ENSURING SECURITY AND PRIVACY IN VEHICULAR EDGE COMPUTING WITHBLOCKCHAIN AND DEEP DETERMINISTIC POLICY GRADIENT

Neelima Sahu[1]
Asst. Professor, East Point College of Engineering and Technology, Bangalore, Karnataka, India
Neelimasahu73@gmail.com

Vaishali Sontakke [2]
Asst. Professor, East Point College of Engineering and Technology, Bangalore, Karnataka, India
vaish.sontakke@gmail.com

Correspond to vaish.sontakke@gmail.com

ABSTRACT:

The accelerated surge in the installation of Internet of things devices in smart vehicles leads to an outburst of the massive quantity of data and bandwidth consumption which finally ends in network congestion. Vehicular Edge Computing (VEC) is a proactive thought pattern to supervise an enormous supplyof data and media-rich content that can be cached between adjacent vehicles. On the other hand, moving vehicles with their wireless domain thrives to have a design with optimal content caching policy. The deep Deterministic Policy Gradient (DDPG) technique comes forth to settle the issue with vast area coverage and time-altering characteristics. To secure a peer2peer transaction network the most preferred choice is Permission Blockchain. Integrated DDPG and Permission Blockchain give a perfect solution to manage the security and privacy of smart content caching.

DOI Number: 10.14704/nq.2022.20.9.NQ44700 **Neuro Quantology 2022; 20(9):6009-6020**

I. Introduction: A powerful unit of future smart transportation systems in vehicular networks. Various mobility services are allowed by the networks, which range from content sharing services to information broadcasting operations. With the massive advancement of touch screen stereos, Onscreen GPS navigation displays give rise to the bulk of data and content in moving automobiles. Because of more distance between cloud servers and vehicles, restricted backhaul link capacity creates significant challenges for favoring massive content delivery while also satisfying the low-latency requirement in vehicular networks [1]. mobile traffic on backhaul links can be significantly reduced with the help of cache content at edge servers.[9].The authors in [10] suggested Femto-node base stations diminish the expected delay in content dispatch. Continuous moving of autonomous vehicles which passes through cross communication topologies gives a new challenge to model an appropriate cache content approach in the wireless channel [13]. Furthermore, content generally is hypersensitive data for the owner.

A.REINFORCEMENT LEARNING

Like a small child depends on learning from communication, imitated actions and corrections with small adaptations, and discovering what works and whatnot. Let's think of a child learning to walk. It tries out various imaginable actions for many days to stand on its feet and then be able to walk. During this course of actions, the child is punished by falling and rewarded for improving in its walk [8]. This rewarding scheme is intrinsically built into animals that spark themselves to do actions that

reap convinced rewards and avoid the behaviors that end with bad rewards (e.g., sad, pained, etc.). This sort of reward-based learning has been designed mathematically that is reinforcement learning. A classic reinforcement learning setting comprises an agent

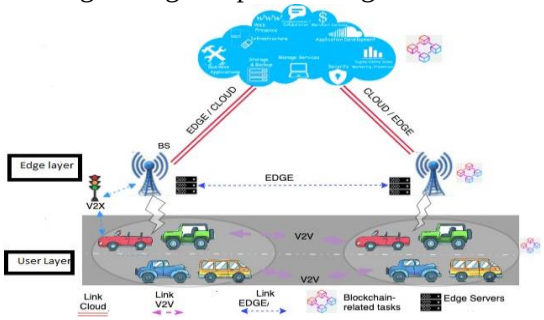

Fig.1: Vehicular Content Caching authorized with Blockchain

Interacting with the domain. After every singles phase p, the agent earns an experience e_p from the domain and takes an action a_p. The agent gets a reward r_p and the next state s_{p+1}.

B.DEEP REINFORCEMENT LEARNING (DRL)

DRL has transformed the field of artificial intelligence (AI) and produce autonomous systems with the optimal perception of the visual world. DRL is the integration of artificial neural networks and reinforcement learning that guides software agents to learn how to achieve revery goal state with corresponding rewards. It combines function approximation, target optimization, scaling states, and total rewards.

DRL can discover the feasible and available content cacheable resources at the edges of networks and blueprints correlated model by collaborating with edge servers [12]. In [16] they have drafted a course of action based on reinforcement learning accompanied by an actor-critic model for sharing of content cache, allocation of resources, and offloading of computations. In the process of selecting an ideal edge server in the network of vehicles, a deep learning method called Q-learning is used for task offloading. Despite the above works failed to ensure security and privacy.

C.A DEEP DETERMINISTIC POLICY GRADIENT (DDPG) APPROACH

The DDPG guideline planned by Lillicrap [2] may be a classic deterministic policy gradient approach supported by Actor-Critic [3] design. This brings out the thought of experimental replay and also the disjunction of the destination network from DQN, and also the distinct behavior area into the continual area. None the less the majority of the present reinforcement learning ways are not applicable throughout the coaching critic module which is controlled by a centralized node. This centralized module administers agents with information concerning the unending observance and probable behavior of all agents, and the intercommunication amid agents converts an unplanned scenario to a certain possible scenario. Once corroboratory the agent cannot see all the facts, and it doesn't understand the behavior of the opposite agents; he uses solely the facts advised by his private critic module to do the required action alone. Henceforth, as related to alternative DRL ways, DDPG is appropriate for designing shared pathways for crowded areas. The simplified picture of the DDPG rule isdepicted in Fig. 3. It comprises the net chunk of every actor and critic.

D. BLOCKCHAIN TECHNOLOGY

Blockchain is an ingenious open database technology that holds transactional records while ensuring security, transparency, and decentralization. [15] Blockchain is a distributed technology that resolves all the limitations of centralization. Given target end-users Blockchain can be divided into categories (i) public Blockchain which is fixed for the general public as end-users and is completely independent of organizations. Here the main disadvantage is network can be very slow and organizations can't restrict access and use. If attackers capture 51 % or more of the computing power of a public chain network they can single-handedly alter it. (ii) Permission Blockchain is geared up towards invitation-only end users. An organization that builds up a permission blockchain network can dictate the visibility of nodes, add or change data. It can also restrict third parties from retrieving certain information. Because they're comparatively small in scope, permission blockchains are very agile and can perform transactions more promptly compared to public blockchains. (iii) Consortium

Blockchain: In this more than one organization shares a common blockchain platform. (iv)Hybrid Blockchain: Here when organizations require the best of both worlds, a hybrid blockchain helps them to set up a Private-permission system in addition to public-permissionless system, restricts who can retrieve distinct information available in the blockchain, and which is available for public. Generally, the content of hybrid blockchain is not are not available in public, However can be verified when the necessity arises, like when retrieving the content of a smart contract.

In this research, we planned to model a smart and safe caching of contents among nearby vehicles. But we may encounter two types of threats. The first kind of threat that may complicate the best approach is the continuous movement of vehicles gives rise to frequent change in the topology of networks with a variation of time. Another threat can be if we aim for minimum expenditure of energy to model a permission blockchain leads to prolongation of time. To eliminate these oBasestationtacles we suggested a combined DDPG-activated caching of content approach and consensus process for permission blockchain. The following points are sum up in this research paper:

(i) The caching of content procedures between vehicles and base stations is secured by an entrusted permission-ed blockchain.

(ii)To secure the approach of caching of contents among vehicles, DRL concepts are followed by taking the movability of vehicles into account.

(iii)To determine the new blockhead Consensus Blockchain methodology is followed to ensure security and privacy in the content caching procedure.

II.RELATED WORK

Thanks to the characteristics of blockchain-like, decentralization, immutability, anonymity, and security which allured researchers in the recent past. The research paper in [4] put forward the working of a blockchain integrated with neural networks for caching of contents among drones. The research paper in [5] suggested a form of dynamic caching of contents among hierarchically structured wireless channels to safeguard self-governing cache dissemination between unreliable groups. The research paper in [6] gives an idea of a distributed information control model empowered with blockchain for a network of vehicles. The attributes of block nodes in edge computing networks of vehicles like computation competence or QoS fulfillment are not necessary for setting up a blockchain. In the union of a network of vehicles and blockchain, some approaches make use of proof-of-stake Consensus to interchange the original PoW model. By not competitive PoW puzzle, the writers of the paper in [26] determined that Delegated Proof-of-Stake (DPoS) is especially applicable for weightless vehicles to provide vehicles empowered with blockchain. One more investigation of [7], the writers in [8] gave the idea of upgraded DPoS consensus for sharing of data approach where the prominence of a node is selected in DPoS to calculate the standard of RSUs. The writers in [3] and [24] also used prominence-based consensus in building a blockchain for vehicular network environments. The research paper in [25] applied proof-of-driving dealt with blockchain to ensure a smart sharing of data between vehicles. The writers in [26] availed proof-of-integrity in the blockchain of vehicles to supervising the gathered information related to vehicles from varied sensors for a privacy-alert traffic accident analysis. Notwithstanding, all of the above research works are not convenient for edge computing network of vehicles, are still not taking into account the edge nodes with limited computing capability which is the real demand of users. This paper suggested confirmation -of -Task-related consensus for vehicular edge caching.

III. Blockchain controlled IoT caching of content model

A new technique of Blockchain-enabled IoT caching of content model that consists of a user layer and an edge layer, as demonstrated in fig(1).The user layer consists of various IoT devices, set up with numerous sensors and operations, can gather a range of collectible content from surroundings, etc. But when maximum capacity exceeds the available cache, It starts looking for nearby devices, that can act as a caching donor.

The edge layer consists of a lot of base stations that are spread evenly in a distinct range to act as edge servers with effective communication and computations. The job of Base stations is to spot free caching reserves of IoT devices and distribute caching requests to the caching donor. Also, Base stations are much capable of anticipating the range as well as the period of content cache between caching seekers and caching donors. The master of edge layer controls safety guidelines and encrypted keys of Base stations and vehicles with manipulation free applications.

A. Content Caching in vehicles with the implementation of Blockchain

Content caching can be either cache seeker or cache donor. Content caching in vehicles can happen by following steps.

(i)Setting up Identity and System Initialization.

The process of content caching begins with the registration of the vehicle with the unique account with a key. To generate a secure and authenticated key elliptic curve digital signature procedure is used. Especially, vehicles and BASESTATIONs will have an registered identity after authentication of master of edge layer. The permitted identified entity comprises of a public key, a private key, and a certificate, which can be set as (PBK, PR, CERT). The source address is set as the public key of the caching transaction which is used to test the authenticity of transactions. To sign a transaction private key can be used and the certificate is used for unique identification of the vehicle by combining registration information of the vehicle.

Every participating vehicle will be having an e-wallet. The wallet address can be created from its public key. At the beginning of content of caching procedure, every vehicle seeks the e-wallet-Id of other vehicles from the mater of edge layer. Every vehicle sends its e-wallet-id to a global account pool and then downloads other vehicle e-wallet-Id for content caching from there.

ii) Generation of Smart Caching of Content :

In this section, every Base station collects cache seek of vehicles and keep oBasestationerving their existing cache reserve under its span. A Vehicle v_x posts its caching demand to the adjacent Base station b_y. The cache seek the message of vehicle v_x includes the desired cache reserve Rv_x, current location $locv_x$, public key $PBKv_x$, signature $Sigv_x$, certificate $Certv_x$, and timestamp ts, which can be denoted as

$$\mathbf{SEEK_{vx\rightarrow by}=EPBKby(Rv_x||locv_x||PBKv_x||Sigvx||certv_x||t_s);}$$

where $EPBKb_y$ denotes that message $SEEK_{vx\rightarrow by}$ is encrypted with public key $PBKb_y$, $Sigv_x = SignSK_{vx}(Rv_x||locv_x)$ stand for the digital signature of Rv_x and $locv_x$ private key SKv_x , and ts is the timestamp of the current message. Vehicle periodically sends its available caching resource to the nearest BASESTATION by for caching resource sharing.

The message about available caching resources includes the available caching resource Rv_d, location $loRv_d$, public key $PBKv_d$, signature $Sigv_d$, certificate $Certv_d$, and timestamp t_s, which can be described as

$$\mathbf{Mesg_{vd\rightarrow by} = EPBKb_y\,(Rv_d\,||loRv_d\,||PKv_d||Sigv_d\,||Certv_d\,||t_s);}$$

where EPBKby denotes that message $Mesv_d\rightarrow b_y$ is encrypted with public key PKb_j, $Sigvd = SignSKv_d\,(Rv_d\,||loRv_d)$ denotes that the digital signature of Rvd and locRvd is with private key SKvd , and ts is the timestamp of the current message. BASESTATIONs first verify the identity of cache seeks and cache donors after accepting their requests. and available caching resources of vehicles, BASESTATIONs first verify their identity. To speed up the verification process, Basestations adopt a batch verification process that can verify the validity of several identities simultaneously [31].A Basestation tracts the verification parameters from the received $SEEK_{vx\rightarrow by}$ and $Mes_{vd\rightarrow bj}$ and constructs the verification parameters as < P, Mes, Sig >, where Mes is the detailed message.Then, BASESTATION calls the batch verification algorithm for identity verification. If all V er(PBKvx ;Mesvx ; Sigvx) = 1

for all i 2 I [P, we have Batch((PKv1 ;Mesv1 ; Sigv1) ; :::; (PBKvx ;Mesvx ; Sigvx); :::) = 1 and batch verification is

passed. If one or more than one of the signatures are invalid, batch verification fails. After the batch verification, BASESTATIONs perform

V2V content caching mechanism to make caching pair matching. After completion of vehicular content caching operation. Every Basestation sends an acknowledgment to the cache seeker and cache donor. $Ack_{by\rightarrow vx}$ is the message that Basestation by is giving acknowledgment to cache seeker v_x and $Ack_{by\rightarrow vd}$ is the message that Basestation b_y is acknowledging caching donor v_j, which can respectively be described as

$$\mathbf{Ack_{by\rightarrow vx} = EPBK_{vx}\ (locRv_d\ ||Medx_p||PKv_d\ ||Sigb_j\ ||t_s);}$$

$$\mathbf{Ack_{by\rightarrow vd}=EPBKv_d\ (Rv_x||locRv_p|Medx_p||Sigb_y\ ||ts);(3)}$$

Where Med_{xp} seek is the wireless communication medium between the cache seeker and the cache donor. $EPBKv_x$ and $EPBKv_d$ denote that $Ack_{by\rightarrow vx}$ and $Ack_{by\rightarrow vx}$ are encrypted with publickey $PBKv_x$ and $PBKv_d$, respectively. The digital signature in $Ack_{by\rightarrow vx}$ seek is denoted as $Sigb_j$ = $SignSKb_j$ ($loRv_d$ || chanip),and the digital signature in $Ackb_{y\rightarrow vd}$pro is denoted as $Sigb_j$ =$SignSKb_j$(Rv_x||$locRv_p$ ||Med_{ip}).Based on the above two messages, vehicles autonomous execute the pre-programmed smart contract. The smart contract consists of two modules. The first module is to carry out content delivery. The second module is to transfer a certain amount of coins from the wallet of content caching seeker v_x to the wallet of content caching donor vehicle v_d.

C. Transactions Registration:

Cache seekers make fees to cache donors at the end of a transaction and register the cache content transaction phase. Then the cache seeker v_x forwards the introduced transaction to the immediate Basestation b_y. The Basestation first authenticates the acquired transaction, followed by encryption and broadcasting of the previous transaction phases in the blockchain domain. The transaction consists of the mutual cache reserve Rv_x , the coins obtained by the cache seeker v_x($coin_{vx\rightarrow vy}$) ,the content seekers wallet addresses, the content seekers signature, and timestamp, specifically,

$$\mathbf{T_{vx\rightarrow by}=EPBKb_y\ (Rv_x||coinv_x\rightarrow v_dwalletv_xaddr\ ||\ walletvdaddr||Sigv_x\ ||\ t_s\ )}$$

(4)

The above $EPBKb_y$ indicates that $T_{vx\rightarrow by}$ is encrypted with public key $PBKb_y$, $Sigv_x$= $SignSK_{vx}$ (Rv_x ||$coin_{vx\rightarrow vd}$) stands for digital signature of Rv_x and transferred coins $coin_{vx\rightarrow vy}$ with private key SKv_x . The recently produced transaction is broadcasted over the whole network for audit and verification. The authenticated transactions are arranged and grouped into a meddle-free data structure, called block. The blocks are allied in a sequential order by hash pointers to form a blockchain.

D. Creation of Blocks for Consensus Process:

Different Basestations create transaction blocks and start the consensus process. Basestation selects a newly created block as a Blockhead. Then Blockhead advertises the block with a timestamp for the process of scrutinization of the block. The newly created block is also authenticated by its adjacent Basestations. The agile block which solves confirmation of task (CoT) becomes the Blockhead to create the new block. But CoT is calculation demanding work, so it is not preferred for vehicular interconnections [16].Consequently the demand of an agile and decisive blockchain consensus algorithm with low energy-engrossing and less time requirement arises.

CONTENT OF CACHE ESTABLISHED WITH DEEP DETERMINISTIC POLICY GRADIENT

Here a DDPG-planted caching of content algorithm is proposed with the consideration of moving vehicles to solve the safety quotient in the user layer.

A. Caching of content accompanied by Manhattan grid model

Here a design of caching of content issue among vehicular interconnection is modeled with the aim of maximizing system use, by concentrating on a particular block with Basestation and its group of vehicles under its own reachable area. A group of cache seekers can be expressed as I = $\{v_1,\ldots,v_x\}$ and the group of cache donor as P = $\{v_1\ldots.v_x\}$ so that $I \cap P = \Phi$ and $I + P = N$. The content produced by caching seeker Rv_x can be depicted as{ Rv_x ,Tv_x }to which Rv_x and Tv_x depict the needed cache reserve and the utmost content delivery latency, respectively.

We can imitate any city as a Manhattan-style grid, with a homogenous block size over a fixed square area. The Manhattan grid model is established as a standard movability model

according to the European Telecommunications Standards Institute (ETSI) [17]. In this model, the map comprises of various parallel and perpendicular streets. Each perpendicular street has two lanes one for north and another for south direction and east and west for horizontal streets. Vehicles move crosswise streets and may curve at merging with a given probability. Given the direction of vehicles is represented by η where its value can be {north||south||west||east}.The probability that every vehicle moves at an intersection can be denoted as

$$P\eta = \frac{1}{\frac{1}{\Delta intV} + \frac{T wait\ Pwait}{2}} = \frac{2}{2 + T wait\ Pwait \delta intV} \quad (5)$$

Where V= constant velocity, δintV = Density of convergence is depicted as ζ .The probability of stoping of a vehicle at convergence is given by

$$P\zeta = 1 - P\eta = \frac{T wait\ Pwait\ \delta intV}{2 + T wait\ Pwait\ \delta intV} \quad (6)$$

V_d is a cache donor vehicle and has a local cache with the capacity of C_d. We define xi ϵ {0,1} as caching of content variable. If the content of cache seeker v_x is cached at cache donor v_d, $x_{id} \epsilon 1$. Otherwise, $x_{id} = 0$. When v_x caches its content on v_d, it has to make a certain fee for caching resource usage. The fee that vx pays to v_d is denoted as $coin_{vx \to vd} = x_{id}\ \zeta\ cv_x$, where $\zeta > 0$ is the amount per part caching reserve. Because of quantity of cache reserve on every vehicle is restricted, the complete available cache reserve v_d of all contents on v_d cannot outstrip its cache capacity, i.e.$\sum_{i\epsilon I}, x_{id}\ c_{vx} <= C_d$ .The possibility of communication between Vehicles depend on interval between them does not surpass the communication distance, i.e., $d_{ip} < \gamma$ [18]. The communication data rate between vehicle v_x and vehicle v_d can be expressed as

$$\mathbf{Rx_d = b \log2 (1 + p_x (hx_d)^{-\grave{\alpha}} / \sigma^2)} \quad (7)$$

Here b is the channel bandwidth. The transmission power of v_x is denoted as p_x. The channel gain is hx_d, ὰ depicted as the path loss exponent, and the noise power is shown as σ^2.

$X_{xd} \epsilon \{0,1\}$, ∀ pϵP ∀ iϵI 10(iii)

According to (7), the V2V content transmission latency is

$$\mathbf{Txd = \frac{Rv_x}{Rx_d}} \quad (8)$$

Because of the reason content of cache seeker vi should be transmitted within τ_{vi},we have

$\sum p\epsilon P X_{xd} Txd <= \tau_{vx}$

The sum of used up energy comprises of two parts: (i)transmission energy utilization (ii) content caching energy utilization.Let ß denote the price per energy utilization.

The energy cost for V2V content caching is

$$Exd = ß\left(dx \frac{Rv_x}{Rxd} + e_0 * Rv_x \right) \quad (9)$$

The above e_0 denotes the unit energy utilization per caching resource. Under the constraints of caching capacity and maximum content delivery latency, the problem to maximize system can be identified as follows:

$$\max \sum_{i\epsilon I} \sum_{p\epsilon P} xd\zeta R_{vx} - X_{xd} E_{xd}$$

$$\sum_{i\epsilon I} X_{xd} R_{vx} \le C_d \quad \forall\ p\epsilon P \quad 10(i)$$

$$\sum_{p\epsilon P} X_{xd}\ T_{xd} \le \tau_{vx} \quad \forall\ i\epsilon I \quad 10(ii)$$

As X_{xd} is a binary variable, the feasible set and objective function of equation (10) is notconvex. Never the less we can apply anapproximate algorithm to solve it, the scalabilityof the solution is still weak as the solution mayfail with the more number of vehicles.Moreover, the movement of vehicles leads totime-varying wireless channels such that theconventional optimization method is impractical.

Since deep deterministic policy gradient algorithm is best for decision-making problems with high-dimensional and time-varying features, here we attempt to utilize it to solve the problem.

V2V Content Caching Solution using deep reinforcement learning

The problem (10) can be modeled by utilizing the concepts of DDPG with system state, action, and reward, as shown in Fig. 3. A system state is a space to reflect the perceived vehicular domain. Let S denote the system state space.The state $S_t \epsilon S$ at time slot t can be defined as

- S_t= {V(t), L(t),E(t), η(t), F_i,C} (11)

where [(V(t) = [V11(t).....VXP (t)] is a vector which denotes V2V communication data rate between vehicles at time slot t;

- T (t) = [T11 (t)…….. TIP (t)] is a vector which denotes content latency vxa V2V transmission at time slot t;
- E(t) = [E11(t)…….EIP (t)] is a vector which represents energy utilization of V2V content delivery at time slot t;
- η(t) = [(η(t)…… η(t) $\eta(t)_{I+1}$(t) …….$\eta(t)_{I+P}$] is a vector which represents every vehicle's driving direction at time t;
- Fi ={ [RVx… ,tvx],…, [RV_x ; tvx] matrix denotes the essential caching resource and the utmost content delivery latency of caching seekers.
- C = [C1,…Cd] is a vector that represents the caching proportions of caching donors.

Because of movability, the location of every vehicle is different at different times such that V2V delivery data rates, content transfer latency, and the energy utilization of V2V content delivery are time-varying. The step of V2V content caching is to match caching pairs. Let A stand for the action space. The action $a_t \in$ A at time slot t is defined as

a_t = [X11(t),… XID (t)] (12)

After taking action, the system will receive an immediate ⍱ reward ⍱ (st, at). Since the objective of the problem (10) is to maximize system use, we define the immediate reward as

$$⍱(st, at) = E\left[\sum_{i\epsilon I}\sum_{p\epsilon P} xd\zeta R_{vx} - X_{xd}E_{xd}\right] dlt \quad (13)$$

the immediate reward is the current system value. Apart from that, the system will get a penalty and ⍱ (st, at) = dlt, where dlt is a negative constant. The optimal V2V content caching strategy is to maximize the long-term reward which can be defined as

$$Reward = \max E\left[\sum_{t=0}^{T-1} \epsilon^t ⍱(st,at)\right] \quad (14)$$

$\epsilon \in [0,1]$ is discount factor . We can utilize DDPG to solve the proposed V2V content caching problem by taking into account System state, action, and reward.

B. Ancestors of DDPG:

Deep Q Network (DQN): Most of the actor citric theory for the continuous problem is for stochastic policies. DQN is a type of deep reinforcement learning algorithm which considers deep neural networks rather than Q-function to investigate actions. DQN is a significant tool that can learn optimal policies with high-dimensional monitoring spaces.

DPG: An efficient gradient computation for deterministic policies, with proof of convergence.

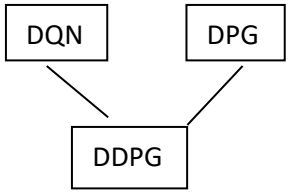

Fig.(2)

C. Vehicular Content Caching with Deep Deterministic Policy Gradient

A caching module can be composed of three modules: initial network, the destination network, and replay memory. The initial network plans to match content caching pairs by the deterministic policy gradient method. The initial network comprises two deep neural networks, namely the initial actor neural network and the initial critic neural network. The destination network is used to generate destination value for training the initial network. The structure of the destination network is similar to the structure of the initial network but with different parameters. Replay memory is used to store perceived tuples. Experience tuples include present state, the chosen action, reward, and next state, which can be randomly sampled for training the initial network and destination network. The detailed interaction processes among these models are shown fig(3)

A parameterized value θ_π denoted as a new explored policy which can map a current state $a=\pi(St|\ \theta_\pi)$ where a is an action generated by mapping with $\pi(St|\ \theta_\pi)$ which is the newly discovered edge caching and content delivery

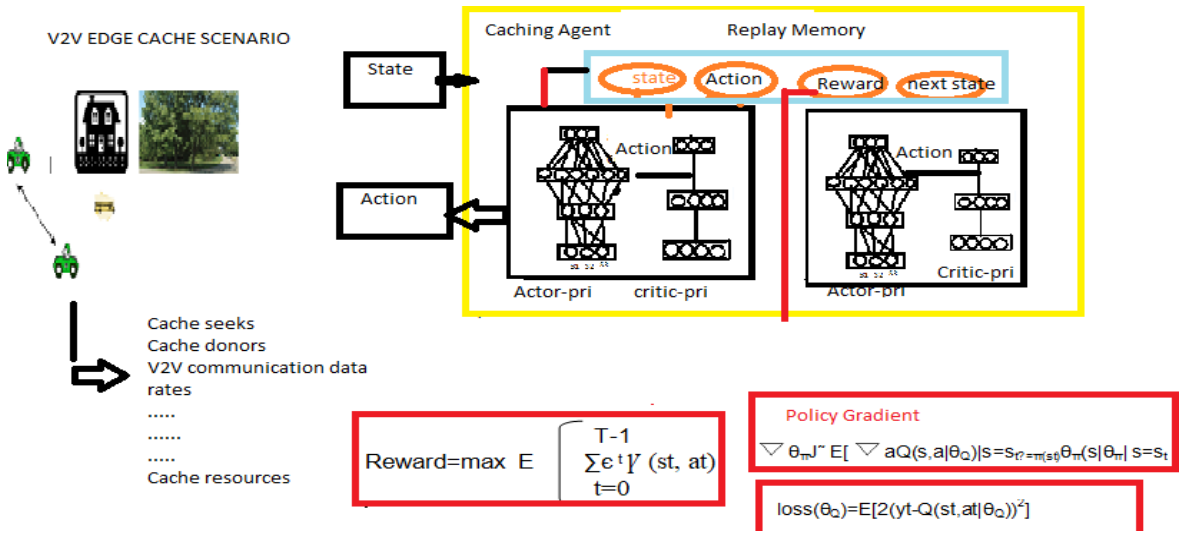

Fig.4 DDPG based V2V Content Caching.

policy produced by primary actor neural network. By adding an Ornstein-Uhlenbeck noise ʁt, the constructed action can be described as [19]

$$\mathbf{a_t = \pi(st|\ \theta_\pi) + ʁt} \qquad \mathbf{(15)}$$

The primal actor neural network improves network parameter θ_π using the sampled policy gradient, computed as

$$\mathbf{\theta_\pi J \approx E[\ \nabla\ aQ(s,a|\theta_Q)|s=s_{tà=\pi(st)}\theta_\pi(s|\theta_\pi|\ s=s_t} \qquad (16)$$

Where $Q(s,a|\theta_Q)$ is an action-value function. At every training phase, at every training step, $\theta_\pi$is updated by a mini-batch experience <st,at,Rimm, St+1 > T ϵ {1…….v} sampled arbitrarily from replay memory

$$\theta_\pi=\theta_\pi - \frac{à_\pi}{v}\sum_{t=1}^{v}\left(\nabla\ aQ(s,a|\theta_Q)|\ s=s_{t,à=\pi(st)}\ \nabla\ \theta_\pi\ (s|\ \theta_\pi|\ s=s_t\right) \qquad (17)$$

Here $à_\pi$ is represented as learning rate of prelilimnary neural network which estimates the performance of the picked out action based on the action-value function. The action-value function is determined by the Bellman optimality equation and can be shown as

$$Q(st,at)|\ \theta_Q)=E\ [Rimm(st,at)+\epsilon Q(st+1,\pi_{st+1}\ |\ \theta_Q\ )] \qquad (18)$$

The network parameter θ_Q is updated by preliminary critic neural network with the help of minimizing the loss function loss(θ_Q)

$$loss(\theta_Q)=E[2(yt-Q(st,at|\theta_Q))^2] \qquad (19)$$

Wher yt is the target value can find out by

$$Yt=Rimm(st,at)+\epsilon Q'(s_{t+1}(\pi)'s_{t+|}|\theta_\pi^T|\theta QT \qquad (20)$$

The gradient of loss function loss(θ_Q) is find out by taking its Ist derivative, which can be calculated as[20

$$\nabla\ \theta_Q Ls=E[2(yt-Q(st,at)|\theta_Q)\ \nabla\ \theta_Q(st,at)] \qquad (21)$$

As given in (21), the parameter θ_Q of primary critic neural network can be upgraded by training step, θ_Q is updated with a tiny-batch experiences<st,at,Rimm, st+1> by randomly sampled from replay memory

$$\Theta_Q = \theta_Q - \frac{à_Q}{V}[\ 2(yt-Q(st,at|\theta_Q))^2\ \nabla\ \theta_Q(st,at)] \qquad (22)$$

Algorithm 1: DDPG-empowered V2V content caching algorithm for vehicular edge computing
Input: Different variables related to the motion of vehicular model δ_{int}, Ř, T_{wait}, and P_{wait} ;
Variables about the connection between vehicles
Ǵα
Transference power, bandwidth, channel gain, and path loss exponent;
The state of the anticipated vehicular surroundings s_t;
Output: The examined caching pairs;
1: load α (s|θ $^{\alpha}$) and Q(s a|θ^{Q}) of the primary network)
with parameters θ^{α} and θ^{Q}
2: load the target network with parameters $\theta^{\alpha'} \leftarrow \theta^{\alpha}$ and $\theta^{Q'} \leftarrow \theta^{Q}$
load replay memory;
3: for every phase do
4: Setup vehicular surrounding;
5: for every time step t do
6: Execute action a_t based on α (s|θ $^{\alpha}$) and state s_t .
7: OBasestationerve reward Ř (s_t; a_t) and state s_{t+1} based on
(13);
8: Store the tuple < s_t; a_t; Ř (s_t; a_t); s_{t+1} > into Replay memory;
9: Sample a mini-batch of tuples from replay memory;
10: calculate the target value y_t and update θ^{Q} by minimize the loss function (19);
11: Update using α (s|θ $^{\alpha}$) using the sampled policy gradient
(16);
12: Update destination networks with:
$\theta^{\alpha'} \leftarrow c\omega\theta^{\alpha}+(1-c\omega)\theta^{\alpha'}$
$\theta^{Q'} \leftarrow c\omega\theta^{Q}{}^{Q}+(1-c\omega)\theta^{Q'}$
13: end for
14: end for

Algorithm 2 Design the edges of bipartite graph G
Input: The outputs of the explored caching pairs;
Output: The building bipartite graph G(I; P; E) ;
1: Set E ?.
2: if P_p 6 1 then
3: There is only one node vd_1 corresponding to caching
donor p.
4: for every x_{0ip} > 0 do
5: Add edge (i; vd_1) into edge set E and let the edge weight as $e_{ip1} = x_{0ip}$.
6: end for
7: else
8: Find the minimum index i_s where $\sum_{i_o=1} x_{0ip} > s$.
9: if i = i_{s1} + 1; ::; i_s 1, and x_{0ip} > 0 then
10: Add edge (i; vd_s) into edge set E with weight $e_{ips} = x_{0ip}$.
11: else if i = i_s then
12: Add edge (i; vd_s) into edge set E with $e_{ips} = 1$ $\sum_{i=1} x_{0ip}$.
13: else
14: Add edge (i; $vd_{(s+1)}$) into edge set E with weight $e_{ip(s+1)} = \sum_{i=1} x_{0ip}$ s.
15: end if
16: end if

of –
and

Computing

here we present the details of CoT mechanism for permission blockchain in the edge layer and try to suggest how to determine BASESTATION utility for block controller selection.

A. CoT Consensus

Blockchain consensus scheme which takes advantage of voting and election to secure blockchain from centralization and unauthorized usage [22]. the number of items competing in DPoS consensus is very less, making it possible to effectively reduce the time consumption and

energy consumption to every consensus. Influenced by DPoS, the proposed CoT consensus consists of two parts: 1) agent election, 2) block construction and verification, as shown in fig(4)In CoT we use the Utility of base stations to elect delegates. The details about CoT consensus are introduced in the following.

1) Delegate election: Given authenticated coin transfer in content caching transactions, among vehicles that may not be trusted, we tag vehicles as token holders to control the delegate selection procedure. Since delegates involve the crucial case of consensus algorithms (i.e., block construction and

Validation), they are by choice neutral nodes [22]. In the content caching process, BASESTATIONs are not straightly engaging in content delivery and coin payment, which implies they do not get any profit from the caching process, such that they are ideal neutral nodes and can be regarded as delegate candidates.

In every election, vehicles vote for their favored BASESTATIONs with the highest utility. The utility is used to calculate the quality of BASESTATIONs. Bigger utility means BASESTATION is provided with more capable computing and processing competency to generate and authenticate blocks. Every vehicle is given one unique vote per phase and the vote count is proportional to the number of coins it has. The top n candidates with the most votes are elected to form a delegation. where n is an odd integer and not higher than the number of BASESTATIONs.

2) Block Production and Verification: In the block construction and verification process, delegates are divided into two types of actors: Blockheader and controller, where block header is responsible for

transaction collection and block production, and the controller is responsible for block authentication At every block production process, one of block node delegate acts as the block header and the other delegates act as controllers. Block header is nominated in a round-robin manner among delegates which indicate every delegate can become a block header to produce the block. For example, in Fig. 3delegate B is the block header responsible for creating Block B. In a specific block production and verification process,

the block header first collects a certain amount of V2V content caching transactions and then calculates a correct hash to create an unverified block.

CONCLUSION

This research work suggested a smart caching of content technique for interconnecting of vehicles by coordinating DDPG with permission blockchain. Here, a blockchain privileged caching of content mechanism where vehicles served as cache seekers and cache donors to do content caching and Basestations treated as reviewers to control and support permission blockchain. To learn continuous changing network topology and time fluctuating wireless network medium we utilized DDPG to model an perfect caching of content proposal, by exercising permission blockchain to achieve the security and privacy of content caching between vehicles and introduced a new controllers selection metric, CoT to accelerate block verification.

REFERENCES

[1]Z. Zhou, C. Gao, C. Xu, Y. Zhang, S. Mumtaz, and J. Rodriguez, “Social big-data-based content dissemination in the internet of vehicles," IEEE Trans. Ind. Informat., vol. 14, no. 2, pp. 768–777, Feb 2018.

[2] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, ``Continuous control with deep reinforcement learning," Comput. Sci., vol. 8, no. 6, pp. 187_201, 2015.

[3] V. Konda and J. N. Tsitsiklis, ``Actor-critic algorithms," in Proc. Adv.Neural Inf. Process. Syst., 2000, pp. 1008_1014.

[4] V. Sharma, I. You, D. N. K. Jayakody, D. G. Reina, and K. R. Choo, "Neural-blockchain-based ultrareliable caching for edge-enabled uav networks," IEEE IEEE Trans. Ind. Information., vol. 15, no. 10, pp. 5723–5736, Oct 2019.

[5]W. Wang, D. Niyato, P. Wang, and A. Leshem, “Decentralized caching

for content delivery based on blockchain: A game-theoretic perspective,"in 2018 IEEE

International Conference on Communications (ICC), May 2018, pp. 1–6.
[6] Z. Yang, K. Yang, L. Lei, K. Zheng, and V. C. M. Leung, "Blockchain-based decentralized trust management in vehicular networks," IEEEInternet Things J., vol. 6, no. 2, pp. 1495–1505, April 2019.
[7]Y. Yuan and F. Wang, "Towards blockchain-based intelligent transportation systems," in 2016 IEEE 19th International Conference on IntelligentTransportation Systems (ITSC), Nov 2016, pp. 2663–2668.
[8] J. Kang, Z. Xiong, D. Niyato, D. Ye, D. I. Kim, and J. Zhao, "Toward secure blockchain-enabled internet of vehicles: Optimizing consensus management using reputation and contract theory," IEEE Trans. Veh. Technol., vol. 68, no. 3, pp. 2906–2920, Mar 2019.
[9]. I. Sorkhoh, D. Ebrahimi, R. Atallah, and C. Assi,"Workload scheduling in vehicular networks with edge cloud capabilities," IEEE Trans. Veh.Technol., vol. 68, no. 9, pp. 8472–8486, Sep 2019
[10.] K. Shanmugam, N. Golrezaei, A. G. Dimakis et al., "Femtocaching: Wireless content delivery through distributed caching helpers," IEEETrans. Inf. Theory, vol. 59, no. 12, pp. 8402–8413, Dec 2013.
[11]. J. Liu, N. Kato, J. Ma, and N. Kadowaki, "Device-to-device communications lte-advanced networks: A survey," IEEE Commun. Surv. Tut.,vol. 17, no. 4, pp. 1923–1940, Fourth quarter 2015.
[12] Y. Dai, D. Xu, S. Maharjan et al., "Artificial intelligence empowered edge computing and caching for Internet of Vehicle," IEEE Wirel.Commun., vol. 26, no. 3, pp. 12–18, Jun 2019
[13]Y. Wei, F. R. Yu, M. Song, and Z. Han, "Joint optimization of caching, computing, and radio resources for fog-enabled iot using natural actor-critic deep reinforcement learning," IEEE Internet Things J., vol. 6, no. 2, pp. 2061–2073, Apr, 2019.
[14] K. Zhang, Y. Zhu, S. Leng, Y. He, S. Maharjan, and Y. Zhang, "Deep learning empowered task offloading for mobile edge computing in urban informatics," IEEE Internet Things J., vol. 6, no. 5, pp. 7635–7647, 2019.
[15] Y. Dai, D. Xu, S. Maharjan, et al., "Blockchain and deep reinforcement learning empowered intelligent 5G beyond," IEEE Network, vol. 33, no. 3, pp. 10–17, May 2019.
[16] Y. Dai, D. Xu, S. Maharjan, et al., "Blockchain and deep reinforcement learning empowered intelligent 5G beyond," IEEE Network, vol. 33, no. 3, pp. 10–17, May 2019.
[17] T. UMTS, "101 112 v3. 2.0 (1998-04),"universal mobile telecommunications system; selection procedures for the choice of radio transmission technologies of the umts (umts 30.03 version 3.2. 0),"," European Telecommunications Standards Institute, Tech. Rep, 1998.
[18] J. Liu, H. Nishiyama, N. Kato, and J. Guo, "On the outage probability of device-to-device-communication-enabled multichannel cellular networks: An rss-threshold-based perspective," IEEE J. Sel. Area. Comm., vol. 34, no. 1, pp. 163–175, Jan 2016.
[19] T. P. Lillicrap, J. J. Hunt, A. Pritzel et al., "Continuous control with deep reinforcement learning," arXiv preprint arXiv:1509.02971, 2015.
[20] T. P. Lillicrap, J. J. Hunt, A. Pritzel et al., "Continuous control with deep reinforcement learning," arXiv preprint arXiv:1509.02971, 2015.
[21] Y. Yuan and F. Wang, "Towards blockchain-based intelligent transportation systems," in 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), Nov 2016, pp. 2663–2668
[22]D. Larimer, "Delegated proof-of-stake (dpos)," Bitshare whitepaper,2014.
[23] J. Kang, Z. Xiong, D. Niyato, Y. Zou, Y. Zhang, and M. Guizani,"Reliable federated learning for mobile networks," arXiv preprint arXiv:1910.06837, 2019.
[24]] J. Kang, Z. Xiong, D. Niyato, S. Xie, and J. Zhang, "Incentive mechanism for reliable federated learning: A joint optimization

approach to combining reputation and contracttheory," IEEE Internet Things J., pp.
1–1, 2019.
[25] M. Singh and S. Kim, "Blockchain-based intelligent vehicle data-sharing framework," arXiv preprint arXiv:1708.09721, 2017.

[26] M. Cebe, E. Erdin, K. Akkaya, H. Aksu, and S. Uluagac,"Block4forensic: An integrated lightweight blockchain framework for forensics applications of connected vehicles," IEEE Communications Magazine, vol. 56, no. 10, pp. 50–57, OCTOBER 2018.