sensitive private data can be retrieved since attackers can easily access real-time data received by ENs. Higher criteria for privacy protection techniques in EC have been recommended as a result of methodologies that assure users can utilise the service without revealing their sensitive location data [19]. Traditional privacy-preserving solutions are impractical for directly addressing the highlighted problem in an effective manner, despite certain existing privacy-preserving models or algorithms being presented [9]. Furthermore, the location privacy disclosure concern identified in EC will have a significant impact on EC application development. It is critical to investigate location-aware privacy preserving solutions for massive EC applications based on location.

Motivated by this, to establish an optimal trade-off on data utilities with a sufficient accuracy and efficiency, we offer a dual-scheme ϵ-customised differential privacy model (DDSDP) based on software-defined EC (SD-EC) services. SD-EC increases the network's privacy protection capabilities by allowing it to be programmed flexibly and dynamically [11,18]. First, we initiate the EN clustering method. This strategy connects users to EC services through a group of ENs rather than a single reliable service provider, while also increasing the complexity of attacking targets for adversaries. Furthermore, the measurements of ϵ-customised differential privacy was determined by the distance between ENs cluster. To quantify data utilities and privacy protection levels, we create a QoS-based mapping function. Our thorough studies illustrate the efficiency and accuracy in real time.

The contributions of this paper can be summarized as follows:

- We offer a dual-scheme ϵ-customized differential privacy model to retain location-aware private data for users. In addition, to offer first scheme protection for location-aware data from users to ENs, we consider a dynamic clustering strategy. This protects against frontal attacks from adaptable foes. Furthermore, We adapt Affinity Propagation methods to boost the effectiveness of the clustering scheme while retaining the greatest degree of privacy protection.
- We designate our second approach as the ϵ-customized protection model with modified Laplacian mechanism by providing noise to the clustered EN. Moreover, we developed a QoS-based mechanism to measure the distance between privacy levels. DDSDP seeks to enhance the utilities of data while minimising privacy costs.
- To illustrate the proposed models, we perform extensive experiments based on real-world dataset. The evaluation results reveal notable significant performances in data utilities and privacy protection level.

This paper is organized as following. Section 2 summarizes the existing research on EC and related privacy protection approaches. The attacking formulation with two popular attack methods is analyzed in Sect. 3. Section 4

formulates our proposed DDSDP model in both schemes. Section 5 performs evaluation results from our extensive experiments by using the DDSDP protection model. Finally, summarization and future works are discussed in Sect. 6.

2 Related Work

EC benefits from dense geographical dispersion, which accomplished by installing ENs in multiple locations and interconnecting each of the nodes to end devices [22]. Moreover, the geographic dispersion of the ENs enables location mobility for IoTs devices, obviating the requirement for devices to traverse the entire network [10]. Besides the tangible benefits of EC's location-aware capabilities, the most immediate challenge is the preservation of users' privacy when it relates to geographical location data.

The user in the EC environment are not expected to reveal actual location to attackers when obtaining services. Location-aware data includes not only the current specific location, but also the contents of the EN that are saved and processed in EC, such as movement patterns and behavioural patterns. Indeed, the dimensionality to which location-aware privacy protection and Quality of Service(QoS) are the result of a sequence of antagonistic connections [8]. Current research on location-aware privacy protection technology for EC is primarily concentrated on two aspects: 1) the the models of privacy-preserving that utilises reputable third-party entities [24], and 2) the data anonymization technique [16].

Moreover, J. Kang [13] proposed privacy-preserving pseudonym method which addressed privacy concerns associated with location-based EC internet vehicles. While these solutions provide acceptable performance, they are more concerned with maintaining a stable network state than with dynamic and tailored EC restrictions. Lyu [15] performed extensive research on the tailored *epsilon*-differential privacy-preserving technique [12], which Qu [20], Badsha [2], and Wang Wang [25] all effectively demonstrated. These strategies are very successful in social networks, recommendation systems, and location-aware apps. They are theoretically sound and include strong privacy safeguards [23].

Furthermore, we examine two typical attack vectors that adversaries are commonly using to target sensitive location data in the EC.

Wormhole Attacks. This type of attacks are very hazardous since they may occur even when all parties to the communication guarantee the message's validity and secrecy. The attacker creates a secret channel between two cooperating malicious nodes with the intent of transferring data groups acquired at one network point to another. J. Zhang et al. [28] developed the grid clustering routing method (FGC) to protect against wormhole attacks, and it is currently extensively used in industrial IoT.

DDoS Attack. Another popular attack technique in SDN-enabled EC is the DDoS assault, which targets the communication layer. According to the location service provider, it prevents radio signals from being transmitted [21,26].

The literature has examined the feasibility and efficacy of DDoS attacks against a variety of transport protocols, including Bluetooth [5]. Along with active interfering attacks, the adversaries may conduct a denial-of-service (DoS) attack by installing a malicious device or router that intentionally violates the communication protocol in order to cause conflicts or disrupt communication.

3 Formulation of Adversaries Attack in Location-Aware

The volume of information that attackers may extract from the location-aware privacy data for the users after its broadcasting influences the effectiveness of the adversary attack model in the EC. The location data captured by the adversaries from user encapsulated within the context of a collection of spatiotemporal data $\{r : r = (P, t)\}$, where r determined a single adversary-collected location data, and P defined the specific location of the user, and the accurate timing of the collection determined as t. Moreover, the disclosure of sensitive location data is structured as a set of $\{s_1, s_2, ..., s_n\}$, where s_n denotes the nth position after numbering.

Adversaries can deduce users' privacy information based on location data by predicting the probability p when user is in a sensitive location s_i at time t. We describe location data that users supply at any point in time t where it does not reveal the θ privacy of users in a sensitive location. Therefore, it will quantify the adversary location data collection in order to identify the user in a specified sensitive location information.

Definition 1 *(Location-Aware Data Privacy)*

In any time t', it represents the rate of probability for each user at the time of t' at location s_i, we use $P\left\{{U_i}^{t'}\right\}$. We also use Ld_t to represent the data at specific location which collected by attackers at certain time t. As a result, we have

$$P\left\{{U_i}^{t'} \mid L_t\right\} - P\left\{{U_i}^{t'}\right\} \leq \theta \tag{1}$$

where θ denotes the user's privacy requirement, $P\left\{{U_i}^{t'} \mid L_t\right\}$ denotes that the adversary acquires location data subsequent to the user at time t and evaluates the posterior probability of the user being in a vulnerable position s_i at specific time t', and $P\{U_i\}$ denotes the adversary's prior probability of speculating that the user is in a perilous position s_i. According to Def 1, the adversary's collection of user location data cannot surpass θ in order for the attacker to deduce the user's sensitive location. Therefore, after aggregating the position sequence from the users, the discrepancy in probability values between the posterior and prior probabilities of the user being in a specific sensitive position at a given moment is less than θ. At each moment t', the adversary has $-\sum_i P\left\{{U_i}^{t'}\right\} logP\left\{{U_i}^{t'}\right\}$ of previous knowledge about the user's location.

Thus, the amount of sensitive user location data exposed to attackers in ENs can be estimated as follows:

$$Comp(s) = \sum_{i} P\left\{{U_i}^{t'}\right\} logP\left\{{U_i}^{t'}\right\} - \sum_{i} P\left\{{U_i}^{t'} \mid L_t\right\} logP\left\{{U_i}^{t'} \mid L_t\right\} \tag{2}$$

As defined in Definition 1, a privacy-preserving technique that meets θ privacy may guarantee the quantity of information included in the published data in the EN is less than $n\theta$, where n is the quantity of sensitive locations. At any point in time, users with strict privacy needs may set *theta* to 0. As a consequence, users must ensure the location data they post does not jeopardise their θ privacy in any sensitive location. They only need to adhere to the $\frac{\theta}{n}$ standard of the privacy level.

4 System Modeling and Analysis

Figure 1 illustrates the proposed DDSDP model utilised in the SDN-enabled EC service. Both schemes secure the data privacy in location-aware applications between users and ENs. We begin with a customized SDN control layer. This novel control layer attempts to offer a real-time clustering solution. It uses the modified affinity propagation (AP) clustering technique. As the clusters are constantly updated, attackers cannot identify the source of the first connected EN. As a result, clustering protects privacy. The security level is also increased by modifying the Laplacian mechanism. We utilise QoS mapping to assess data utility and privacy protection. Thus, SDN-enabled EC services provide the highest privacy level and data utility protection.

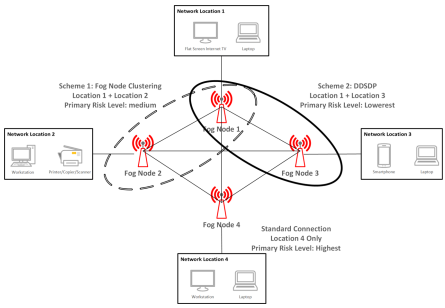

Fig. 1. DDSDP framework in SD-enabled EC.

4.1 Edge Nodes (ENs) Clustering Modelling

We use EN clustering to conceal the real position data from users in the first scheme of proposed DDSDP model. This preservation approach successfully avoids the opponent from quickly invading and allows for more complex attack measures. Each cluster is composed of at least two or more ENs. Instead of direct connections, users are assigned to EN clusters that span a larger region under the adversary attacking paradigm. Adversaries need necessary analytical stages as precise selection gets increasingly difficult.

Weight Factors and Unification Process. Distance-based location is a popular location-aware privacy-preserving technique in EC. We determine the distance based on the user situation by analysing the arrival time and the its difference between the arrivals. Typically, each EN calculates the user position via a distance-based location method. If location information is exposed, user privacy will be jeopardised. Moreover, in order to establish the current location of the user, the selected EN must be aware of the location of each reference node. The geolocation information for the reference node are revealed if the adversary perform spoofing on the location data and other attacks against the reference node.

We calculated the position distances between specified ENs and the initial anchor node using the matrix S where $S = S_1^1, S_2^1, S_3^1, S_\beta^1$ denotes distinct factors associated with the same EN, whereas the elements $S = S_1^1, S_1^2, S_1^3, S_1^\alpha$ denotes the same factor for all ENs. As a result, we construct the matrix as

$$S = \begin{bmatrix} S_1^1 & S_2^1 & S_3^1 & \dots & S_\beta^1 \\ S_1^2 & S_2^2 & S_3^2 & \dots & S_\beta^2 \\ & & \dots\dots\dots & & \\ S_1^\alpha & S_2^\alpha & S_3^\alpha & \dots & S_\beta^\alpha \end{bmatrix} \tag{3}$$

ENs have a variety of indices based on their features, such as access points and servers. Before the aggregate indicator can be computed, a unification procedure is needed to identify the distances. To address the issue of differing absolute values for various indices, the absolute values must be transformed to relative values. S' denotes the matrix S after the unification procedure.

$$S' = \begin{bmatrix} \frac{S_1^1}{\sum_{i=1}^{\alpha} S_1^i} & \frac{S_2^1}{\sum_{i=1}^{\alpha} S_2^i} & \frac{S_3^1}{\sum_{i=1}^{\alpha} S_3^i} & \cdots & \frac{S_\beta^1}{\sum_{i=1}^{\alpha} S_\beta^i} \\ \frac{S_1^2}{\sum_{i=1}^{\alpha} S_1^i} & \frac{S_2^2}{\sum_{i=1}^{\alpha} S_2^i} & \frac{S_3^2}{\sum_{i=1}^{\alpha} S_3^i} & \cdots & \frac{S_\beta^2}{\sum_{i=1}^{\alpha} S_\beta^i} \\ & & \dots\dots\dots\dots & & \\ \frac{S_1^\alpha}{\sum_{i=1}^{\alpha} S_1^i} & \frac{S_2^\alpha}{\sum_{i=1}^{\alpha} S_2^i} & \frac{S_3^\alpha}{\sum_{i=1}^{\alpha} S_3^i} & \cdots & \frac{S_\beta^\alpha}{\sum_{i=1}^{\alpha} S_\beta^i} \end{bmatrix} \tag{4}$$

Cluster Triggering Process. Primarily, all ENs are geographically allocated. When the potential attacks are detected, we initially propose a cluster triggering

technique by allocating measurements according to the entropy weight determination. The EWM's fundamental premise is to derive objective weights of certain variables. The EWM determine the weight of each element in each factor of the attacking formulation. In general, lower e^j value indicates the importance of the element including dataInformation, quantity of the data, and therefore merits a greater weight in the relevant factor. In comparison, a higher entropy e^j implies that an element has a lower value, offers less information, and contributes less to the overall assessment, and therefore should have a lower weight. Due to the unification procedure, the weight factor for each component j in each dimension i must be computed, where $j = 1, 2, 3, ..., m$, and $i = 1, 2, 3, ..., n$.

$$p_{ij} = \frac{S'_{ij}}{\sum_{i=1}^{n} x_{ij}} \tag{5}$$

Each factor j must have an entropy value computed, where $j = 1, 2, 3, ..., m$. $k = 1/\ln(n) > 0$ in this case, and $e_j >= 0$.

$$e_j = -k \sum_{i=1}^{n} p_{ij} \ln p_{ij} \tag{6}$$

A redundancy rate is determined throughout this procedure to minimise variance. The redundancy rate for $j = 1, 2, ..., m$ is $d_j = 1 - e_j$. Following the redundancy correction, the following weight factors are calculated:

$$w_j = \frac{d_j}{\sum_{j=1}^{m} d_j} \tag{7}$$

After assigning weights towards each element per dimension, the clustering procedure is carried out by the weight factors. Assume that t_{tg} is the threshold value for clustering that is dependent on the edge network's distance. T_{tg} is the outcome of the triggering process and is determined by the weight factor computed for each element. Clustering process is defined by the unique EN factor abstracted with the maximum value.

$$T_{tg}^1 = \frac{S_1^1}{\sum_{i=1}^{\alpha} S_1^i} \times t_{tg}$$

$$T_{tg}^2 = \frac{S_2^1}{\sum_{i=1}^{\alpha} S_2^i} \times t_{tg}$$

$$\cdots$$

$$T_{tg}^{\beta} = \frac{S_{\beta}^1}{\sum_{i=1}^{\alpha} S_{\beta}^i} \times t_{tg}$$

$$T_{tg} = max(T_{tg}^1, T_{tg}^2, \ldots, T_{tg}^{\beta}) \tag{8}$$

4.2 Affinity Propagation-Based Clustering

On the basis of a modified AP mechanism, we develop a clustering model. AP is a semisupervised clustering method based on closest neighbour propagation developed by Frey et al. [6]. In comparison with other clustering techniques, AP does not need a final number of clusters to be specified. Rather of creating new data points, the cluster centres are chosen using existing geographical data points. AP approach is less reliant on the initial location information input and does not require a symmetric data similarity matrix. In EC, the input data may be of various types as a result of our triggering mechanisms weight factor-based selections. As a result, the AP method is the optimal choice for grouping ENs.

Preference. We begin by examining the parameter for the preferences. The similarity between each cluster centres is stated as $sim(d, p)$, which reflects the similarities between the data points p and d. The Euclidean distance is used to determine this similarity:

$$sim(d,p) = \sqrt{\sum_{r=1}^{n}(d-p)^2} \times T_{tg} \tag{9}$$

Responsibility. $sim(d, p)$ indicates the degree to which data point p is appropriate for designation as the cluster centre for data point d and represents a message sent from d to p, where $p \in 1, 2..., N$ and $p \neq p'$.

$$r(d,p) = (s(d,p) - max\left\{a(d,p') + sim(d,p')\right\}) \times T_{tg} \tag{10}$$

where $a(d, p')$ is the value showing the accessibility of point i to all points except $k$with a starting value of 0. $s(d, j)$ denotes the responsibility of a point by points other than p, where points other than d are in rivalry with d for ownership. $r(d, p)$ denotes p's cumulative duty to act as the cluster centre for d. When $r(d, p) > 0$, it indicates that p has a greater responsibility to act as the cluster centre.

Availability. To analysis the availability aggregration, $a(d, p)$ indicates the probability that data point d should always choose data point p as its cluster centre and is equivalent to a message delivered from p to d.

$$a(d,p) = min\left\{0, r(d,p') + \sum_{p}\left\{max(0, r(d',p))\right\}\right\} \times T_{tg} \tag{11}$$

$$a(p,p') = (\sum_{p}\left\{max(0, r(d',p))\right\}) \times T_{tg} \tag{12}$$

where $r(d', p)$ indicates the responsibility value of point p as the cluster centre for points other than d. We aggregate all responsibility values greater than or equal to 0, and we also include the responsibility value p as a cluster centre in its

own right. Specifically, all data points with matching responsibility values larger than 0 support point p, and data point d chooses p as a cluster centre based on its cumulative value.

λ Damping Factor. A damping factor is used as the algorithm repeatedly updates the availability and responsibility values. This factor λ enables the AP method to converge more quickly. The damping factor may take on values ranging from 0 to 1. λ operates on the responsibility and availability values throughout each iteration of the algorithm to weight the update in relation to the previous iteration.

$$d_n = (1 - \lambda) \times d_n + \lambda \times r_{n-1} \tag{13}$$

$$p_n = (1 - \lambda) \times p_n + \lambda \times p_{n-1} \tag{14}$$

4.3 ϵ-Customized Differential Scheme

Each user in EC exposes their sensitive location data during connection establishment. However, before this location data is released, it must be protected from disclosure. We developed the first clustering method in order to enhance the difficulty level when an opponent intends to attack. Additionally, the model secures the released data with a tailored degree of protection based on the cluster distances. Additionally, our second scheme seeks to offer consumers with the utmost in privacy protection. To get the greatest protection, we compromised ϵ-customized differential privacy and introduced Laplacian noise to the cluster.

QoS Data Utility Mapping. The previous paragraph specified the distance between each cluster as $sim(i, k)$, and the Softmax function was utilised to describe the data utility with QoS and privacy preservation level as ϵ in the DDSDP model. In a multi-class problem, the Softmax function assigns decimal probability to each class. Moreover, it is often used to visualise the utility of data and the level of privacy protection provided by the QoS mapping. The mapping function is represented as follows:

$$\mathrm{QoS}(\epsilon_i) = k \times \frac{\exp(\theta_i^t sim_{ik} \cdot x)}{\sum_{k=1}^{K} \exp(\theta_k^t sim_{ik} \cdot x)} \tag{15}$$

where $k \in K$ is the parameter used to modify the maximum amplitude value, θ indicates the curve's steepness, and x indicates the position.

Laplacian Mechanism and Laplacian Noise. On preserve anonymity of location, we use probabilistic clustering to the initial results of the single clustering query. To protect users' privacy when it comes to location-aware information, we utilise the Laplacian technique to change the actual value by adding Laplacian noise to the original clustering result data, guaranteeing differential privacy both before and after noise addition.

$$
\begin{aligned}
&M(D) = f(D) + Y \\
&\text{s.t.} \\
&Lap(\alpha) = \frac{p_x(z)}{p_y(z)} = exp(\frac{\epsilon \cdot \parallel f(x) - f(y) \parallel}{\triangle f})
\end{aligned}
\tag{16}
$$

where ϵ specifies the privacy budget and ϵ may be adjusted to obtain a better privacy budget outcome owing to the clustering requirement. The Laplacian distributed noise is determined by Y. $Lap(\alpha)$ denotes the mechanism's probability density, while α denotes the noise's magnitude.

ϵ-Customizable Differential Privacy Formulation. To prevent data release in ENs, we model the ϵ-customized differential privacy method. We map the privacy protection level using the clustering method and multihop with QoS. Users submit sensitive location data under the EC clustering paradigm. These data must be protected against attackers. ENs also vary in capacity and processing capability. As a result, we construct the second scheme as follows.

We utilise ϵ-customizable differential privacy. In the case of $M \rightarrow \theta(\chi)$, we defined the mechanism as follows:

$$
\begin{aligned}
&Pr\left[M(D) \in \Omega\right] = \exp(QoS(\epsilon_i)) \cdot Pr\left[M(D') \in \Omega\right] \\
&= \exp(k \times \frac{\exp(\theta_i^t sim_{ik} \cdot x)}{\sum_{k=1}^{K} \exp(\theta_k^t sim_{ik} \cdot x)}) \cdot Pr\left[M(D') \in \Omega\right] \\
&\text{s.t.} \\
&\forall \Omega \subseteq \chi, \\
&\forall (D, D') \subseteq \psi,
\end{aligned}
\tag{17}
$$

where χ represents the result of the nosied location and D denotes the location sensitive data's storage space. $\epsilon \geq 0$ signifies the proximal relationship between the data, and $\psi \subseteq \forall (D, D') \subseteq \psi$ denotes the proximal relationship between the data. We treat D_t and D_{t+1} as changeable datasets to allow the proposed model to include additional dynamic characteristics.

Three criteria are defined in order to undermine the configurable privacy protection approach and clustering model. Initial ϵ-customizable protection is provided by the first qualifier. Each piece of sensitive location data is referred to as p_i, and $\epsilon(\frac{1}{d_{ik}})$ should be fulfilled by anticipations $\{y_{ik}\}$ from p_k. The second qualification's purpose is to specify a maximum degree of privacy protection for the upper bound EN which where the sources are.

The second requirement specifies the degree of privacy protection that the upper limit EN should be maximum. As stated above, the composition in its entirety represented as:

$$
com(\epsilon) = \sum_{i=1, k \neq 1}^{n} M_D(\epsilon(\frac{1}{d_{ik}}))
\tag{18}
$$

In the third qualification, we optimise user-published data utilities. The response of the real output x_d should be the most accurate noisy n from mechanism M. Additionally, various approximations $\{y_{dp}\}$ result in a range of data utility optimisation values. Thus, we can represent the total usefulness of the data as follows:

$$\sum_{d=1}^{n}\sum_{p\neq 1}^{n} E\| \, y_{dp} - x_i \, \|_2^2 \tag{19}$$

Thus, optimal tradeoffs will be evaluated based on the anticipated greatest degree of privacy protection and lowest data usefulness.

$$\begin{aligned} \epsilon = k \times & \frac{\exp(\theta_i^t sim_{ik} \cdot x)}{\sum_{k=1}^{K}\exp(\theta_k^t sim_{ik} \cdot x)} \\ & \sum_{i=1,k\neq 1}^{n,n} M_D(\epsilon(\frac{1}{d_{ik}})) \leq maxM_D(\epsilon(\frac{1}{d_{ik}})) \\ & \sum_{i=1}^{n}\sum_{k\neq 1}^{n} E\| \, y_{ik} - x_i \, \|_2^2 \geq min(DU) \end{aligned} \tag{20}$$

5 Performance Evaluation

To evaluate the performance of our proposed DDSDP model, we examine a series of simulations in this section. We first assess data utilities by sampling time periods at various places, then privacy protection levels at different locations. Finally, the experiment assesses the clustering approach's performance, including clustering, transmission and loading outcomes for the total ENs. We used the VicFreeWiFi Access Point Locations dataset [1] to validate these findings. The dataset covers over 300 kms distances geographically and minimum of 250 MB data flow per device every day. In the dataset, it contains 391 nodes are located in the city centre, 44 nodes are located in the northbound zone, and 82 nodes are located in the west-northbound region. This data impairs adversaries ability to identify the location of users.

Due to the fact that SDN-enabled EC should allow more customization choices without sacrificing original performance, we begin by analysing clustering efficiency and evaluating network performance as a result of SDN-enabled clustering. Furthermore, we evaluate our DDSDP model against a range of ϵ values in order to optimise the ϵ-specific differential privacy protection technique. Additionally, two additional methods are compared: classic customisable differential privacy (CCDP) and classic ϵ-differential privacy (CDP). In CDPs, the Laplacian process produces noise. The CCDP provides a customizable degree of privacy and adheres to Laplacian noise.

5.1 Clustering Analysis

Figure 2 illustrates the node clustering results from our DDSDP model's initial scheme clustering technique. The expected number of clusters is 16, created from

517 accessible ENs. We assess system performance using three similarity values. The lowest similarity is 2.00000e-9, the median is 0.017874, and the highest is 1.276488. Likelihood is dependent on location, length, and connection speed. The homogeneity rate is 0.513, which shows how closely related nodes are grouped. This dataset's rate is heavily influenced by connection speed. It has a V-measure and completeness of 0.333 and an adjusted Rand index of 0.080. The first scheme clustering system produces excellent clustering results.

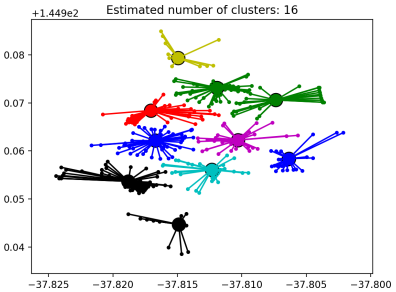

Fig. 2. Edge Node (EN) clustering results.

5.2 Data Utilities Performance

Figure 3 illustrate the outcomes of our DSDP data utilities. The figure depicts the overall QoS functionality. We compare our findings with raw data values for $\epsilon = 0.1$, $\epsilon = 0.5$, and $\epsilon = 1$, which makes it relevant to different situations. We begin by aggregating 20 clustered EC nodes based on QoS metrics derived from the distances between cluster. The Laplacian process generates a large amount of noise in the form of responses. As illustrated in the figure, decreasing the value of ϵ results in improved overall data utility performance. When $\epsilon = 0.1$, we use clustering time slot 5 to achieve the peak value of 1.7.

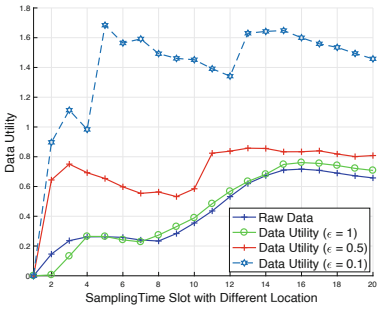

Fig. 3. Data utilities performance with three ϵ values.

Furthermore, we examine various clustering scenarios in terms of privacy protection performance. To build up the customised ϵ, we enable three sample parameter values from the data utilities evaluation. In order to mimic the Laplacian mechanism's unpredictability, three ϵ values were chosen. Figure 4 compares privacy protection levels in terms of configurable ϵ. It achieves a maximum privacy protection level of 1.5 while sampling time slot 7, while sampling time slot 20 maintains the greatest degree of privacy protection. Although the performance for three parameters varies across clusters, it demonstrates the significance of customisation. Moreove, we set $\epsilon = 1$ for the cluster in time slot 4.

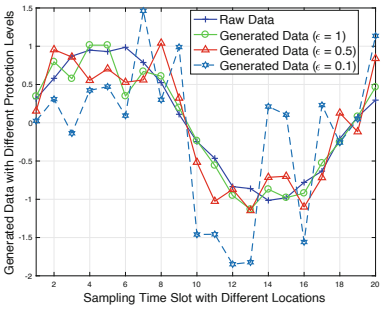

Fig. 4. Different locations privacy levels.

5.3 Performances Against Attacking Scenarios

Figure 5 demonstrates the performance level of privacy protection in a wormhole attack scenario against solitary and clustered ENs. The DDSDP model is compared to the other two classic models. Two dotted green lines show the adversary's data's ϵ value. A lower ϵ results in greater privacy protection. Since our approach relies on ϵ-customized differential privacy, a lower value of ϵ indicates less data given to adversaries. ϵ has two values. When ($\epsilon = 0.45$), the adversary can identify most of the location data and familiar with the other attacking models. ($\epsilon = 2.05$) is the point when CDP and CCDP are completely functioning and can offer comprehensive protection for consumers. All ϵ values between the dashed green lines except DDSDP model provide optimum protection. As a result, DDSDP can offer stronger protection strategies when assaults occur where CDP and CCDP suffer.

Figure 6 illustrates the performance of proposed DDSDP model against a DDoS attack. We utilise ϵ values of 0.45 and 2.05 from previous figures. Because the composition process is still free, CDP is unaffected by DDoS attacks. On top of that, the ϵ rises from 0.45 for DDSDP and CCDP. However, DDSDP shows that two or more opponents will not collect any additional information following an assault. *epsilon* total = maximum ϵ value at time of assault. As a result, the DDSDP model we presented above outperforms both common attacking techniques.

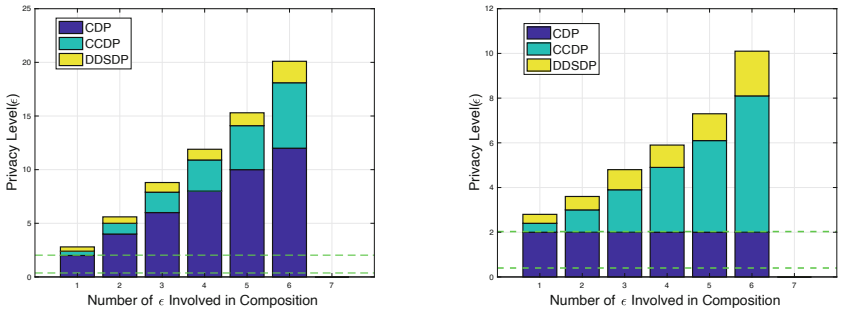

Fig. 5. Attacking Scenario 1: (a) Location data shared by multiple users to individual Edge Node (EN); (b) Location data shared by multiple users to clustered Edge Nodes (ENs).

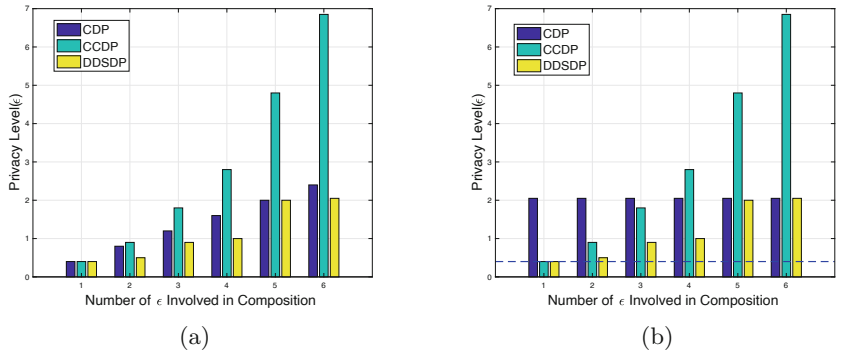

Fig. 6. Attacking Scenario 2: (a) Location data shared by multiple users to individual Edge Nodes (ENs); (b) Location data shared by multiple users to clustered Edge Nodes (ENs).

6 Conclusion and Future Works

In this paper, we proposed a privacy protection mechanism (DDSD) for SDN-enabled EC. Our proposed model includes two schemes: SDN based clustering EN and ϵ-customized differential privacy to defend against two common attacker techniques. As a first step, we created our first clustering-based protection system. As a result, attackers cannot identify the location data in the first place. Furthermore, we integrated ϵ-customizable differential privacy model where adding noises into the SDN based EN cluster. Our extensive experimental findings established that the proposed model and clustering approach not only exhibit high reliability in specified situations, but also offer configurable location-aware data privacy protection. More over, we consider to further optimize the data utilities and privacy loss by considering Markov Decision Process

(MDP) where will provide optimal tradeoff solutions. Moreover, reinforcement learning methods will also be tested to optimize the convergence speed.

References

1. Vicfreewifi access point locations - victorian government data directory, July 2017. https://discover.data.vic.gov.au/dataset/vicfreewifi-access-point-locations
2. Badsha, S., et al.: Privacy preserving location-aware personalized web service recommendations. IEEE Trans. Serv. Comput. **14**(3), 791–804 (2018)
3. Bonomi, F., Milito, R., Natarajan, P., Zhu, J.: Fog computing: a platform for internet of things and analytics. In: Bessis, N., Dobre, C. (eds.) Big Data and Internet of Things: A Roadmap for Smart Environments. SCI, vol. 546, pp. 169–186. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-05029-4_7
4. Dang, T.D., Hoang, D.: A data protection model for fog computing. In: 2017 Second International Conference on Fog and Mobile Edge Computing (FMEC), pp. 32–38 (2017)
5. Deepali, Bhushan, K.: DDoS attack defense framework for cloud using fog computing. In: 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information Communication Technology (RTEICT), pp. 534–538 (2017)
6. Frey, B.J., Dueck, D.: Affinity propagation (2007)
7. Xia, Q., Tao, Z., Li, Q.: Privacy issues in edge computing. In: Chang, W., Wu, J. (eds.) Fog/Edge Computing For Security, Privacy, and Applications. AIS, vol. 83, pp. 147–169. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-57328-7_6
8. Gu, B., Wang, X., Qu, Y., Jin, J., Xiang, Y., Gao, L.: Context-aware privacy preservation in a hierarchical fog computing system. In: 2019 IEEE International Conference on Communications (ICC), ICC 2019, pp. 1–6. IEEE (2019)
9. Gu, B.S., Gao, L., Wang, X., Qu, Y., Jin, J., Yu, S.: Privacy on the edge: customizable privacy-preserving context sharing in hierarchical edge computing. IEEE Trans. Netw. Sci. Eng. **7**, 2298–2309 (2019)
10. Hasan, K., Ahmed, K., Biswas, K., Islam, M.S., Kayes, A.S.M., Islam, S.M.R.: Control plane optimisation for an SDN-based WBAN framework to support healthcare applications. Sensors **20**(15) (2020). https://doi.org/10.3390/s20154200. https://www.mdpi.com/1424-8220/20/15/4200
11. Hasan, K., Ahmed, K., Biswas, K., Saiful Islam, M., Ameri Sianaki, O.: Software-defined application-specific traffic management for wireless body area networks. Future Gener. Comput. Syst. **107**, 274–285 (2020). https://doi.org/10.1016/j.future.2020.01.052, https://www.sciencedirect.com/science/article/pii/S0167739X19322587
12. Ho, S., Qu, Y., Gu, B., Gao, L., Li, J., Xiang, Y.: DP-GAN: differentially private consecutive data publishing using generative adversarial nets. J. Netw. Comput. Appl. **185**, 103066 (2021)
13. Kang, J., Yu, R., Huang, X., Zhang, Y.: Privacy-preserved pseudonym scheme for fog computing supported internet of vehicles. IEEE Trans. Intell. Transp. Syst. **19**(8), 2627–2637 (2018)
14. Lu, R., Heung, K., Lashkari, A.H., Ghorbani, A.A.: A lightweight privacy-preserving data aggregation scheme for fog computing-enhanced IoT. IEEE Access **5**, 3302–3312 (2017)
15. Lyu, L., Nandakumar, K., Rubinstein, B., Jin, J., Bedo, J., Palaniswami, M.: PPFA: privacy preserving fog-enabled aggregation in smart grid. IEEE Trans. Ind. Inf. **14**(8), 3733–3744 (2018)

16. Ma, L., Liu, X., Pei, Q., Xiang, Y.: Privacy-preserving reputation management for edge computing enhanced mobile crowdsensing. IEEE Trans. Serv. Comput. **79**, 500–513 (2018). Part 2
17. Mouradian, C., Naboulsi, D., Yangui, S., Glitho, R.H., Morrow, M.J., Polakos, P.A.: A comprehensive survey on fog computing: state-of-the-art and research challenges. IEEE Commun. Surv. Tutor. **20**(1), 416–464 (2018)
18. Nafi, N.S., Ahmed, K., Gregory, M.A., Datta, M.: Software defined neighborhood area network for smart grid applications. Future Gener. Comput. Syst. **79**, 500–513 (2018). https://doi.org/10.1016/j.future.2017.09.064, https://www.sciencedirect.com/science/article/pii/S0167739X17311007
19. Ni, J., Zhang, K., Lin, X., Shen, X.: Securing fog computing for internet of things applications: challenges and solutions. IEEE Commun. Surv. Tutor. **20**(1), 601–628 (2018)
20. Qu, Y., Yu, S., Gao, L., Zhou, W., Peng, S.: A hybrid privacy protection scheme in cyber-physical social networks. IEEE Trans. Comput. Soc. Syst. **5**(3), 773–784 (2018)
21. Rasool, R.U., Ashraf, U., Ahmed, K., Wang, H., Rafique, W., Anwar, Z.: CyberPulse: a machine learning based link flooding attack mitigation system for software defined networks. IEEE Access **7**, 34885–34899 (2019). https://doi.org/10.1109/ACCESS.2019.2904236
22. Stojmenovic, I., Wen, S.: The fog computing paradigm: scenarios and security issues. In: 2014 Federated Conference on Computer Science and Information Systems, pp. 1–8, September 2014. https://doi.org/10.15439/2014F503
23. Wang, Q., Chen, D., Zhang, N., Ding, Z., Qin, Z.: PCP: a privacy-preserving content-based publish subscribe scheme with differential privacy in fog computing. IEEE Access **5**, 17962–17974 (2017). https://doi.org/10.1109/ACCESS.2017.2748956
24. Wang, T., Zhou, J., Chen, X., Wang, G., Liu, A., Liu, Y.: A three-layer privacy preserving cloud storage scheme based on computational intelligence in fog computing. IEEE Trans. Emerg. Top. Comput. Intell. **2**(1), 3–12 (2018)
25. Wang, W., Zhang, Q.: Privacy preservation for context sensing on smartphone. IEEE/ACM Trans. Netw. **24**(6), 3235–3247 (2016). https://doi.org/10.1109/TNET.2015.2512301
26. Wibowo, F.X., Gregory, M.A., Ahmed, K., Gomez, K.M.: Multi-domain software defined networking: research status and challenges. J. Netw. Comput. Appl. **87**, 32–45 (2017). https://doi.org/10.1016/j.jnca.2017.03.004, https://www.sciencedirect.com/science/article/pii/S1084804517300991
27. Yannuzzi, M., Milito, R., Serral-Graci, R., Montero, D., Nemirovsky, M.: Key ingredients in an IoT recipe: fog computing, cloud computing, and more fog computing. In: 2014 IEEE 19th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), pp. 325–329 (2014)
28. Zhang, J., Feng, X., Liu, Z.: A grid-based clustering algorithm via load analysis for industrial internet of things. IEEE Access **6**, 13117–13128 (2018)