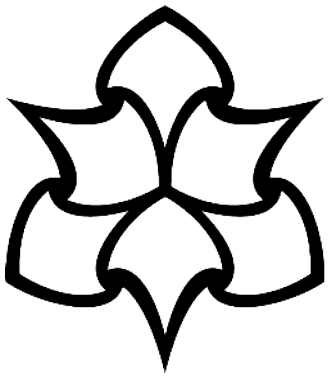

Energy-Efficient End-to-End Security for Software Defined Vehicular Networks

Gunasekaran Raja, *Senior Member, IEEE*, Sudha Anbalagan, Geetha Vijayaraghavan, Priyanka Dhanasekaran, Yasser D. Al-Otaibi and Ali Kashif Bashir, *Senior Member, IEEE*

***Abstract*—One of the most promising application areas of the Industrial Internet of Things (IIoT) is Vehicular Ad hoc NETworks (VANETs). VANETs are largely used by Intelligent Transportation Systems (ITS) to provide smart and safe road transport. To reduce the network burden, Software Defined Networks (SDNs) acts as a remote controller. Motivated by the need for greener IIoT solutions, this paper proposes an energy-efficient end-to-end security solution for Software Defined Vehicular Networks (SDVN). Besides SDN's flexible network management, network performance, and energy-efficient end-to-end security scheme plays a significant role in providing green IIoT services. Thus, the proposed SDVN provides lightweight end-to-end security. The end-to-end security objective is handled in two levels: i) In RSU-based Group Authentication (RGA) scheme, each vehicle in the RSU range receives a group id-key pair for secure communication and ii) In private-Collaborative Intrusion Detection System (p-CIDS), SDVN detects the potential intrusions inside the VANET architecture using collaborative learning that guarantees privacy through a fusion of differential privacy and homomorphic encryption schemes. The SDVN is simulated in NS2 & MATLAB, and results show increased energy efficiency with lower communication and storage overhead than existing frameworks. In addition, the p-CIDS detects the intruder with an accuracy of 96.81% in the SDVN.**

***Index Terms*—Green IIoT, Vehicular Ad hoc NETworks, Software Defined Networks, Energy Efficiency, Group Authentication, Differential Privacy, Homomorphic Encryption**

I. Introduction

The Industrial Internet of Things (IIoT) is a new ecosystem that combines intelligence fetched from the Internet of Things (IoT) devices to improve performance. One of the key application areas of IIoT is the Intelligent Transportation System (ITS) that relies on Vehicular Ad hoc Network (VANET) for improved road safety and driving assistance to their IoT users [1]. The adoption of VANET in ITS technologies will reduce the number of accidents to 1 million per year by 2020, with an economic benefit of $25.6 billion per year [2]. With an increased demand for ITS, the CO_2 footprint also increases. Current research focuses on green IoT, where energy-saving solutions are at the core of the design and development of the system. Thus, ITS requires integrated energy-efficient security features suited to its dynamic nature [3], [4].

Software Defined Network (SDN) is a fast-growing networking paradigm that allows flexibility and network configuration by separating data and control planes [5], [6]. The SDN based VANET architecture consists of vehicles with On-Board Unit (OBU), and RoadSide Unit (RSU). The RSUs are connected to the SDN controller and act as switches to obtain global network information. The SDN in VANET provides the following advantages: i) simplifies network management and ensures VANET elasticity [7], ii) global network knowledge from RSUs avoids periodic beacon messages among them. As a result, the network burden is substantially reduced and provides efficient routing decisions [8]. SDN integrated with Edge or Fog computing leverage the potential of pervasive technologies to provide several vehicular services such as location-based services, content sharing services, and so on.

Besides SDN's flexible network management, it is also essential to secure the vehicular network. A secure VANET encourages the participants to take part in it. Moreover, security in VANETs is of particular concern as human lives are frequently at risk [9]. Authentication acts as a primary defense mechanism to guarantee that a received message originates from an authenticated source [10]. Furthermore, the RSUs are capable of authenticating the vehicles in its range on the fly [11]. Still, VANETs are vulnerable to many kinds of attacks by a malicious insider node [12]. Some rogue vehicles after authentication pose as legitimate VANET users and send messages to interrupt the network communication. Thus, VANETs need reactive mechanisms such as Intrusion Detection Systems (IDS) in addition to the authentication mechanism to detect potential intruders [13]. Typically, a centralized IDS analyzes a dataset present in a central database, to search for an intrusion-related pattern. But such IDS are susceptible to performance bottlenecks, single-point failure, scalability issues, and often prone to data privacy risk [14]. To overcome the challenges faced by centralized IDS and to improve the classifier performance, a collaborative IDS is suitable for a dynamic network like VANETs.

Our contributions to the aforesaid problems in VANET are:

- In the proposed RSU-based Group Authentication (RGA) technique, the RSU provides a group ID and group key pair for each vehicle in its range to ensure further secure communication among vehicles with reduced network overhead.
- A private-Collaborative IDS (p-CIDS) is proposed to de-

Gunasekaran Raja, Geetha Vijaraghavan and Priyanka Dhanasekaran are with NGNLab, Department of Computer Technology, Anna University, Chennai, India. (e-mail: dr.r.gunasekaran@ieee.org; geethu15@gmail.com; priyankasekard2511@gmail.com).

Sudha Anbalagan is with Department of Computer Science and Engineering, SRM Institute of Science and Technology, Kattankulathur, India. (e-mail: sudhaa@srmist.edu.in)

Yasser D. Al-Otaibi is with Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah. (e-mail: yalotaibi@kau.edu.sa)

Ali Kashif Bashir is with the Department of Computing and Mathematics, Manchester Metropolitan University, UK. (e-mail: dr.alikashif.b@ieee.org).

tect potential attacks using a Collaborative Learning (CL) model. The p-CIDS in each vehicle learns collaboratively by co-ordinating with other vehicles.

The rest of this paper is structured, as section II discusses the literature related to VANET security solutions. The SDVN framework overview and its preliminaries are presented in section III. In section IV, the RGA mechanism is discussed. The p-CIDS based on CL is detailed in Section V. Section VI discusses the empirical findings and presents them in detail. Finally, the work is concluded in section VII.

II. Related Work

Numerous works study the importance of authentication in networking technologies such as LTE [15], VANETs [16], and so on. Threshold-based authentication [17] and Bilinear Pairing (BP) scheme [9] achieves anonymity, unforgeability, and revokes malicious node via traceability. A BP scheme and several trusted authorities were used to have a decentralized system [18]. Because of BP cryptography, the computational overhead is increased. In [16], securing group communication for SDN based 5G-VANET environment is focused but failed to support scalability. Elliptic Curve Cryptography (ECC)-based authentication, and Fuzzy C-means clustering for intrusion detection is used to prevent and detect the intruders in the network, respectively [19]. In [1], the El-Gamal signature is used, but such researchers face difficulty in cluster formation and inter-cluster communication.

There are a lot of recent research findings for the IDS system as security solutions in VANET. For example, the system in [13] uses a novel feature extraction technique, and the classification algorithm is based on improved growing hierarchical self-organizing map. A hierarchical growing neural gas network-based IDS is proposed in [12] that uses a semi-cooperative feature extraction algorithm, where the current location information is acquired from the neighboring vehicles in a co-operative fashion.

To secure the data and operate on ciphertext space, a homomorphic encryption method is discussed in [20]. Privacy preservation in Machine Learning (ML) is addressed using a differential privacy paradigm, which deals with adding a statistically-designed noise to the exchanged functions or states to protect the sensitive data [21]. Alternatively, in [22], a cryptographic image classification algorithm is proposed on a multi-layer extreme learning system that is capable of specifically classifying encrypted images without decryption.

The proposed SDVN provides energy-efficient authentication mechanisms and intrusion detection that makes the system more secure and robust against VANET cyber-security attacks.

III. SDVN Framework

Software Defined Vehicular Network (SDVN) framework provides end-to-end security and privacy using both proactive and reactive mechanisms in an energy-efficient manner. For proactive security, RGA authentication is designed as a lightweight authentication mechanism with reduced communication costs. The reactive security, p-CIDS, is a CIDS system that uses Collaborative Learning (CL), which reduces the storage cost of the CIDS.

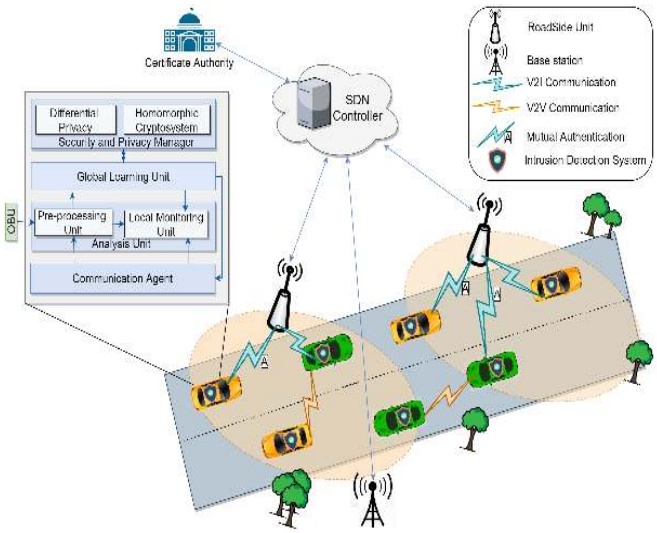

Fig. 1. System model of SDVN

The SDVN framework uses distributed SDN controllers with flat network topology, in which each of them is responsible for a specific segment such as the city. As shown in Fig. 1, the data plane constituted by the RSUs and BS is connected to the SDN controller and facilitates it with the global network information. The control plane provides policies like mobility management, authentication, routing, and so on. The flat design of the controllers enables them not only with a local and global view of the network but also reduces the complexity of the computation. The reliability of VANET can be achieved through the integration of security features into SDVN. The RGA scheme increases the trustworthiness of the network by ensuring messages received from an authenticated source. However, a semi-honest vehicle can perform malicious activities, which can be detected using CL based p-CIDS present in each vehicle.

The SDVN security architecture consists of Certificate Authority (CA), and VANET nodes (RSU & vehicles).

1) Certificate Authority: CA registers VANET nodes, stores identities, issues private credentials, and also capable enough to revoke certificates. CA is a fully trusted entity by VANET nodes and is interfaced with the control layer of the SDN controller. In the practical scenario, there will be 'n' number of CA available and each responsible for one particular region. Each VANET node should register exactly with one CA.

2) RoadSide Unit: RSU is a fixed infrastructure component in SDVN that is connected to the SDN controller. The controller is aware of interconnection among RSUs and transfers the logic to RSU for executing control layer commands. The security model is modeled as hierarchical trust, in which CA acts as a fully trusted entity, and RSUs act as the next level of trust. RSUs support and act as the upper level of trust for the vehicles in the network. The CA revokes the certificate of compromised RSUs to maintain trust.

3) Vehicles: Each vehicle is equipped with sensors and Tamper-Proof Devices (TPD). The TPD is capable of storing

cryptographic material generated by CA or RSU. Each vehicle in SDVN is registered with the CA before it joins the VANET. After successful registration, CA generates private credentials along with certificates, which are issued to each vehicle user in the network. The exchanged message format for V2V communication includes vital elements such as group id, pseudo id, payload, timestamp, and payload hash to achieve the integrity of the message.

In SDVN, vehicles engage in an authentication process, and if successful, it receives the group id-key pair for further network communication. Due to ECC-based cryptography, the attacker can't retrieve the secret key from the public key. To achieve message integrity, SDVN also uses key hashed functions. However, even after authentication, a vehicle can become semi-honest to the network through malicious activities like spreading misinformation. In such cases, the malicious activity is identified using p-CIDS and shared with the CA, which further revokes the credentials of the adversary. Thus, the SDVN framework prohibits the participation of malicious vehicles, thereby providing end-to-end security.

IV. RGA: RSU-BASED GROUP AUTHENTICATION

The source authentication by RGA act as a first-level defense mechanism in VANET. The SDN controller in SDVN runs the authentication module and responsible for controlling the network. The control plane provides general policies for authentication, mobility management, and routing. The CA provide the key for its registered users, deployed RSUs, and also capable of revoking the malicious node. In traditional, revoked identity is stored in the revocation list, which consumes storage space and increased searching time. To reduce the storage space and the searching time, CA constructs the Id revocation polynomial ($A_i^{i'}$) and key revocation polynomial ($K_i^{i'}$) using the identities and secret key of revoked vehicles in the SDVN network respectively. Besides, the proposed RGA process uses ECC-based authentication mechanism because of its fast computation and robustness to attacks. It achieves mutual authentication and confidentiality between the participants and also resists a reply attack.

A. RGA System Initialization

Let F_p be a finite field where p denotes the large prime number and the elliptic curve defined as $E : y^2 = x^3 + ax + b \ mod \ p$, $a, b \in Z_q^*$. The CA in SDVN selects the group G on E, where the order of the group as q and generator as g. CA construct the function $f(x, y) = b_0x + b_1y + c$, where $b_0, b_1, c \in Z_q^*$ are constants. After initializing these system parameters, CA generates public-private credentials for all the network entities. First, CA generates its public and private key based on these system parameters and performs ECC multiplication on a chosen random number. A similar procedure gets repeated to generate public and private keys for each network node, and also they get the unique Id. The private and public keys, as well as unique Id securely shared with their corresponding entities. Each node in the SDVN network is capable of holding TPD to store the details like private key, unique Id, etc.

Algorithm 1 RGA System Initialization Process

Input: Order of elliptic curve n, VID_i, RID_i
Output: Nodes' public-private key pairs, RSUs' member key

1: Set the public parameters $p,\ a,\ b$ and g
2: Generate CA public and private key $K = k.g$
3: **for** all vehicles and RSUs **do**
4: Allocate unique Id for vehicle as VId_i, RSU as RId_i
5: Generate vehicle pseudo Id as $h(VId_i||k||n_j)$, $j = 1, 2$
6: Generate secret key for vehicle (SV_i) & RSUs (SR_i)
7: Compute public keys $PV_i = SV_i.g$ and $PR_i = SR_i.g$
8: **end for**
9: Generate the member key for RSUs (MR)
10: Compute the public key $MP = MR.g$

In Algorithm 1, step 2 specifies the key generation for CA, and from step 4 to 7, CA generates the key pairs for vehicles and RSUs. The unique Id is obtained by hashing the corresponding id with the private key of the CA as $VId_n = h(VID_i||k)$ and $RId_i = h(RID_i||k)$ for vehicles and RSUs respectively. Step 9 and 10, specifies the member key generation for RSUs in the VANET to perform communication among them.

RSU Initialization: The RSUs in the SDVN network receives its own public and private credentials and stores it in TPD. To initialize the RGA scheme, RSU computes the unique group id using its private key $GId_j = h(RId_j||SR_j||r)$ where $r \in Z_q^*$. In addition, RSU selects the random number $d_1, t_1 \in Z_q^*$ and computes the backward hash chain of length n, as $t_{n-1} = h(t_1), t_n = h(t_{n-1})$, and $d_{n-1} = h(d_1), d_n = h(d_{n-1})$, where $n = 3$. To have secure communication among authenticated vehicles, each RSU in the SDVN network computes its own group key as $GK_j = h(d_1, t_n)$. CA generated Id revocation polynomial $A_i^{i'}(x) = (x - VId_1)(x - VId_2)...(x - VId_k)$, where $VId_1, VId_2, ..., VId_k$ are the identities of revoked vehicles. This revocation polynomial gets shared with all RSUs in the SDVN network to reject malicious vehicle at the time of authentication process.

B. RGA Authentication Process for Secure Communication

The vehicles in the SDVN network initiates the authentication process to get the group id and group key from the nearby RSU. With the help of group id-key pair, the vehicles ensure the authenticity of the received message. The detailed process of the proposed RGA mechanism is shown in Fig. 2.

Step 1: The vehicle V_i initiates the RGA authentication process by sending its VId_i and Loc_i to the nearby RSU_j.

Step 2: The RSU_j checks the Id revocation polynomial, if vehicle id is revoked then $A_i^{i'}(x) = 0$ and also checks the location is within the range, then generates the random number (R_1) by performing ECC multiplication on random number (r_1) and time stamp (TS_j), else tears down the connection. If valid, then R_1 and TS_j is relayed to V_i.

Step 3: The V_i which received the time-stamp TS_j checks the validity of time falls within the permissible range, further communication gets established, otherwise cut down the connection. The V_i computes the temporary token $T_1 = h(VId_i)$,

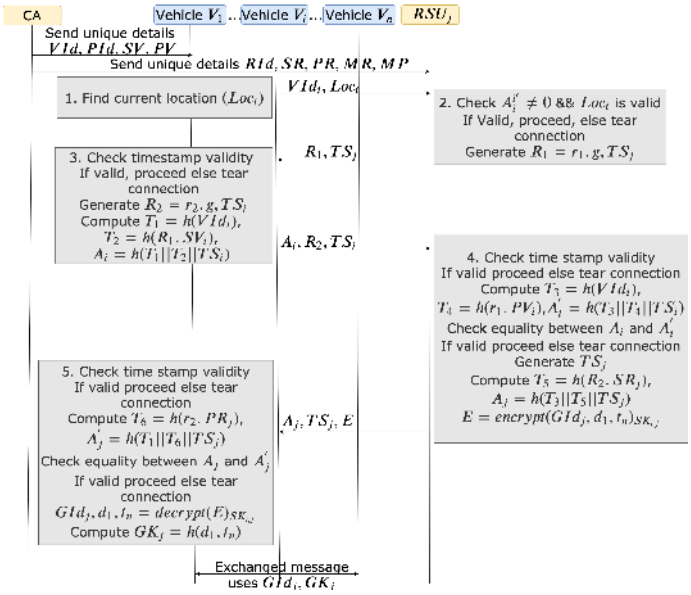

Fig. 2. RGA authentication process between RSU and vehicles

$T_2 = h(R_1.SV_i)$ and generate the time stamp TS_i, authentication token $A_i = h(T_1||T_2||TS_i)$ as well as the random number (R_2). The Authentication token A_i, R_2 and TS_i gets relayed to RSU_j.

Step 4: The RSU_j checks the validity of time, if valid accepts the message otherwise breaks the connection. The RSU_j computes the token $T_3 = h(VId_i)$, $T_4 = h(r_1.PV_i)$ and $A_i^{'} = h(T_3||T_4||TS_i)$. If the equivalence of A_i and $A_i^{'}$ is true, proceed; otherwise break the connection. The RSU_j computes the temporary token $T_5 = h(R_2.SR_j)$, generate the time stamp TS_j and Authentication token $A_j = h(T_3||T_5||TS_j)$. The RSU_j computes the session key using the public key of the vehicle and private key of the RSU $SK_{i,j} = PV_i.SR_j$. Then encrypts the group id and elements of group key pair (d_1, t_n) using session key $(SK_{i,j})$, which makes the shared information secure and confidential. The Authentication token A_j, TS_j and encrypted message gets relayed to V_i.

Step 5: The V_i checks the validity of time, if valid accepts the message otherwise breaks the connection. The V_i computes the token $T_6 = h(r_2.PR_j)$ and $A_j^{'} = h(T_1||T_6||TS_j)$. If the equivalence of A_j and $A_j^{'}$ is true, then proceed; otherwise break the connection. After mutual authentication, V_i decrypts the message which consists of the group id and elements of group key of RSU_j using session key $SK_{i,j} = SV_i.PR_j$. The vehicle computes the group key of RSU_j as $GK_j = h(d_1, t_n)$ and stores it in TPD.

The authenticated vehicles have the group id and group key to perform further V2V communication in the SDVN network. The V2V communication uses group id (GId_j), group key (GK_j) as follows: the vehicles send the message with group id GId_j, pseudo id PId_i, Message M, and hash the message using group key $h(M)$, the received vehicle checks the GId_j and if it is valid accept the message otherwise rejects the message. The vehicle also checks for integrity $h^{'}(M)$, if $h(M)$ equals $h^{'}(M)$, which ensures that received message is not modified by others. Further, confidentiality is achieved by encrypting the message, which consists of PId_i, M, and $h(M)$, using the group key. If malicious activity found, the report is sent to nearby RSU along with the PId_i and GId_j. The CA has the ability to open the true vehicle id and punish the node by invalidating the certificate. The CA includes the misbehaved vehicle id and secret key into the Id revocation polynomial and key revocation polynomial, respectively. Then sends the updated Id revocation polynomial to all RSUs and key revocation polynomial to reported group id of RSU. The reported RSU in SDVN creates new group id-key pair and also updates its previous group key to protect the group from a rogue vehicle. The following subsection describes the previous group key update process of RGA.

C. RGA Group Key Update Process

The RSUs of SDVN network receives the key revocation polynomial $K_i^{i'}(x)$ from CA initializes the RGA group key update process, where $K_i^{i'}(x) = (x - SV_1)(x - SV_2)...(x - SV_k)$, where $SV_1, SV_2, ..., SV_k$ are the secret key of revoked vehicles. Assume RSU_j receives the key revocation polynomial, then it computes the updated group key as $GK_j = h(d_2, t_2)$ and a masking polynomial $\varphi(x, y) = K_i^{i'}(x).t_{n-1} + d_{n-1}.f(x, y)$. The RSU_j sends the reported group id, key revocation and masking polynomial to the base station in the SDVN network, which further broadcast the received message to the vehicles in its range. The vehicle which receives the message checks the group id, if it valid then it under goes the group key update process. The vehicle V_i updates it key by computing $d_{n-1} = h(d_1), K_i^{i'}(SV_i), \varphi(SV_i, d_{n-1})$ and $f(SV_i, d_{n-1})$. Then computes $t_{n-1} = \frac{\varphi(SV_i, d_{n-1}) - d_{n-1}.f(SV_i, d_{n-1})}{K_i^{i'}(SV_i)}$. The vehicle also verifies the $t_n = h(t_{n-1})$, if it is valid then vehicle V_i computes the updated group key as $GK_j = h(d_{n-1}, t_{n-1})$. If the vehicle is revoked, $K_i^{i'}(SV_i) = 0$, and the rogue vehicle cannot get the updated key.

V. Private-Collaborative IDS (p-CIDS)

The RGA authentication process prevents the outside attackers, but some rouge users perform attack after joining the SDVN network. To detect such intruders in the dynamic VANET environment, p-CIDS is developed, as shown in Fig. 1. In p-CIDS, the vehicles share their knowledge to its nearby users with the help of Distributed Machine Learning (DML) techniques. When vehicles collaborate, DML's model parameters are shared to generate the global model. The designed CIDS works collaboratively to detect attacks in the VANET environment, such as sybil attack, wormhole, blackhole, denial of service attacks, and so on.

For DML, each vehicle in SDVN has a partitioned dataset and performs the ML steps to obtain their loss function. The global model is obtained by minimizing the sum of loss function of all the vehicles in the SDVN network. Alternating Direction Method of Multipliers (ADMM) is used for CL, as ADMM convergence is more rapid with a standard convergence rate of $O(1/t)$ [23], [24].

A. ADMM based Collaborative Learning Problem

Let us consider a VANET network, which consists of $\boldsymbol{N}$ vehicles which can be represented as a undirected Graph $\mathcal{G}(\mathbb{N}, \mathbb{E})$. $\mathbb{N} = \{1, 2, 3, ..., \boldsymbol{N}\}$ represents the number of vehicles in the VANET and $\mathbb{E}$ represents the set of edges connecting the vehicles. A vehicle $m \in \mathbb{N}$ can exchange information only with its neighbour $i \in \mathscr{N}_m$, where $\mathscr{N}_m$ is the set of neighbouring vehicles to vehicle m and N_l is the total number of neighbouring vehicles to vehicle m. Each vehicle m contains a dataset $D_m = \{(x_{jm}, y_{jm}) \subset X \times Y : j = 0, 1, ..., R_m\}$, in which R_m is the training data size containing data instances $x_{jm} \in X \subseteq \mathcal{R}^d$, where d refers to dimensional vector space of the instances and corresponding label $y_{jm} = \{0, 1\}$. The total dataset of the entire network is thus, $\hat{D} = \cup_{l \in \mathbb{N}} D_m$. Let us consider the Empirical Risk Minimization (ERM) problem for a regularized binary classification as follows:

$$\min_{f_m} Z_{ERM}(f_m, \hat{D}) = \sum_{i=1}^{\boldsymbol{N}} \frac{C}{R_m} \sum_{n=1}^{N_m} \mathscr{L}\left(y_i^n f_n^T x_i^n\right) + \tau R(f_m) \tag{1}$$

Here $C \leq R_m$ and τ are constants; C is called regularization parameter, τ controls the effect of regularization. The Loss function $\mathscr{L}(.)$ is a measure of the classifier accuracy. The function $R(.)$ assists in mitigating the over-fitting problem. Thus the goal is to learn a global classifier f_m over the total training dataset $\hat{D}$ in a distributed fashion using a ADMM and also provides privacy assurance to each data sample.

To apply ADMM-based distributed learning algorithm, the objective function is reformulated to be solved for collaborative nodes and solved using ADMM using the following equations [25]:

$$\begin{aligned} f_m(t+1) \;=\; & arg \min_{f_m} \{Z(f_m, D_m) + 2(\lambda_m(t))^T f_m \\ & + \eta \sum_{i \,\in\, \mathscr{N}_m} \frac{1}{2} \| (f_m(t) + f_i(t)) - \; f_m \|^2 \} \end{aligned} \tag{2}$$

$$\lambda_m(t+1) = \lambda_m(t) + \frac{\eta}{2} \sum_{i \,\in\, \mathscr{N}_m} (f_m(t+1) - f_i(t+1)) \tag{3}$$

where f_m and λ_m are called global classifier and dual variable of the ADMM algorithm, respectively. Each vehicle in the SDVN network runs the CL algorithm to detect the intruders. The IDS in each vehicle uses a pre-processed dataset and runs the CL algorithm, which minimizes the loss function of f_m. If neighbor available in its coverage, then it broadcast the f_m. Each vehicle that receives f_m undergoes the operation in Eq. (2) and Eq. (3) to compute the global classifier model by iterating for a particular threshold of iterations (50 iterations). The updated classifier f_m is used to predict the future traffic data instances in the SDVN network.

B. Secure and Private Collaborative Learning

The CL methodology results in a scalable IDS for the SDVN, which best suits the dynamic VANET system. The CL is secure than centralized learning as training data is not shared directly, but the classifier only is shared. CL is energy-efficient than centralized learning because of the reduced storage overhead and computational efficiency. Yet this private solution results in privacy leakage if an adversary can gather information through statistical inferences of the data. Simple anonymization techniques are not sufficient to provide a barrier against such privacy leakages. It is, therefore, becomes a need to protect the SDVN system from such privacy leakages. We formally describe our Secure and Private CL (SPCL) with the notion of differential privacy and homomorphic cryptosystem as a means of guaranteeing privacy against inference attacks. The p-CIDS unit in each vehicle has the following components: communication agent, analysis unit, Global Learning Unit (GLU), security and privacy manager. The analysis unit consists of pre-processing unit and Local Monitoring Unit (LMU). The security and privacy manager consists of Differential Privacy (DP) and Homomorphic Cryptosystem (HC).

Algorithm 2 Secure and Private Collaborative Learning

Initial Stage: Each vehicle has $f_j^0, \lambda_j^0 = 0, j \in N$
Input:: Network traffic data **Output:** updated classifier

1: **if** traffic data received **then**
2: Pre-process (Collected traffic data)
3: **for** $t = 0, 1, 2..., Th$ **do**
4: **for** $m = 1, 2, ..., N$ **do**
5: Compute the classifier $f_m(t+1)$ using Eq. (2)
6: Vehicle m encrypts f_m^{t+1} with member public key M_{rp} and adds laplacian noise e:
$$f_m^{t+1} \rightarrow E(-f_m^{t+1}) + E(e)$$
7: Vehicle m broadcasts f_m^{t+1} to its neighbor i
8: Compute $\lambda_m(t+1)$ using Eq. (3)
9: **end for**
10: **end for**
11: **end if**
12: The classifier f_m is sent to RSU
13: At RSU, decrypt the classifier and send the classifier to the vehicle

Each vehicle monitors the network traffic and application traces in the vehicle using LMU, which consists of a classifier. The LMU will generate an alert if the classifier indicates an intrusion. The classifier in the LMU can be updated using CL in GLU. Once initialized, the GLU runs the SPCL algorithm and updates the current classifier in LMU. The SPCL algorithm uses DP and HC components of the security and privacy manager for securing the training data used in the CL. Any communication for CL uses the communication agent of the IDS.

1) Homomorphic Cryptosystem: If a system uses a separate key for encryption and decryption, then it is called a public key cryptosystem. In SDVN, public key (PH_k) is used for encryption and a secret key (SH_k) is used for decryption process [20].

Definition 1: Homomorphic Cryptosystem

A public key cryptosystem (Gen, Enc, Dec) is known as homomorphic if for all message x in plain text space $\mathbb{P}$ with encryption/decryption key pair (PH_k, SH_k), it is possible to define groups $\mathbb{P}$, $\mathbb{C}$ such that:

(i) The message in plain text space $\mathbb{P}$, and all ciphertexts

output by encryption algorithm are elements of cipher text space $\mathbb{C}$.

(ii) For any $p1, p2 \in \mathbb{P}$ and their corresponding cipher texts c1,c2 $\in \mathbb{C}$, it should satisfy the following criteria: $Dec_{SH_k}(c1.c2) = p1 + p2$.

The multiplication of cipher text $(c1, c2)$ is equivalent to the cipher text obtained by encrypting sum of $p1$ and $p2$.

2) Differential Privacy:

Definition 2: Neighboring Dataset

The datasets D and D' have the same symmetry and attribute structure, which is denoted as $\mid D \Delta D' \mid$. We call D and D' neighbour datasets if and only if: $\mid D \Delta D' \mid = 1$.

Definition 3: ϵ-Differential Privacy

A randomized mechanism S gives ϵ-DP for every set of outputs X, and for any neighbor data set of D and D', if S satisfies the following condition: $Pr[S(D) \in X] \leq exp(\epsilon) \times Pr[S(D') \in X]$.

Definition 4: Laplace - Differential Privacy Mechanism

For a dataset D and a function $f : X \rightarrow Y$, a privacy mechanism $M(D) = f(D) + e$ provides $\epsilon-$DP where $e \in \mathbb{R}$ has Lap(σ) distribution, if its density function is given by $\frac{1}{2\sigma} exp(-|x|/\sigma)$.

The CL algorithm requires the vehicles to collaborate and disclose intermediate classifiers in each iteration with the neighboring vehicles to reach an agreement on an optimal final classifier. The security and privacy manager provides DP and HC functions to provide privacy preservation to the CL of the CIDS used in the SDVN. SPCL is based on a paillier homomorphic cryptosystem [20], which is public key cryptography with a pair of keys, namely public key and private key. The reliability of these systems is bound to the hardness of solving the factorization problem.

The privacy preservation mechanism is applied to CL using a combination of DP and HC, as shown in Algorithm 2. In Algorithm 2, the broadcast classifier is encrypted before sending it to other vehicles with encryption function E(.). A laplacian noise is added to the encrypted classifier and broadcasted to neighboring vehicle i. Thus, CL happens in the ciphertext space and also perturbed using a laplacian noise to apply the DP paradigm. At any vehicle i, the dual variable $\lambda_m(t+1)$ is calculated from Eq. (3) from the received classifier. Once the number of iterations reaches its threshold, the classifier is sent to RSU of SDVN network, which decrypts the model and sends it back to each vehicle. The LMU uses the received classifier for future predictions.

VI. Implementation and Results

The proposed mechanisms are simulated in NS2 and MATLAB on a four-core 3.2 GHZ machine with an 8 GB RAM. For the experimental analysis of p-CIDS, the NSL-KDD dataset is used [26]. In this section, we present the experimental analysis of the authentication mechanism, security, and privacy of training data in collaborative IDS.

A. Security and overhead analysis of RGA

In this subsection, we highlight the Security Features (SF) of the RGA mechanism in the SDVN framework towards VANET cyber-security attacks.

TABLE I
Comparative analysis of security features

Scheme	SF1	SF2	SF3	SF4	SF5	SF6	SF7	SF8
Zhong et al [9]	×	✓	✓	✓	×	✓	✓	×
Azees et al [18]	×	✓	✓	×	✓	✓	✓	✓
Dua et al [1]	✓	×	✓	✓	×	✓	✓	✓
Sahil et al [19]	✓	✓	✓	✓	✓	×	×	✓
RGA scheme	✓	✓	✓	✓	✓	✓	✓	✓

1) Support mutual authentication (SF1): In the RGA method, authentication occurs between vehicles and RSU, ensures the authenticity of participating vehicles by means of the authentication token. To create the token, ECC multiplication is performed on random number with the private key, thereby ensuring that only authenticated vehicles have a genuine private key. In ECC, extracting the private key from public key is impossible.

2) Resist eavesdropping (SF2): The use of the random number, timestamps, and private key in the authentication process avoids eavesdropping attacks in VANET. The attacker can't extract exchanged messages because of the above attributes as well as group id and key pair, which is encrypted using the shared key.

3) Support anonymity (SF3): Anonymity accomplished through fresh tokens during the authentication process. In each run between RSU and vehicles, a fresh token is generated by means of a random number, time stamp, location, and ECC-based private key.

4) Resist replay attack (SF4): The timestamp attribute in the authentication process resists the VANET replay attack. The participating nodes in the SDVN network drop the delayed transmitted messages. The replay attack is avoided because of a random number in the authentication token for each run.

5) Resist spoofing attack (SF5): The attacker cannot spoof the identity of CA, RSU, and vehicles in the VANET because the token is generated by the private key of the nodes. Under the ECC cryptosystem, it is impossible to find a private key from a public key.

6) Support message authentication (SF6): After authentication, group id and key pair have been used to indicate the authenticity of the message. The vehicle checks the group id and integrity of the message using the group key. The revoked vehicles can't get this pair to participate in the SDVN network.

7) Resist man-in-the-middle attack (SF7): The authentication token created by the respective private and public credentials of the participants in the authentication process. The group id-key pair gets encrypted using the session key their genuine participants generate. The intermediate nodes, therefore, can not be able to forge them.

8) Support forward secrecy (SF8): The RGA authentication and group update process support forward secrecy, which uses a random number, and timestamp attribute to enhance the security of the SDVN framework. The attacker cannot retrieve the previously exchanged messages, even though they aware of the current system information. Table I summarizes the comparative analysis of the security features provided by the proposed RGA system with the current state-of-the-art VANET security mechanism. From Table I, it is inferred that the RGA

TABLE II
COMPARATIVE ANALYSIS OF OVERHEADS IN AUTHENTICATION PROCESS

Scheme	Communication cost (in bits)	Computation cost (in seconds)	Number of exchanged messages
Zhong et al [9]	$823^{*}, 832n^{**}$	$0.0171n + 0.1197^{**}$	1
Azees et al [18]	7488	0.1302!, $(n+1)0.0171 + (n+4)0.0192$!!	1
Dua et al [1]	2144	0.1406	3
Sahil et al [19]	1568	0.1061	4
RGA scheme	1632	0.117	4

where *: single message verification, **: batch verification of messages, !:single certificate and signature, !!: n certificate and signature.

scheme performs better in terms of resistance to cyber-security attacks and also provides supportive security features.

B. Comparative analysis

In this subsection, the cost of computing and transmitting the RGA scheme is compared with other schemes in the authentication process of VANET. Vehicle registration is a one-time operation in the RGA scheme, and thus, the cost of authentication between vehicles and RSU is considered. The RGA process takes advantage of ECC cryptography and a one-way hash function. Our scheme uses 160-bit ECC, which is equal to a 1024-bit RSA cryptosystem and 160-bit output Secure Hash Algorithm (SHA-1).

1) Comuptation cost: The computation time in seconds for 160-bit ECC multiplication (T_{eccm}), one-way hash function (T_h), and symmetric encryption (T_{senc}) or decryption (T_{sdec}) is 0.0171, 0.00032, and 0.0056 respectively. The RGA authentication process between vehicles and RSU requires the total computation cost of $6T_{eccm} + 10T_h + T_{senc} + T_{sdec} \approx 0.117$.

2) Communication cost: Assume that the number of bits used to represent identity, timestamp, location and hash output as 160, 32, 32 and 160 respectively. In the RGA authentication process, the exchanged messages in bits are $(VId_i, Loc_i), (R_1, TS_j), (A_i, R_2, TS_i)$ and $(A_j, TS_j, enc(GId_j, d_1, t_n))$, which needs (160+32) = 192, (320+32) = 352, (160+320+32) = 512 and (160+32+128+128+128) = 576 respectively. The total communication cost of RGA authentication process is 1632 bits. Table II presents the comparative overhead analysis with the existing authentication process, which makes use of ECC [1], [19], and bilinear pairing [9], [18] as a security mechanism in VANET. From Table II, it is inferred that the RGA scheme requires a bit more computational and communication overhead compared with [19] but achieves significant security features, which best suits the dynamic VANET environment.

C. Security and Privacy analysis of p-CIDS

Logistic regression in ML is used to predict the probability of the occurrence of an event by fitting a logistic function. For the SPCL, we assume a binary logistic regression, where the output variable is one of two possible classes 0,1. The Logistic regression algorithm aims to find the optimal parameters by

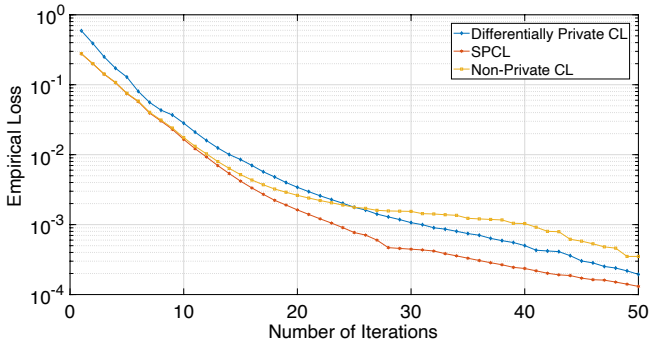

Fig. 3. Error Vs Number of Iterations in CIDS

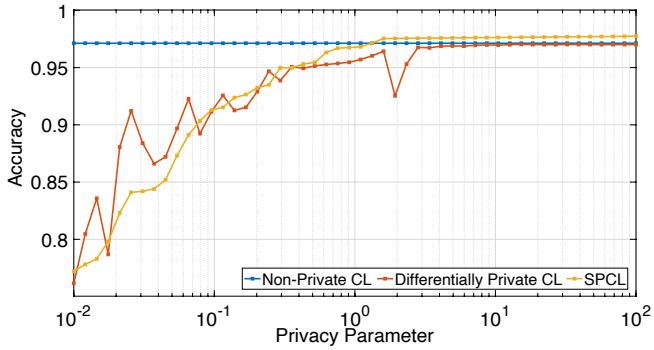

Fig. 4. Privacy Parameter Vs Accuracy of CL in CIDS

minimizing a loss function. The loss function $J(\theta)$ of the logistic regression is as follows

$$J(\theta) = \frac{1}{n}\sum_{i=1}^{n} log\left(1 + e^{(-y\theta^T x)}\right) \quad (4)$$

The p-CIDS is simulated with a penalty parameter (τ) to 0.1. The parameter for regularization is selected as 10^{-6} with standard 10-cross validation. The simulation uses the empirical risk function of logistic regression. As Fig. 3 shows, in SPCL methodology, the empirical risk is close to the non-private CL technique and performs much better than the differentially private CL as well. As shown in Fig. 4, with increasing privacy parameter ϵ, the accuracy also increases. But as the accuracy increases, it is statistically easier to infer the data from the intermediate states shared among the vehicles. With decreasing ϵ, the noise in the results increases, which results in increased security but potentially degrades the utility of the model. Thus there should be a balance between the security and accuracy of the model, which is achieved through the ϵ parameter. Therefore, the privacy parameter for SPCL learning, ϵ is set as 1. The p-CIDS is evaluated using the metrics: precision, recall, F1-score, and cross-validation score. The detection accuracy of the p-CIDS system is at 96.81% as seen in Fig. 5, and it also summarizes the evaluation scores of various learning models.

VII. CONCLUSION

In the proposed SDVN framework, the security of the VANET achieves energy efficiency towards green IIoT using

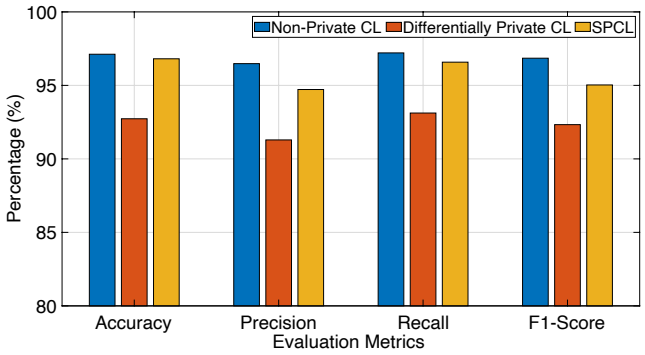

Fig. 5. Evaluation of Collaborative Learning Models

two strategic levels. In the first level, pre-trusted RSUs provide an RGA authentication mechanism to the vehicles in the VANET. The RGA schemes reduce the network overhead with increased security by preventing several attacks. The vehicle enters into the VANET by receiving a group id-key pair through RGA authentication. But, some semi-honest vehicles provide misinformation or drop the packets, such vehicles become a network threat. The SPCL based p-CIDS is used to identify these semi-honest vehicles in the network and report to CA via RSU. In effect, the CA will remove the vehicle from the green IIoT network, and revocation is achieved using the RGA group key update process.

ACKNOWLEDGEMENT

This Publication is an outcome of the R&D work undertaken in the project under the Visvesvaraya PhD Scheme of Ministry of Electronics & Information Technology, Government of India, being implemented by Digital India Corporation (formerly Media Lab Asia).

REFERENCES

[1] A. Dua, N. Kumar, A. K. Das, and W. Susilo, "Secure message communication protocol among vehicles in smart city," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 5, pp. 4359–4373, 2018.

[2] L. A. Maglaras, "A novel distributed intrusion detection system for vehicular ad hoc networks," *International Journal of Advanced Computer Science and Applications*, vol. 6, no. 4, pp. 101–106, 2015.

[3] R. Arul, G. Raja, A. O. Almagrabi, M. S. Alkatheiri, S. H. Chauhdary, and A. K. Bashir, "A quantum-safe key hierarchy and dynamic security association for LTE/SAE in 5g scenario," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 1, pp. 681–690, 2020.

[4] M. Hayes and T. Omar, "End to end VANET/ IoT communications a 5g smart cities case study approach," in *IEEE International Symposium on Technologies for Homeland Security (HST)*, pp. 1–5, 2019.

[5] G. Raja, A. Ganapathisubramaniyan, S. Anbalagan, S. B. M. Baskaran, K. Raja, and A. K. Bashir, "Intelligent reward-based data offloading in next-generation vehicular networks," *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 3747–3758, 2020.

[6] S. Anbalagan, D. Kumar, G. Raja, and A. Balaji, "SDN assisted stackelberg game model for LTE-WiFi offloading in 5g networks," *Elsevier-Digital Communication & Networks*, vol. 5, pp. 268–275, 2019.

[7] S. Anbalagan, D. Kumar, G. Raja, W. Ejaz, A. K. Bashir, *et al.*, "SDN-assisted efficient LTE-WiFi aggregation in next generation iot networks," *Future Generation Computer Systems*, 2017.

[8] S. Anbalagan, D. Kumar, D. Ghosal, G. Raja, and V. Muthuvalliammai, "SDN-assisted learning approach for data offloading in 5g hetnets," *Mobile Networks and Applications*, vol. 22, no. 4, pp. 771–782, 2017.

[9] H. Zhong, J. Wen, J. Cui, and S. Zhang, "Efficient conditional privacy-preserving and authentication scheme for secure service provision in VANET," *Tsinghua Science and Technology*, vol. 21, no. 6, pp. 620–629, 2016.

[10] S. B. M. Baskaran, G. Raja, A. K. Bashir, and M. Murata, "QoS-aware frequency-based 4G+relative authentication model for next generation LTE and its dependent public safety networks," *IEEE Access*, vol. 5, pp. 21977–21991, 2017.

[11] J. Li, H. Lu, and M. Guizani, "ACPN: a novel authentication framework with conditional privacy-preservation and non-repudiation for vanets," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 4, pp. 938–948, 2015.

[12] A. Ayoob, G. Su, and G. Al, "Hierarchical Growing Neural Gas Network (HGNG)-Based Semicooperative Feature Classifier for IDS in Vehicular Ad Hoc Network," *Journal of Sensor and Actuator Networks*, vol. 7, p. 41, 2018.

[13] J. Liang, J. Chen, Y. Zhu, and R. Yu, "A novel intrusion detection system for vehicular ad hoc networks (VANETs) based on differences of traffic flow & position," *Applied Soft Computing*, vol. 75, pp. 712 – 727, 2019.

[14] S. Han, U. Topcu, and G. J. Pappas, "Differentially private distributed constrained optimization," *IEEE Transactions on Automatic Control*, vol. 62, no. 1, pp. 50–64, 2017.

[15] R. Arul, G. Raja, A. K. Bashir, J. Chaudry, and A. Ali, "A console GRID leveraged authentication and key agreement mechanism for LTE/sae," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 6, pp. 2677–2689, 2018.

[16] C. Lai, H. Zhou, N. Cheng, and X. S. Shen, "Secure group communications in vehicular networks: A software-defined network-enabled architecture and solution," *IEEE Vehicular Technology Magazine*, vol. 12, no. 4, pp. 40–49, 2017.

[17] J. Shao, X. Lin, R. Lu, and C. Zuo, "A threshold anonymous authentication protocol for vanets," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 3, pp. 1711–1720, 2016.

[18] M. Azees, P. Vijayakumar, and L. J. Deboarh, "EAAP: efficient anonymous authentication with conditional privacy-preserving scheme for vehicular ad hoc networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 9, pp. 2467–2476, 2017.

[19] S. Garg, K. Kaur, G. Kaddoum, S. H. Ahmed, and D. N. K. Jayakody, "SDN-based secure and privacy-preserving scheme for vehicular networks: A 5g perspective," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 9, pp. 8421–8434, 2019.

[20] P. Paillier, "Public-key cryptosystems based on composite degree residuosity classes," in *International conference on the theory and applications of cryptographic techniques*, pp. 223–238, Springer, 1999.

[21] E. Nozari, P. Tallapragada, and J. Cortés, "Differentially private distributed convex optimization via functional perturbation," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 1, pp. 395–408, 2018.

[22] H. Zhu, X. Liu, R. Lu, and H. Li, "Efficient and privacy-preserving online medical prediagnosis framework using nonlinear SVM," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 3, pp. 838–850, 2017.

[23] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via alternating direction method of multipliers," *Foundations & Trends in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.

[24] Z. Huang, R. Hu, Y. Guo, E. Chan-Tin, and Y. Gong, "DP-ADMM: ADMM-Based Distributed Learning With Differential Privacy," *IEEE Transactions on Information Forensics & Security*, vol. 15, pp. 1002–1012, 2020.

[25] P. A. Forero, A. Cano, and G. B. Giannakis, "Consensus-based distributed support vector machines," *Journal of Machine Learning Research*, vol. 11, no. May, pp. 1663–1707, 2010.

[26] *NSL-KDD data set for network-based intrusion detection systems,* Available on: http://nsl.cs.unb.ca/KDD/NSLKDD.html, March 2009.