Scalable Private P2P network for distributed and Hierarchical Machine Learning in VANETs

Jeong Min Jeon,*Choong Seon Hong
Department of Computer Engineering
Kyung Hee University, Korea
jmjeon0212,cshong@khu.ac.kr

Abstract—With the recent development of the Internet of Things (IoT), applications are becoming smarter and connected devices are being used in all aspects. As the amount of collected data increases, machine learning (ML) technology has been applied and is being used as a useful tool for extracting vast amounts of information. If the data set is wide and distributed, old machine learning algorithms cannot be used because the whole training data should be centralized in one location. Therefore, distributed learning, federated learning, and circular learning are being used. In this paper, we propose a new Trust-based Edge network architecture that is suitable for distributed learning and hierarchical machine learning in a Vehicular ad-hoc network(VANETs) it is inspired by Dempster-Shafer theory with Scalable Chord Peer to Peer Network. In order to cut down on computation, communication costs, and time.

Index Terms—chord P2P Network, Network Architecture, Dempster-Shafer theory, Communication cost.

I. INTRODUCTION

The total Internet of Things (IoT) connected devices is used to amount to 21.5 billion entities worldwide by 2025 [1]. The biggest challenges the IoT confront now is how to handle and utilize the number of data generated by a bunch of connected devices. Machine Learning(ML) has been developed as a useful tool to extract information from extensive data(e.g. videos, pictures) [2]. When datasets are extremely large-scale and distributed, conventional machine learning algorithms cannot be used, because they want the whole training data to be centralized at one location. Particularly in the vehicular network, the vehicles communicate to Road Side Units(RSU).

This is a kind of mobile distributed edge system, it is not allowed to send all of the data to a server-side location. the reason is that the amount of data generated by the bunch of devices is too big to be sent using the communication network. Further, there is a possibility of privacy reasons. one may not want to send original data to the central server. These motivations the need for federated learning, distributed or hierarchical machine learning [3]. So this emphasizes the need for intelligent distributed learning mechanisms that effectively use restricted resources. In this paper, we proposed scalable and trust-based vehicular edge network architecture in VANETs for reducing the communication, computational

cost, and time to send their accuracy, learning parameter, gradient, or something. [4] [5] use the blockchain technology approach, but it is computation, complexity is high. Therefore, we applied a private Peer to Peer(P2P) Network for the social-trust vehicle cluster group. In this trust-group, they are connected with Chord P2P Network which applied Dempster-Shafer theory of evidence to reduce the communication, computational cost, and time. If we applied the theory, we can choose the best outcome for all vehicles in the group. Which one is more reliable for sending the model to the global. There should be many considering points such as accuracy, gradient information, learning parameters, etc. We can not only depend on accuracy. We have to depend on the other trust parameters based on and previous experience of each Trust Group. That's why we applied evidence theory.

The rest of the paper is organized as follows: Section II presents the related work. Section III introduces the system model. Section IV explains the working of the proposed scheme. We conclude the paper with future research directions in Section VI.

II. RELATED WORK

A. Demster-Sharfer Theory(DST)

Dempster-Shafer(D-S) [6] [7] evidence theory as a Bayes probability theory is an useful means of handle unsure information [8] [9]. D–S evidence theory has been applied in various fields, such as pattern classification [10] and decision-making [11] [12].

Definition 2.1: Assume that $\Omega = \{ \theta_1, \theta_2, ...\theta_i, ...\theta_N \}$ is a set of N elements which represent exhaustive and mutually exclusive events, Ω is the frame of discernment(FOD). The power set of Ω consists of 2^N elements denoted as follows :

$$2^\Omega = \begin{cases} \emptyset, \{\theta_1\}, \{\theta_2\}, ..., \{\theta_N\}, \{\theta_1, \theta_2\}, \\ ..., \{\theta_1, \theta_2, ..., \theta_i\}, ..., \Omega \end{cases} \quad (1)$$

Definition 2.2: A mass function m is a mapping from the power set 2_Ω to the interval [0,1]. m satisfies:

$$m(\emptyset) = 0, \sum_{A \in \omega} m(A) = 1. \quad (2)$$

if $m(A) > 0$, then A is called a focal element. $m(A)$ reflects the degree of the evidence supports the proposition A.

Definition 2.3: The body of evidence(BOA), also known as the basic probability assignment(BPA) or basic belief assignment(BBA), is defined as the focal sets and the corresponding mass functions:

$$(R, m) = \{< A, m(A) >: A \in 2^\Omega, m(A) > 0\}. \quad (3)$$

where R is a subset of the power set 2^Ω.

Definition 2.4: A BPA m can also be represented by the belief function Bel or the plausibility function Pl, defined as follows:

$$Bel(A) = \sum_{\emptyset \neq B \subseteq A} m(B), Pl(A) = \sum_{B \neq A \subseteq \emptyset} m(B). \quad (4)$$

Definition 2.5: In Dempster-Shafer evidence theory, 2 independent mass functions m_1 and m_2 can be fused with Dempster's rule of combination:

$$m(A) = (m_1 \bigoplus m_2)(A) = \frac{1}{1-k} \sum_{B \cap C = A} m_1(B)m_2(C), \quad (5)$$

where k is a normalization factor defined as follows :

$$k = \sum_{B \cap C = \emptyset} m_1(B)m_2(C). \quad (6)$$

In this paper, we applied this evidence theory. it can be applied to quantify the probabilistic properties of simulation results generated from incomplete knowledge and information. it can be explained that the focal element represents the degree to which the position of a side (x-value, Quantile) on the cumulative probability distribution is supported by the evidence obtained. In VANETs, There are many Trust-Group or Cluster. in this case, we can extract the best outcome from those groups by using this theory. it can reduce the computational time and cost.

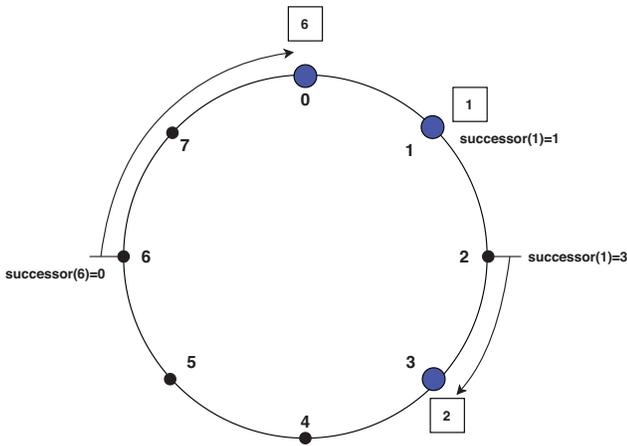*B. Chord P2P Network with Consistent Hashing*

Fig. 1. The example of an identifier circle consisting of the 3bit address space

Chord is a method created at the University of Berkeley, and is a method of assigning key values for each node and data to an m-bit$(0 - 2^m - 1)$ circular address space. Mapping from each node and data key value to the address space is performed using a hashing function, and a node that is greater than or equal to the data key k in the address space is named as a successor of k to increase the efficiency of routing. Used to fig 1 is an example of 3 bit address space that composes Chord network. Three nodes of 0,1, and 3 exist on the network, and as successors of data keys 1,2, and 6,3 nodes of 1,3 and 0 participate in the system. In Chord, each node maintains routing information called finger tables. The size of the finger table varies according to the size of the address space. In the case of m-bit,it consists of m rows. Each row has successor information of data corresponding to the n(1-M) power of 2, based one the location of the current node. In this paper Based on the finger table [13].

In addition, in the case of departure due to node failure, the ripple effect can be minimized by periodically calling the stabilize() function.
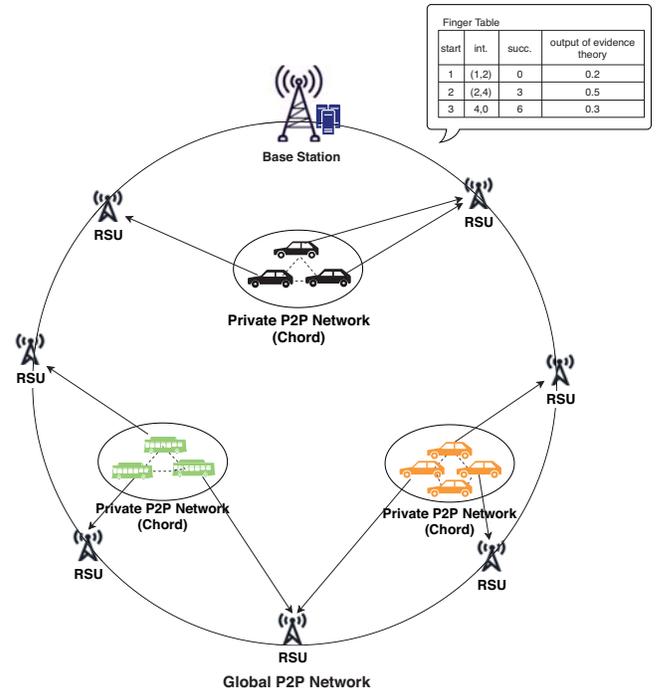
III. SYSTEM MODEL

Fig. 2. the overview of our proposal system model

The overview of our proposal system model is shown fig 2, There are the different small Group(Trust Group). It is connected with P2P Network. In this paper, we applied evidence theory. Based on this theory. We choose the best outcome from all of the vehicle in these group. Which one is more reliable for sending the model to the global. There should be many considering point such as accuracy, gradient information, learning parameters, etc. it can not only depends on the accuracy. we have to depend on the other trust parameters

based one previous experience of each trust group. these are the reason why we apply the evidence theory.

IV. PROPOSED FRAMEWORK

fig 3 Based on this theory, the outcome is saved in the finger table. There are Belief and Plausibility.

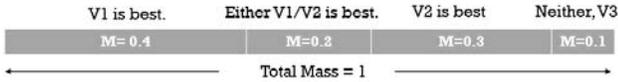| V1 is best.	Either V1/V2 is best.	V2 is best	Neither, V3
M= 0.4	M=0.2	M=0.3	M=0.1

Total Mass = 1

Fig. 3. The proposed Network Architecture based on evidence theory

Witness : The V1 or V2 is the best accuracy, parameter, gradient value or anything based on the distributed learning model. Probably V1, though it could be V2. Now we need to choose which one is the best model.

TABLE I
PARAMETERS FOR THE EVIDENCE THEORY

Belief	Plausibility
Bel(V1) = 0.4 Bel(V2) =0.3	Pl(V1) = 0.4 +0.2 = 0.6 = 1-Bel(not V1)
Bel(not V1) = 0.3+0.1 = 0.4 Bel(not V2) = 0.5	Pl(V2) = 0.2+0.3 = 0.5 = 1- Bel(not V2)

Initially, based on the current outcome we choose the initial vehicle. After that, we update by combining previous experience(i.e., history) and current outcome.
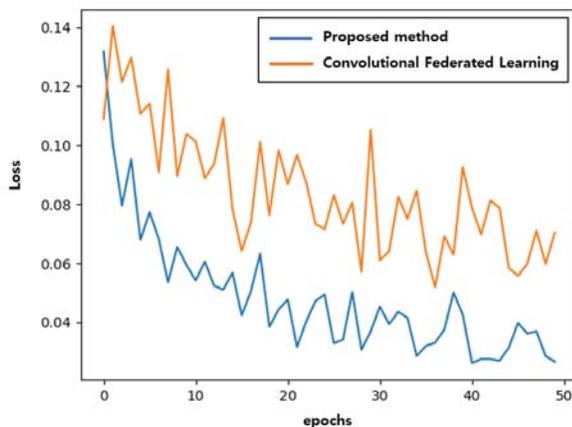
V. PERFORMANCE AND EVALUATION

Fig. 4. The proposed method vs Convolutional Federated Learning loss result

fig 4 compares the loss of our proposed method with convolutional federated learning. In this paper, we use the model of a fully connected neural network, and the MNIST Dataset was used. The number of edges is 100. And epochs are 50 times. Each edge is connected to a different number of users. It is shown that learning can be performed at a faster rate than the convolutional federated learning method, and communication and computational cost and time can be reduced. we use the parallel method to save time and cost. the optimal outcome can be extracted through Trust-Group and Evidence theory. Accelerated training is possible in Federated learning hierarchical or distributed learning.

VI. CONCLUSION

In this paper, We introduce a new network architecture for vehicular edge network for the distributed and hierarchical machine learning in VANETs. Firstly, we proposed a Trust-based P2P vehicular edge computing overlay network for creating the vehicular cluster. secondly, we apply a consistent hashing (chord) technique to manage this overlay network with less complexity. Each Trust-based P2P group uses evidence theory and control-flow to select the best outcome for each group to send to the global server. Considering accuracy, gradient information, learning parameters with prior experiences. It is more robust and less complex. We can add and remove when it is necessary by using our proposed network architecture. Also, the proposed network architecture can reduce computational and communication cost and time.

REFERENCES

[1] Statista, "Iot number of connected devices worldwide" Available Online: https://www.statista.com/statistics/1101442/iot-number-of-connected-devices-worldwide/statisticContainer, (Accessed on 23 Oct, 2020).
[2] Marjani, Mohsen Nasaruddin, Fariza Gani, Abdullah Karim, Ahmad Hashem, Ibrahim Siddiqa, Aisha Yaqoob, Ibrar. (2017). Big IoT Data Analytics: Architecture, Opportunities, and Open Research Challenges. IEEE Access. 5. 10.1109/access.2017.2689040.
[3] L. Liu, J. Zhang, S. H. Song and K. B. Letaief, "Client-Edge-Cloud Hierarchical Federated Learning," ICC 2020 - 2020 IEEE International Conference on Communications (ICC), Dublin, Ireland, 2020, pp. 1-6, doi: 10.1109/ICC40277.2020.9148862.
[4] H. Kim, J. Park, M. Bennis and S. Kim, "Blockchained On-Device Federated Learning," in IEEE Communications Letters, vol. 24, no. 6, pp. 1279-1283, June 2020, doi: 10.1109/LCOMM.2019.2921755.
[5] S. R. Pokhrel and J. Choi, "Federated Learning With Blockchain for Autonomous Vehicles: Analysis and Design Challenges," in IEEE Transactions on Communications, vol. 68, no. 8, pp. 4734-4746, Aug. 2020, doi: 10.1109/TCOMM.2020.2990686.
[6] Dempster, Arthur P. "Upper and lower probabilities induced by a multivalued mapping." Classic works of the Dempster-Shafer theory of belief functions. Springer, Berlin, Heidelberg, 2008. 57-72.
[7] Shafer, Glenn. A mathematical theory of evidence. Vol. 42. Princeton university press, 1976.
[8] Deng, Xinyang, Wen Jiang, and Zhen Wang. "Zero-sum polymatrix games with link uncertainty: A Dempster-Shafer theory solution." Applied Mathematics and Computation 340 (2019): 101-112.
[9] Jiang, Wen, Chan Huang, and Xinyang Deng. "A new probability transformation method based on a correlation coefficient of belief functions." International Journal of Intelligent Systems 34.6 (2019): 1337-1347.
[10] Liu, Zhun-Ga, et al. "Combination of classifiers with optimal weight based on evidential reasoning." IEEE Transactions on Fuzzy Systems 26.3 (2017): 1217-1230.
[11] Fu, Chao, et al. "Multiple criteria group decision making with belief distributions and distributed preference relations." European Journal of Operational Research 273.2 (2019): 623-633.
[12] He, Zichang, and Wen Jiang. "An evidential Markov decision making model." Information Sciences 467 (2018): 357-372.
[13] Ion Stoica, Robert Morris, David Karger, M. Frans Kaashoek, and Hari Balakrishnan, "Chord: A Scalable Peer-to-Peer Lookup Service for Internet Applications," Proc. of the 2001 ACM SIGCOMM Conf., 2001.