Data and Location Privacy of Smart Devices over Vehicular Cloud Computing

Hani Al-Balasmeh, Maninder Singh
Department of Computer Science and Engineering
Thapar Institute of Engineering and Technology, India
{hbalasmeh_phd17 & Msingh}@thapar.edu

Raman Singh
School of Computing, Engineering, and Physical Sciences
University of the West of Scotland, UK
Raman.Singh@uws.ac.uk

Abstract—In this paper, we have addressed the problem of data and location privacy in smart devices over vehicular cloud computing (VCC). We proposed a framework to identify and register the smart IoT GPS devices over VCC service and allow the users to monitor their devices in real time. The proposed framework divides into three parts: First, data anonymization of users’ information over VCC, by masking the original data of the user and replaced with fake data. The proposed technique will remove the user identity and other linkers to identify the users. Second, proposed a technique using asymmetric cryptography (RSA) technique, the proposed technique provides location privacy of users’ trajectories before requesting point of interest (POI) from location-based services (LBS). Third, secure communication between users and the VCC, based on Token-based authentication by authenticating the trusted users while requesting a location from the VCC service. The proposed framework shows the efficiency and reliability of responding to user trajectories from different sources of IoT GPS devices and datasets.

I. INTRODUCTION

The smart technologies and internet of things (IoT) devices paradigm have experienced fast growth in popularity and attention recently in our daily life. IoT is defined as objects or people or devices that are provided with unique identifiers of each and the ability to transfer data over different types of networks into the cloud without requiring human interaction. Different types of application fields the IoT devices play such as smart cities, smart roads, smart vehicles, smartphones, smart health care, etc. Most of these applications are required from IoT devices sending and receiving data between device and server or requesting from the user to enable some services such as location services in smart vehicles and smartphones. Increasing the leaking of sensitive data of users from different sources in cloud computing is a critical problem that will make the user’s life in danger. In this paper, we address this problem of data and location privacy of the user’s over VCC. We proposed a framework ‘TIETGTS’ to register and monitor the smart IoT GPS devices and propose a method to provide data anonymization and location privacy over VCC. The framework has been developed to register the different types of smart IoT GPS devices; we have tested our framework with the smart IoT GPS device ‘GTU’ and smartphone android and data-set ‘Taxi Trajectory Data. The framework shows the efficiency of receiving the data from different sources of smart deceives in VCC. In our proposed method of location privacy, we used the asymmetric RSA algorithm over VCC of receiving the user’s trajectories, where each data row in the database will encrypt with different public and private keys. Also, we proposed data anonymization of sensitive data of the users over VCC. Fig. 1. shows the registers and monitoring of the IoT GPS device over VCC using the TIETGTS framework. For secure communication between the client and server-side over VCC the TIETGTS framework provides authentication communication by generating token authentication of the user’s login and hashing the user login details using ‘JWT-HS256’ over hypertext transfer protocol HTTP. Where the token authentication will be destroyed once the user logout from the system.

II. RELATED WORK

Addressing the threat data in smart home IoT, the authors [1] propose an algorithm for distributed data integration and describe the challenge and analysis of the threat of information linkage of smart devices in the IoT ecosystem. The description applied in the smart home data and showing the processing tasks lead the unintended privacy breaches of the identification of the objects of smart home and the linkage information between their availability and resource use. For privacy preservation in software-defined networks (SDN) the smart city the authors [2] proposed privacy-preserving for equipped IoT devices based on the environment using the SDN paradigm. In the proposed the IoT devices will decide to route the data depending on sensitivity information by splitting sensitive data into two parts and sending the split
part through a secure VPN that will be created between IoT and SDN. To provide the privacy preservation of IoT data environment on cloud computing, the authors [3] proposed a systematic model for cloud data centers (CDCs) of deploying the IoT data over cloud environment and improving average resource utilization and accessing the performance of CDCs while reducing the energy consumption using genetic algorithm II. With enhancing the data transmission privacy in the E-health IoT application, the authors [4] proposed a novel building path algorithm “BPA” to ensure content privacy by encrypting the data transmitting using a symmetric key “L-IBE” encryption scheme. With enhancing the anonymity and authentication in fog IoT systems, the authors [5] proposed an anonymous privacy-preserving scheme with authentication in fog-enhanced IoT systems, by providing authentication in multi-layers of IoT devices using a pseudonym and pseudonym certificate with a non-unique user data in smart devices (SDs). To provide the security and privacy of user trajectories, the authors [6] proposed an anonymous framework model to provide privacy of trajectory data generated in the intelligent transportation system of IoT devices, by replacing the location with an anonymous location using the personal correlated location by taken the correlation between locations and individuals using k-correlation region. For hiding the real location before requesting a POI the authors [7] proposed a novel model to provide location privacy protection in a trusted third party (TTP) between the user and LBS by collecting the real user location and sending it to the LBS, where is the model is based on loss of services quality by finding the obfuscated location and region. The model will add a noising location to improve the quality of the services. By generating dummy locations around a real location, the authors [8] proposed query anonymization method and dummy generation location algorithms to achieve the k-anonymity privacy-aware area in LBS and generate a random dummy location near the reallocation of the user. With the increase in the leakage of user location in LBS, the authors [9] proposed anonymous entropy-based location privacy in mobile social network “MSN” to solve the problem of leakage of user location privacy by using two types of methods K-DDCA for densely populated region and K-SDCA for the sparsely populated region and generate dummy locations algorithms. The protection of location privacy of users that sacrifice quality services (QoS) does obtain accurate query results. To address this problem, the authors [10] proposed location privacy preservation based on the k-anonymity method credible chain with different two features of optimal k value of current user and second, using anonymizing spatial region (ASR). The trusted third-party server (TTP) will be generating fake trajectories based on the number of user locations. Many leaking location information of users in mobile networks while requesting the users to share their location in TTP, by addressing this problem the authors [11] proposed a framework model of various location privacy-preserving mechanisms and describe the characterizing the effectiveness of the different location privacy approaches. The method will give the difference between the initial uncertainty trace before and after the adversary’s attack. With the problem of location privacy in mobile networks and using the nearest neighbor queries with the current location of the user while requesting a query from an untrusted location server, the authors [12] proposed a method to protect location privacy as a hidden ring and hidden forest (HRHF), by providing effectiveness and efficiency of protecting location privacy. With the increased, profound implications on the privacy of personal information in LBS, the authors [13] proposed a framework for preserving location privacy based on modifying the user location while sending it to the services provider. And prevent the service provider from knowing the exact user’s location movement information, also from being disclosed to other users who are not authorized to access this information. With increasing the different technologies based in LBS, However, their services are exposing the privacy and security of location context users and other information, to address this problem the authors [14] proposed an efficient location privacy-preserving based on location cloaking, by using cloaked region construction method (CRCRA) to protect the user’s query and linking location, such as center point or boundary and trajectory’s location from attacks. The authors in [15] propose a framework TIEFTGEO to provide location obfuscation of user trajectories in a vehicular (VCN) cloud network. The proposed framework will generate an obfuscation of user movements to hide their true trajectories. The researchers in [16] proposed a mix-zone scheme using the pseudonym technique to provide location privacy in VCN. The proposed scheme will anonymous the pseudonym vehicles’ keys when they move from one zone to another. In a dynamic Mix-zone, the researcher in [17] proposed a dynamic Mix-Zone for location privacy by predicting the location of the vehicle in a specific zone by encrypting the messages of the vehicles when transmitted within the Mix-zone. On the internet of vehicles, the researcher in [18] proposed a novel location privacy preservation when the vehicles request location from the LBS services. In vehicular ad hoc networks (VANET), the researcher in [19] a proposed scheme to protect the vehicles on a beginning track from the attacker by changing the pseudonym key effectively [20] to avoid linking messages of the vehicle. Protected the user privacy when requesting POI from the LBS, the researcher proposed reinforcement learning (RL) based on a sensitive semantic location privacy protection [21]. The proposed scheme is based on releasing the semantic historical location of the vehicle. To provide the personalized location privacy of users when requesting POI from the LBS, the researcher in [22] proposed a location privacy protection scheme for user movements. The proposed scheme will anonymous the location movement when requesting a POI from LBS. Table I and II show the summary of the related works.
### Table I
#### A. Summary of the Related Works

<table>
<thead>
<tr>
<th>Auth</th>
<th>Proposed Method</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1]</td>
<td>ecosystem k-anonymity</td>
<td>for distributed data integration the threat of information linkage of smart devices IoT in the smart home.</td>
</tr>
<tr>
<td>[2]</td>
<td>privacy-preserving scheme</td>
<td>k-anonymity privacy preservation in software-defined networks (SDN) the smart city.</td>
</tr>
<tr>
<td>[3]</td>
<td>a systematic model</td>
<td>k-anonymity To improving average resource utilization and accessing the performance of CDCs.</td>
</tr>
<tr>
<td>[5]</td>
<td>anonymous privacy-preserving scheme</td>
<td>pseudonym enhancing the anonymity and authentication in fog IoT systems.</td>
</tr>
<tr>
<td>[6]</td>
<td>anonymous framework</td>
<td>k-anonymity, dummy generation To provide the security and privacy of user trajectories in the intelligent transportation system.</td>
</tr>
<tr>
<td>[7]</td>
<td>a novel location privacy protection</td>
<td>k-anonymity, dummy generation For hiding the real location before requesting a POI from the LBS.</td>
</tr>
<tr>
<td>[8]</td>
<td>query anonymization method and dummy generation location</td>
<td>k-anonymity, dummy generation to achieve the k-anonymity privacy-aware area in LBS and generate a random dummy location near the reallocation of the user.</td>
</tr>
<tr>
<td>[9]</td>
<td>anonymous entropy-based location privacy</td>
<td>K-DDCA, K-SDCA to solve the problem of leakage of user location privacy in the MSN.</td>
</tr>
</tbody>
</table>

### Table II
#### B. Summary of the Related Works

<table>
<thead>
<tr>
<th>Auth</th>
<th>Proposed Method</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>[10]</td>
<td>location privacy preservation scheme</td>
<td>k-anonymity To protection of location privacy of users that sacrifices quality services (QoS).</td>
</tr>
<tr>
<td>[11]</td>
<td>various location privacy-preserving mechanisms</td>
<td>k-anonymity Provide a difference between the initial uncertainty trace before and after the adversary's attack.</td>
</tr>
<tr>
<td>[12]</td>
<td>method to protect location privacy</td>
<td>k-anonymity providing effectiveness and efficiency of protecting location privacy in untrusted services.</td>
</tr>
<tr>
<td>[13]</td>
<td>preserving location privacy framework</td>
<td>k-anonymity increased, profound implications on the privacy of personal information in LBS.</td>
</tr>
<tr>
<td>[14]</td>
<td>efficient location privacy-preserving CRCA, k-anonymity</td>
<td>To protect the user's query and linking location, such as the center point or boundary.</td>
</tr>
<tr>
<td>[15]</td>
<td>DLP Obfuscation</td>
<td>To obfuscate the user trajectories in the VCN.</td>
</tr>
<tr>
<td>[16]</td>
<td>mix-zone scheme</td>
<td>mix-zone, pseudonym will anonymize the pseudonym vehicles’ keys when they move from one zone to another.</td>
</tr>
<tr>
<td>[17]</td>
<td>dynamic Mix-Zone Mix-Zone, RSA</td>
<td>To predict the location of the vehicle in a specific zone.</td>
</tr>
<tr>
<td>[18]</td>
<td>novel location privacy-preservation</td>
<td>k-anonymity To provide the location privacy-preservation in the internet of vehicles.</td>
</tr>
<tr>
<td>[19]</td>
<td>Location privacy scheme</td>
<td>pseudonym To protect the vehicles on a beginning track from the attacker.</td>
</tr>
<tr>
<td>[20]</td>
<td>Location privacy scheme</td>
<td>pseudonym To avoid linking messages of the vehicle.</td>
</tr>
<tr>
<td>[21]</td>
<td>RL</td>
<td>k-anonymity Protected the user privacy when requesting POI from the LBS.</td>
</tr>
<tr>
<td>[22]</td>
<td>location privacy protection scheme</td>
<td>k-anonymity To anonymous the location movement when requesting a POI from LBS.</td>
</tr>
</tbody>
</table>
III. METHOD OF DATA AND LOCATION PRIVACY OVER VCC

In this section of the paper, we are going to address the methods of data and location privacy over VCC.

A. Data privacy over VCC

Cloud technology has given opportunities to different types of businesses in our daily life. While it provides various advantages, such as data storage, data management, business automation, and resource optimization. With include all the cloud advantages provided still there is a critical issue and risk of using these different services. One of these issues is data privacy; every cloud service is facing the challenge of proving the security and data privacy of the user information. Different challenges of data privacy such as:

- Data Replication: mostly of online services face this challenge while storing the data in the cloud. But still aware of where these data have been stored and who has control to access and see them? Whoever is there is an identification of unauthorized access or copy of the data.
- Data Loss: virtual data can be lost easily as it transfers between different virtual machines (VMs) or in the cloud service itself, and it will generate a disaster for any business.
- Internal Threat: this is the biggest challenge the organization can face, where the authorized employees (system administrators) can have a chance to access the data every time in cloud-based services.
- An insecure Application Programming Interface (API): APIs can be a threat to cloud security by allowing different types of users to customize their cloud computing activities, by integrating their applications with other software.

In this part of the paper, we introduce a method to provide data anonymization techniques over VCC based on two approaches Static Data Mask (SDM) and Dynamic Data Mask (DDM). There are two basic forms of data masking: static and dynamic. Static Data Masking (SDM) permanently replaces sensitive data by modifying it at rest. Dynamic Data Masking (DDM) is designed to replace sensitive data in transit, leaving the original data intact and unmodified. The proposed framework will provide the data anonymization of the sensitive information of the user’s details, such as name and email address and phone number, login details of the users’ and masking it with fake data by removing any linker to identity to identify the user while publishing the data over VCC. Figure 2 is showing the data anonymization of original data and hiding the real identity of the user by replacing it with fake data by removing and linker in data that discover the original identity of the user.

B. Location privacy using the RSA method

The smart devices start transmitting the data through Transport control protocol (TCP) and internet protocol (IP) by specifying the port number in VCC for receiving the trajectory of smart devices in real-time as shown in figure 2 The data processing of smart device trajectories will be inserted into the data privacy block to encrypt the data and inserts it into the database. In this paper, we introduced an asymmetric cryptography algorithm using the RSA method by [23], to encrypt the recording data and location trajectories of smart devices over VCC. The proposed method will be no exchanging of keys between the server and client-side as the encrypted data will hide the public key inside it and insert it into the database, where every row of data in the database will have different public and private keys. Figure 3 For explaining the steps of encrypted trajectories data in VCC. The proposed method will be selecting the receiving trajectory data from IoT devices and inserting it into the encrypts block. RSA technique is based on generating two pair keys for encrypting and decrypting data, the proposed method is based on generating different pairs of keys for each inserting data row in VCC storage. The proposed method will be to encrypt the data in three steps: The first step is, Select the location data from the VCC storage and encrypts the data using the private key, the second step is, attaching the public key to the text string in the data row and calculating the string length. The third step, Encoding the encrypted data and storing it in the secondary VCC storage.

Enhancing secure accessing the GPS location information of the user’s mobile devices, the researchers proposed a method to secure GPS data information of users based on using the Identity Based Mediated RSA algorithm (IB-MRSA) by using the ID of the users to do public key generation and private keys for securing the GPS data [24]. Securing data, management of the authorized access of the users, and applying different maintaining the privacy of their data in the cloud storage services. The researchers [25] proposed a method based on using the asymmetric algorithm RSA, to secure the distributed database of users’ information. The algorithm 1 describes the steps of data encryption of the location information (lat, lon).
Algorithm 1 Algorithm Data Encryption

Input: lat, lon,
Output: Encryption Data

Function Rsa_encryption(list[lat], list[lon]):
  for i in list[lat, lon]
    M = lat, lon
    Prime p, q = Generate random(p, q)
    if p ≠ q then
      return n = p * q
    else
      Get to Rsa_encryption(list[lat], list[lon])
    end
  end
  Fn = (p - 1)(q - 1)
  V = e [gcd(, e) - 1; 1 < e < (Fn)]
  Compute: d = mod Fn = 1
  KU = e, n
  RU = d, n
  cipher = e * m^2(mod n)
  Size_m = length(m)
  merg_m = m(d, n, size_m) at the end
  Store → merg_m in VCC_DB

End Function

Algorithm 2 Algorithm Data Decryption

Input: merg_m
Output: Decryption Data

Function Rsa_decryption(list[merg_m]):
  for i in list[merg_m]
    ML = length(merg_m)
    Get last 2bit(ML) → merg_m
    split(M) → (merg_m) from right to left by ML
    split(N) → (merg_m) from left to right by 2bit from ML
    split(d) → (merg_m) from left to right from split(N)
    d, n = split(d), split(N)
    KU = d, n
    E_m = split(m)
    cipher = E_m^2(mod n)
    send → cipher to user
  end

End Function

Fig. 5 shows the steps of decrypting the data per request from the user in VCC. Decrypting part goes into three steps: The first step is, Select the encoding data from the VCC storage and decode it. The second step is, to split the data by 2-bit to get the public key from the decoding string, third step, Decrypting the data with the public key and sending the location data to the user.

The algorithm 2 describe the steps of data decryption of the location information (lat, lon).

C. Authentication users over VCC

Token authentications improve the quality and efficiency of the website over the internet, by reducing the users to constantly log in and human-made security. Token-based authentication it’s based on a web authentication technique by letting the users access the profile page using a username and password and generating an encrypted token between the client and server side. The login credential will use the token key to protect accessing authorized pages and the designated period of it. With enhancing the privacy authentication of users in cloud computing, the researchers [26] proposed a technique to verify the authorized privilege of users by using their Passwords, access tokens, and private keys, which are required for authenticating and authorizing access in the cloud services. Token-based authentication is a type of authenti-
cation that is a stateless process without carrying out any information about the user. The researchers [27] tested the authentication users’ requests based on JSON-web token on cookie storage using Cross-Site Request Forgery (CSRF) using fake requests from the attack by forcing authenticated users to submit a request to a Web application against in the cloud services which they are currently authenticated. Using two-factor authentication based on JWT, the researcher [28] proposed a technique using Hyperledger Fabric v1.0 with the JWT authentication method in the cloud service. The method is based on the Time-Based One-Time Password (TOTP), which generates OTP tokens for user authentication when accessing the services. Enhancing the privacy access of the management for cloud SaaS applications. The researchers [29] proposed a model that harnesses the stateless and secure nature of JWT for user authentication and session management in cloud services. To improve the CIDM (Consolidated Identity Management) referred to as EIDM (Ethereum-based Identity Management) for user authenticate protocol. The researchers [30] proposed a method to improve the protocol based on the JWT (JSON Web Token) in OAuth 2.0 to provide a credible identity authentication protocol for cloud users and service providers using the smart contracts into EIDM protocol and the credit management system. Using SWIM (System Wide Information Management) in the network system, the researchers [31] proposed a method to provide the security gateway for SWIM to authentication and authorization process of users’ activities in the services based on JWT (JSON Web Tokens) technique. Authenticated the users in the IoT-cloud services, the researchers [32] proposed a method for authenticating the subsequent user’s requests without making frequent calls to the resource server or database service. the proposed method was based on using the JWTs (JSON Web Token) by generating users’ requests based on the timestamp value to enhance the authentication technique. Enhancing the quantum authentication in the network services, the researcher [33] proposed a protocol to realize network authentication utilizing quantum tokens, by the validity of the period session time of the token by checking it by the server when the user accessing the service.

The stages of generating tokens will allow the website to add more layers of security to enhance the privacy login, in figure 6 shows the process of generating the token key between the client and server.

1) Enter login details of the users, such as (username, and password).
2) Verify the server-side of login details and generates a secure signed token at a particular login time.
3) The signed token will send it to the user back through the browser HTTP/HTTPS.
4) The Server-side will decode and verify the attached token when the user browsing the particular pages on the server.
5) The lifetime of destroying the token will be at the time of login out from the page.

Fig. 5. Token authentication steps

Our proposed method builds using JSON Web Token “JWT” by consists of three parts:

- **Header:** the header part will consist of two parts, token and hashing algorithm. In our method hash token will be based64 on “HMAC256”, where the JSON format will be: “Algo”: “HS256”, “Type”: “JWT”
- **Payload:** will contain the actual login details and send them to the server for verification. And we assume the credentials are valid for the user, and the server will return a new JSON Web token, containing some options:
  - Iss: contains the details of the user login. {
    - “UserName”: “Hani”
    - “Password”: “hani12345” }
  - Exp: Expiration time of the token by specific the validity of the token key.
  - Admin: Boolean description of the role of the user in the system. ({ “Iss” : “Hani”
    - “Exp” : 1550946580”
    - “Admin”: false }
- **Signature:** verify successful sending message from client to server using the private key, and encoding the data based on base64-URL.

Several authors applied different techniques of token authentication such as in smartphones to improve the security of the users, the authors [34] proposed a prototype that performed a secure authentication without requiring root access to a smartphone by locking and locking the smartphone using hash-256. For data exchange between client and cloud server, the authors [35] proposed the mechanism “disposable Token” using token-based authentication using RESTful web services API, by requesting the client to store the public and private key of token-pair while establishing the communication between client and cloud server, where the public key will be stored in client-side and verification it by private key by current timestamp to check the validity of the key.
IV. RESULTS

The proposed approach exhibits flexibility and efficiency in registering and receiving location trajectories of different IoT devices. The approach is tested using Amazon Web Services on a 2016 windows server with the following specifications: Intel® Xeon® CPU E-5-2676 v3® with 2.40 GHz and 1 GB RAM. Figure 6 shows the number of successes receiving the GPS trajectory from different sources of smart devices such as smart android phones and IoT GPS device model ‘GT06’, and we used a dataset of GPS trajectory to test our efficiency framework.

![Fig. 6. The number of receiving Trajectories](image)

Fig. 6. The number of receiving Trajectories

Fig. 7 shows the responding time of generating the fake data based on both data masking approaches of the user’s information. Two data masking approaches using in this paper SDM and DDM to anonymize the data in the VCC when the user requests data. The proposed approach shows that the consuming time to generate the SDM for data masking is higher than that of the DDM approach. The DDM approach shows the stability of increasingly consuming time based on the number of data packages required to mask.

![Fig. 7. Time analysis of generate fake data](image)

Fig. 7. Time analysis of generate fake data

Fig. 8 shows the number of encryption and decryption of the GPS trajectories over VCC. The framework shows the efficiency of encrypting and decryption the user trajectories on VCC before requesting POI from the LBS. The proposed method will be to encrypt the user trajectories and store them in VCC storage, and decryption their location based on the users’ request. The time response to the decryption of the trajectory will be higher than that of the encryption part based on the total number of trajectories as shown in the figure.

![Fig. 8. Time analysis of encrypts & decrypts data](image)

Fig. 8. Time analysis of encrypts & decrypts data

Fig. 9 shows the number of generating token keys between the user and the VCC. The proposed framework shows the flexibility of generating the token of user authentication in the system when the users access the system. The token key will be generated on the server side based on using JWT-256HA to enhance the privacy of user login details when authenticated to the system. The signed token will send it to the user back through the browser HTTP/HTTPS. The Server-side will verify the validity of the attached token if the user browses the on the server. The lifetime of destroying the token will be at the time when the user login out from the system page.

![Fig. 9. The number of generating token keys](image)

Fig. 9. The number of generating token keys

V. CONCLUSIONS

With increasing the leaking of information of users over cloud computing nowadays, in this paper we have proposed a framework to register and identify different types of smart
IoT GPS devices over VCC. Also, we propose a framework TIIETGTS based on a method to anonymous the user information over VCC. The proposed method will replace the real user details with fake information, such as username, email, address, and phone number, and remove any linker to identify the original user information using SDM and DDM techniques. Proposed a method to provide location privacy of user trajectories by encrypting the location trajectory using the asymmetric RSA algorithm and storing the data in the VCC storage before requesting POI from the LBS. Where each row of data will be encrypted with different private and public keys. For secure communication between the client and server side, we proposed a method to generate token authentication of authorized users by hashing the user login details using the “JWT-HS256” method when requesting a service from VCC. The proposed framework shows the efficiency and reliability of responding to the user trajectories from different sources of IoT GPS devices and datasets.

REFERENCES


