

Energy Utilization and Transmission Delay Optimization in Cox VANET networks

Adil Boumaalif^{1,*} and Ouadoudi Zytoune^{1,**}

¹ENSA, Ibn Tofail University, Kenitra, Morocco

Abstract. Among the challenges in VANETs networks there is power control, which permits to achieve better performance in terms of Signal to Interference plus Noise Ratio (SINR), Energy Efficiency (EE), Energy Utilization (EU),... Another important point to address is the transmission delay, which we should adapt it to be under an acceptable threshold. For that, we model the network as a Cox Line-Point process, and the transmitter as an M/D/1 queue server. To solve this tradeoff problem, we use machine learning techniques, especially Q learning algorithm. It is shown, via simulations, that through our algorithm, the vehicular transmitter be able to learn its transmit power in an autonomous way, and achieve better performance for the energy utilization rate, the system waiting time and the area spectral efficiency.

1 Introduction

Recently, vehicular ad hoc networks (VANETs) are becoming one of the most active research fields. VANETs are the corner stone in the building of intelligent transportation systems (ITSs). They permit to the vehicles and the road side units (RSUs) to exchange important information such as the traffic status, the roads conditions and accidents occurrence to warn vehicle's drivers of potential risk [1]. These networks are a sub-class of mobile ad hoc networks (MANETs) that consider vehicles as mobile nodes to provide communications between vehicles, i.e. vehicle to vehicle (V2V) and between vehicles and RSUs, i.e. vehicle to infrastructure (V2I) [2]. However, these networks have their own characteristics as compared to MANETs. Even nodes are mobile as in MANETs, the vehicles are expected to move in the road topology instead of moving randomly [3]. Thus, VANETs have a special spatial geometry where the locations of nodes are present uniquely in the roadways lane. Recent model based on stochastic geometry is given to represent the road layouts and the location of vehicles is given [4, 5]. Due to the principally linear nature, roads can be modeled using a Poisson Line Process (PLP) in \mathbf{R}^2 that can be deduced from a Poisson Poisson Processes, where instead of points, the undirected lines (road lanes) are randomly distributed on plane [6]. Thus, the roadways are modeled as a network of lines that are distributed on the plane according to a PLP.

One of the challenging topics in VANETs networks is the Power control, it is defined as the mechanism that permits to each node to adapt its transmission power to achieve better communication without perturbing others communicating nodes. A lot of power control were proposed in the literature. One can review works as [7–10]

that tackle the problem of selecting the transmit power of vehicles. Particularly, the paper [11] proposed a power control algorithm to find the optimum transmission power by adding a power tuning feedback beacon during each beacon message transmission. Thus, the sender estimates the distance to the receiver and predicts the required transmission power. Even the algorithm optimizes the transmission power, it generates additional transmission delay due to the control messages exchanged.

Another important key of performance in the VANETs is the message transmission delay, which should be always under an acceptable threshold. Hence, many works were proposed to deal with this issue [12–14]. The key idea behind the messages transmission delay is to propose techniques that handle this issue in Network or MAC layers. In [15], authors proposed a routing protocol that minimizes the end to end delay by selecting the appropriate relays nodes. In [16], authors proposed to use machine learning to enhance routing in order to reduce data transmission latency. The papers [17, 18] gave a survey on multi-layer techniques and MAC protocols respectively, to improve message transmission delay.

Both these issues can be solved with reinforcement learning (RL). One of the most famous algorithm if RL is Q learning, which is an off-policy control algorithm, i.e. it does not depend on the policy the agent uses to navigate the environment [19].

In this paper, we propose a reinforcement learning based algorithm to handle the transmission power control that aims to optimize the transmission energy while guaranteeing that the transmitted messages are arriving with a limited delay.

The rest of this paper is organized as follows. Section II describes the system model and the problem formulation. Section III details the reinforcement learning algorithm

*e-mail: adil.boumaalif@gmail.com

**e-mail: zytoune.ouadoudi@uit.ac.ma

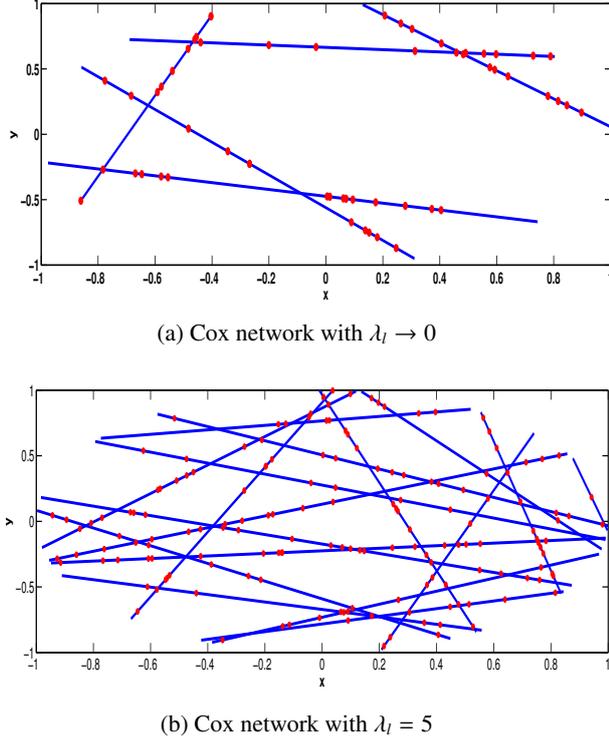


Figure 1: Cox network with different values of λ_l

used to solve the optimization problem defined in Section II. Section VI presents numerical and simulation results and finally, Section V concludes this paper and gives directives for future work.

2 System Model and Problem formulation

2.1 Cox VANET Network

We consider a Poisson line process PLP ϕ_l with line density μ_l to model the road system. Hence, the density of the equivalent point process of ϕ_l is $\lambda_l = \mu_l/\pi$. Besides, the locations of the nodes on each line form a 1D PPP with density $\lambda_v = \lambda_n + \lambda_r$, where λ_n is the density of vehicular nodes and λ_r is the density of RSUs. So, we form a Cox line-point process ϕ_c as illustrated in Fig.1. We consider that just $p\lambda_v$ can transmit in each time slot, where $0 \leq p \leq 1$. Without loss of generality, we consider a typical link, where the receiver is located at the origine, and the corresponding transmitter is located at a distance d from the origin on the same line.

Assuming that all links are affected by Rayleigh fading and the fading powers are exponentially distributed with mean 1 and considering a standard path loss model with $2 < \alpha < 6$. We further assume that all nodes transmit with the same power P_t . So, as proven by Chetlur in [10], The success probability of the typical link P_s is:

$$P_s = \exp\left[-\frac{\beta N_0 d^\alpha}{P_t} - 2p\lambda_v \int_0^\infty \frac{\beta d^\alpha x^{-\alpha}}{1 + \beta d^\alpha x^{-\alpha}} dx - 2\pi\lambda_l \int_0^\infty 1 - \exp\left(2p\lambda_v \int_0^\infty \frac{\beta d^\alpha (x^2 + y^2)^{-\alpha/2}}{1 + \beta d^\alpha (x^2 + y^2)^{-\alpha/2}} dx\right) dy\right] \quad (1)$$

Where $P_s = \mathbb{P}(SINR > \beta)$, with β is the SINR threshold, N_0 is the noise power.

In this work, we are interested by the special case when the density of lines in the network is very sparse, i.e, $\lambda_l \rightarrow 0$. So, the success probability of the typical link converges to that of a 1D PPP and is given by the following reduced equation:

$$P_s = \exp\left[-\frac{\beta N_0 d^\alpha}{P_t} - \frac{2p\lambda_v d\beta^{1/\alpha}}{\text{sinc}(1/\alpha)}\right] \quad (2)$$

Based on this result, we define the Energy Utilization Rate (EU_r) as the ratio of the energy consumption per unity of time to the transmission rate, and assuming that each message is transmitted in one second [20], so:

$$EU_r = \frac{P_t}{P_s} \quad (3)$$

Another important factor is the area spectral efficiency (ASE), it is defined as the average number of successfully transmitted bits per unit time per unit bandwidth per unit area in the network. Assuming that all the transmitted symbols are from Gaussian codebooks, It's given that [10]:

$$ASE = \lambda P_s \log_2(1 + \beta); \text{ bits/s/Hz/km}^2 \quad (4)$$

where λ is the average number of active transmitting nodes per unit area, as we are interested by the special case in which our model converges to a 1D PPP scenario, the equation of λ is given by $\lambda = p\lambda_v$, and:

$$ASE = p\lambda_v \log_2(1 + \beta) \exp\left[-\frac{\beta N_0 d^\alpha}{P_t} - \frac{2p\lambda_v d\beta^{1/\alpha}}{\text{sinc}(1/\alpha)}\right] \quad (5)$$

As the transmission probability p impact the overall spectral efficiency with a tradeoff manner, there is an optimum value of transmission probability p^* that maximizes the ASE of the network, in the case of 1D PPP:

$$p^* = \frac{\text{sinc}(1/\alpha)}{2\lambda_v d\beta^{1/\alpha}} \quad (6)$$

2.2 Queueing Model

In our model, we represent the typical transmitter as a system with single server having a queue length represented with $M/D/1$ queueing, where arrivals are determined by a Poisson process and job service times are fixed (deterministic) [21]. We should define the principles metrics of the queueing system according to our system model. If we represent by λ_a the rate with which the arrivals occur, and μ is the service rate, so, in our case: $\mu = P_s$, and the utilization of our server is:

$$\rho = \frac{\lambda_a}{\mu} = \frac{\lambda_a}{P_s} \quad (7)$$

Another important metric which we will use in our study is the average waiting time in the system, denoted by ω , and defined as: $\omega = \frac{1}{\mu} + \frac{\rho}{2\mu(1-\rho)}$. Finally, we got:

$$\omega = \frac{2P_s - \lambda_a}{2P_s(P_s - \lambda_a)} \quad (8)$$

2.3 Problem Formulation

Based in this system model, our objectif is to minimize the energy utilization rate, while ensuring that the system waiting time is always under a predefined threshold and implicitly the transmit power is always under a maximum value. To achieve this goal, the following optimization problem must be solved:

$$\begin{aligned} \min_{P_t} \quad & EU_r \\ \text{s.t.} \quad & \omega \leq \tau \\ & 0 \leq P_{t,i} \leq p_{max}, \forall \end{aligned} \quad (9)$$

Where τ is the system waiting time threshold, and $P_t = (P_{t,1}, \dots, P_{t,i}, \dots, P_{t,N})$ is the set of the node transmit powers.

We can see clearly that when the transmit power of the transmitting node decreases, the energy utilization rate will decrease automatically, while the system waiting time will increases and may exceed the predefined limit. On the other hand, to ensure the condition of proceeding time, the transmit powers should be always greater than a certain value. In order to find the optimal transmit power, a Q-learning based power control algorithm will be used.

3 Q-learning energy/delay optimization algorithm

We use a simple Q learning algorithm to solve our optimization problem defined in the previous section. The **Agents** here is the transmitting node that want to learn its transmit power which minimizes the energy utilization rate while guaranteeing the condition on the system waiting time. The other elements of Q learning algorithm are defined as follows:

State: We define the state as:

$$S = \begin{cases} 1 & \text{if } \omega \leq \tau \\ 0 & \text{otherwise} \end{cases}$$

Action: The action of each agent consists of a set of transmitting power levels. It is denoted by

$$A = (P_{t,1}, \dots, P_{t,i}, \dots, P_{t,N})$$

where every agent has N power levels.

In this paper we use $\epsilon - greedy$ strategy to choose actions based on the current Q -value estimation, which is

described as follows:

- choose action randomly with probability ϵ from the action space,
- choose action according to $a = \arg \max_{a \in A} Q(s, a)$ with probability $1 - \epsilon$

Reward: The reward function reflects the learning objectives of our algorithm, so we define the reward as

$$R = \begin{cases} 1/E_u & \text{if } \omega \leq \tau \\ -1 & \text{otherwise} \end{cases}$$

Our power control strategy is described in the following algorithm.1. in the beginning, we associate for each vehicular transmitter a set of feasible transmit power levels A . Then, in the learning process, we use $\epsilon - greedy$ technique to select the future action to execute. The selected action is then performed, and the related reward is obtained. Thus, the Q-table is updated. This process is repeated till the number of iteration is achieved.

Algorithm 1 Q learning algorithm

1. Initialization:

for each state $S_i \in S$ and each action $a_j \in \mathcal{A}$ **do**

 Initialize $Q(S_i, a_j)$ to zeros

end for

evaluate the starting state $S_i \in S$

2. Learning:

while MaxIteration not reached **do**

 choose $a_j \in \mathcal{A}$ using the ϵ -greedy policy based on Q

 Take action a_j and observe the immediate reward R_t and next state S'

$$Q_{t+1}(S_i, a_j) \leftarrow Q_t(S_i, a_j) + \alpha[R_t + \gamma \max_{a'_j} Q_t(S', a'_j) - Q_t(S_i, a_j)]$$

$$S_i \leftarrow S'$$

end while

4 Numerical and Simulation Results

4.1 Simulation Parameters

In this work, we consider a VANET formed by Vehicles and RSUs with density $\lambda_v = 0.1$ located in roads map with density $\lambda_l \rightarrow 0$. We assume that the transmission probability $p = 0.3$, the distance between typical nodes is $d = 10m$, the path loss exponent is $\alpha = 3$, the SINR threshold is $\beta = 9dB$, and the noise power is $N_0 = -100dB$.

The Q-learning parameters are as follows: the learning rate is $\alpha = 0.1$, the discount factor is $\gamma = 0.9$ and the ϵ greedy parameter $\epsilon = 0.1$. We define the action space as a vector of discrete values of transmit power levels, i.e. $A = [-60 : 20dBm]$ with step of 10 dBm. Also, the parameters of the queuing model are $\lambda_a = 10 \text{ messages/s}$ and authorized system waiting time $\tau = 3s$.

4.2 Results and Discussion

In Fig.2, we plot the Q values as a function of learning iterations. We can see clearly that Q values become slightly

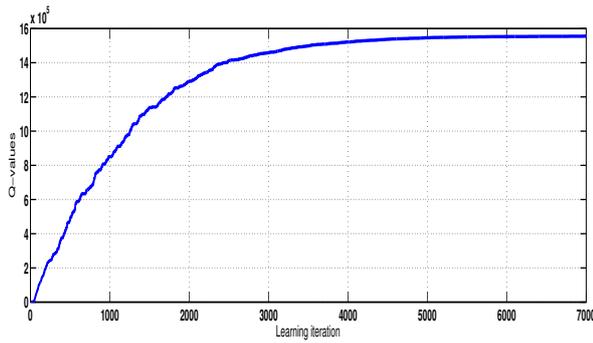


Figure 2: Q-values vs learning time

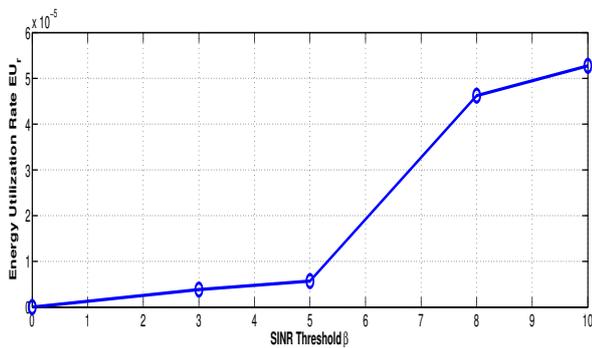


Figure 3: Energy utilisation vs SINR threshold β

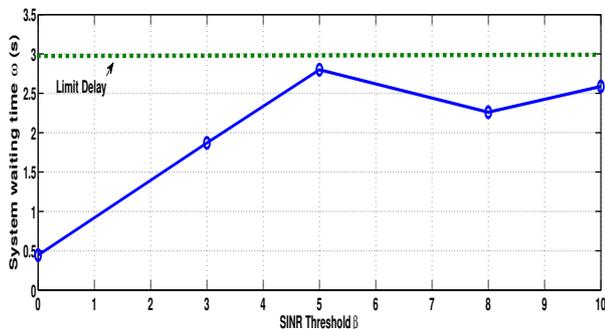


Figure 4: System waiting time vs SINR threshold β

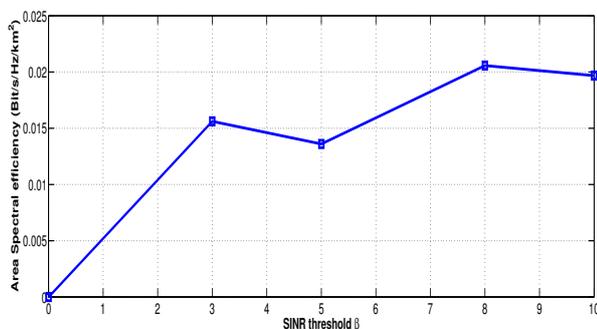


Figure 5: Area Spectral Efficiency vs SINR threshold β

constant after about 4000 iterations, which means that our algorithm converges at this moment. So, the obtained results can be considered as optimal.

In Fig.3, we plot the energy utilization rate with regards of SINR threshold. As depicted, EU_r increases when β increases. That is justified by the fact that we need more transmit power to achieve the required SINR when the threshold becomes greater.

By the way, Fig. 4 gives the evolution of system waiting time with regards to the SINR threshold. We can observe, that even ω varies in function of β , it remains always under the predefined value $\tau = 3s$.

Finally, Fig. 5 represents the variation of the area spectral efficiency vs the SINR threshold. As depicted in this figure, we can observe that ASE increases as β increases with some small fluctuations. In the expression of ASE we have β in two times: in the log function, and in the P_s expression. the observed fluctuations are due to the chosen steps of transmit power.

5 Conclusion

In this paper, we presented a Q-learning based technique to optimize simultaneously the transmission energy usage and the message latency in a VANET. The network is considered as a Cox Poisson Line Process. We modeled the message transmission as a queue $M/D/1$. Each packet waiting to be transmitted is stored in this FIFO queue. Based on the success transmission probability, we tried to select the best transmission power that permits to minimize the energy usage rate while guaranteeing that the system waiting time is always under a predetermined value. Numerical studies proved the effectiveness of our proposition. As a future work, we plan to consider improving the learning method to allow the scalability of our proposition.

References

- [1] H. Hartenstein and L. P. Laberteaux, "A tutorial survey on vehicular ad hoc networks," *IEEE Commun. Magazine*, vol. **46**, no. 6, pp. 164-171, (2008).
- [2] M. Sivasakthi and S. Suresh, "Research on vehicular ad hoc networks (VANETs): an overview," *Journal of Applied Sciences and Engineering Research*, vol. **2**, no. 1, pp. 23-27, (2013).
- [3] H. Moustafa and Y. Zhang, "Vehicular Networks: Techniques, Standards and Applications," CRC Press, Boca Raton, Fla, USA, (2009).
- [4] Dhillon, H.S., and Chetlur, V.V. (2020), "Poisson Line Cox Process: Foundations and Applications to Vehicular Networks. Poisson Line Cox Process,".
- [5] Bartlomiej Blaszczyszyn, Paul Muhlethaler, "Random linear multihop relaying in a general field of interferers using spatial Aloha," (2012).
- [6] D. Stoyan, W. S. Kendall, J. Mecke, and L. Ruschendorf, "Stochastic geometry and its applications," Wiley New York, vol. **2**, (1987).

- [7] C. Wang and G. Yu, "Power Control and Channel Assignment Mechanisms for Cluster-Based Multichannel Vehicular Ad-Hoc Networks," 2013 12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications, (2013), pp. 1762-1767.
- [8] Facchina, Caitlin, "Adaptive Transmission Power Level with Vehicle Speed Approximation of Density for VANET Congestion Control" (2020). Electronic Theses and Dissertations. 8298.
- [9] C. Dai, X. Xiao, L. Xiao and P. Cheng, "Reinforcement Learning Based Power Control for VANET Broadcast against Jamming," 2018 IEEE Global Communications Conference (GLOBECOM), (2018), pp. 1-6.
- [10] Selo Sulistyono, Sahirul Alam, "SINR and throughput improvement for VANET using fuzzy power control," International Journal of Communication Systems. **Volume. 31**, Issue: 10, Page: e3579, (2018).
- [11] Xu Guan, Sengupta, R., Krishnan, H., Fan Bai, "A feedback-based power control algorithm design for VANET, Mobile Networking for Vehicular Environments, IEEE INFOCOM. IEEE, (2007).
- [12] Jin Tian and Qianru Han 2019, "The priority assignment of messages effects on delay performance in VANET ", J. Phys.: Conf. Ser. 1187 042106.
- [13] Jin Tian, Qianru Han, Shangjing Lin, "Improved Delay Performance in VANET by the Priority Assignment", (2019), IOP Conference Series Earth and Environmental Science 234:012081.
- [14] B. Ravi and J. Thangaraj, "End-to-end delay bound analysis of VANETs based on stochastic method via queueing theory model," 2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), pp. 1920-1923, (2017).
- [15] S. Peng, R. Chai, Q. Chen and Y. Qin, "Minimum end-to-end transmission delay based routing algorithm for VANETs," 2017 9th International Conference on Advanced Infocomm Technology (ICAIT), pp. 176-181, (2017).
- [16] Lai WK, Lin M-T, Yang Y-H, "A Machine Learning System for Routing Decision-Making in Urban Vehicular Ad Hoc Networks," International Journal of Distributed Sensor Networks, (2015).
- [17] Ma M, Liu K, Luo X, Zhang T, Liu F. "Review of MAC Protocols for Vehicular Ad Hoc Networks," Sensors (Basel). 2020;20(23):6709. Published (2020).
- [18] Usman Ali, Khan and Sang Sun Lee, "Multi-Layer Problems and Solutions in VANETs: A Review", Electronics (2019), 8, 204.
- [19] C. J. Watkins and P. Dayan, "Q-learning," Machine learning, **vol. 8**, no. 3-4, pp. 279-292, (1992).
- [20] Y. Luo, P. Hong and R. Li, "Energy Efficiency-Delay Tradeoff in Energy-Harvesting-Based D2D Communication: An Experimental Learning Approach," in IEEE Communications Letters, **vol. 22**, no. 8, pp. 1704-1707, (2018).
- [21] O. Brun, J.M. Garcia, "Analytical Solution of Finite Capacity M/D/1 Queues," Journal of Applied Probability, pp. 1092-1098, (2000).