Effect of Packet Loss and Delay on V2X Data Fusion

Tzu-Kuang Lee,*, Jen-Jee Chen†, Yu-Chee Tseng*, Cheng-Kuan Lin‡
*Department of Computer Science, National Chiao Tung University, Hsinchu, Taiwan
†College of Artificial Intelligence, National Chiao Tung University, Hsinchu, Taiwan
‡College of Mathematics and Computer Science, Fuzhou University, Fuzhou, China

***Abstract*—Sensing data fusion is one of the most important technologies in autonomous driving. Its performance depends on advance communication technology. Cellular-Vehicle to Everything (C-V2X) initially defined as LTE V2X in 3GPP Release 14 is a solution for vehicle communication that includes Vehicle-to-Infrastructure (V2I), Vehicle-to-person (V2P), and Vehicle-to-Vehicle (V2V). Although 4G LTE and 5G provides high-speed transmission, packet loss and delay are still inevitable. Packet loss and delay affect the safety of autonomous driving, especially for the judgment of emergency. In this paper, we compare the accuracy of data fusion under different rate of packet loss and broadcast frequency on the simulated platform CALAR. And we propose a skill to improve accuracy. Experiments show that the proposed skill significantly alleviates the effect of communication packet loss and delay on the accuracy of V2X data fusion.**

***Index Terms*—V2X communication, Data Fusion, Autonomous**

I. Introduction

Surrounding vehicles' motion data and traffic information helps autonomous driving to make better decisions to achieve comfortable driving and avoid accidents. Sensing data fusion aggregates all information and integrates accurate environmental parameters, allowing the autonomous driving system to complete human-like driving decisions. With the advancement of wireless communication, computer vision (CV), and sensing technology, sensing data fusion has become one of the most important technologies of autonomous driving.

In the field of wireless communications, several countries just turned on their 5G mobile networks, and some start development of 6G. The 3GPP releases 14, 15, and 16 [1]–[3] proposed Vehicle-to-Everything (V2X) to solve the needs of vehicle communication that include Vehicle-to-Infrastructure (V2I), Vehicle-to-person (V2P), and Vehicle-to-Vehicle (V2V). They divided into 3 stages to complete the standardization of V2X.

Another major vertical focus area in release 16 is intelligent transportation systems (ITS). ITS provides a range of transport and traffic-management services, improves traffic safety, and reduces traffic congestion, fuel consumption and environmental impacts [8]. Efficient communication between vehicles and fixed infrastructure, but also between vehicles is the key to ensure ITS working well. 3GPP TR 38.885 [4] defined 25 use cases for advanced V2X communications, including vehicle platooning and cooperative communication using extended sensors.

Corresponding author: Yc-Chee Tseng (yctseng@cs.nctu.edu.tw)

Although 4G LTE and 5G provides high-speed transmission, low latency, and large bandwidth, packet loss and latency are still inevitable. Especially for information that has high demand for real-time, such as emergency or accident warning, delay or loss of packets will cause the judgment of data fusion being inaccurate. Lee et al. [9] propose a fusion algorithm which integrates V2X communications, GPS, camera information, and magnetometer data. It made the driver can visually see its surrounding vehicles' driving states. In their experiment environment, they don't consider the effect of loss of packets and delay on V2X. In this work, we point out the problem of sensing data fusion with packet loss and delay by implementing an augmented reality (AR) of car through integrating V2X communications and own vehicle information. We use simulate V2X data generated by CALAR simulator [6] to show the impact of packet loss and delay on data fusion.

The rest of this paper is organized as follows. Section 2 reviews some related work. Section 3 defines our problem and gives notations which uses in this paper. Our V2X data fusion scheme is shown in Section 4. Section 5 discusses our simulation results. Final section concludes this paper.

II. Related Work

Environmental awareness is an important issue for the safety and comfortable driving of self-driving cars. If we can correctly detect and classify the objects around the vehicle, then we can respond appropriately in all possible traffic scenarios. Sensing surrounding objects (such as vehicles, people, dogs, obstacles, etc.) and lane deviation warning are currently popular application scenarios. The self-driving Bertha of the Mercedes-Benz research [7] is proposed a multi-sensor network that combined 4 short-range radars, 3 long-range radars, and 4 video cameras to more accurately detect traffic conditions. A prototype of a high-resolution radar sensor is proposed in [10] , which can measure distance, speed and angle in an instant.

V2X can improve road traffic safety and improve the safety and comfort of autonomous driving. A visible light communication warning system is studied in [13]. Brake lights of a vehicle can be used to transmit messages for emergency hard brake to warn the the following vehicle. Sharing sensor data between vehicles and Road Side Units (RSUs) is studied in [11]. It can solve the blocking problem of ego perspective sensors. A V2X communication-centric traffic light controller system is proposed in [14], which exchanges information on all vehicles and traffic controlling system to implement a V2X-integrated traffic light controller system.

Data fusion in self-driving makes integrated utilization of the information obtained by different sensors and multi-sources, which avoids the perceptual limitations and uncertainties of a single-source, such that improves the external perception ability to ensure safe and comfortable driving. Fusing LIDAR point clouds and camera images is proposed for road detection [5]. Using multi-sensors (radar, lidar, and camera) is proposed to detect and classify moving objects including pedestrian, bike, car, and truck [12]. In this paper, we implement an AR of car through integrating V2X communications and own vehicle information to show the effect of packet loss and delay on the result of data fusion.

III. Notation and Problem Definition

We assume that the user's car has an RGB-D camera in the front to identify nearby car objects, information such as color, type, distance, angle, and license plate will be recorded. The calculation may be inaccurate due to shooting distance, shooting angle, weather, and obstructions etc. We also assume that each car can exchange their car profiles and driving states (such as speed, emergency braking, intention to turn, etc.) by V2X communication. Due to privacy protection, some vehicles may provide incomplete information. And the loss and delay of packets, some vehicles' status are not updated instantly. Our goal is to enhance the car's video information by mapping each driving information on surrounding vehicles who are coveraged by the car's RGB-D camera. When we receive several pieces of information from surrounding vehicles, the challenge is to tag the right information on the right car.

We consider a vehicle x. At time t, it will retrieve its current speed $x.sp(t)$, location $x.loc(t)$, and orientation $x.ori(t)$. Its surrounding image at time t is denoted by $I(t)$. We set $V(t) = \{v_1(t), v_2(t), \ldots\}$ being the set of all vehicles who identified in $I(t)$. Each $v_i(t)$ associated with the following information.

- $v_i(t).type$: the type of vehicle such as car, bus, truck, etc.
- $v_i(t).c$: the vehicle color which contains 3 parameters $v_i(t).c.R$, $v_i(t).c.G$, and $v_i(t).color.B$ (red, green, and blue)
- $v_i(t).lic$: the license plate
- $v_i(t).dist$: the distance between x and $v_i(t)$
- $v_i(t).angle$: the angle of $v_i(t)$ with respect to x's camera view

Through V2X communication, x receives a set of number of broadcasts $B(t) = \{b_1(t), b_2(t), \ldots\}$ at time t. Based on privacy protection, each $b_i(t)$ may associates with the following information.

- $b_i(t).time$: its data generating time
- $b_i(t).type$: the type of vehicle
- $b_i(t).c$: the vehicle color which contains 3 parameters $b_i(t).c.R$, $b_i(t).c.G$, and $b_i(t).c.B$
- $b_i(t).lic$: the license plate
- $b_i(t).loc$: the location of $b_i(t)$

To ensure that the data from V2X is kept up to date, x maintains a set of data $C = \{c_1, c_2, \ldots\}$. Each c_i is associated the following information.

- $c_i.time$: its update time
- $c_i.type$: the type of vehicle
- $c_i.c$: the vehicle color which contains 3 parameters $c_i.c.R$, $c_i.c.G$, and $c_i.c.B$
- $c_i.lic$: the license plate
- $c_i.loc$: the location of $b_i(t)$
- $c_i.dist$: the distance between x and this vehicle
- $c_i.angle$: the angle between x and this vehicle

We set d being the maximum tardiness. x confirms each c_i in C regularly. If current time is more than d difference from $c_i.time$, then x withdraws c_i form C. When x receives $b_i(t)$ at time t, it compares $b_i(t).lic$ with each c_j's $c_j.lic$. Then there are 2 cases to be dealt with.

- **UPDATE:**
  1) If $b_i(t).lic = c_j.lic$ and $b_i(t).time > c_j.time$, then x sets $c_j.time$ as $b_i(t).time$, $c_j.type$ as $b_i(t).type$, $c_j.c$ as $b_i(t).c$, $c_j.lic$ as $b_j(t).lic$, and $c_j.loc$ as $b_i(t).loc$.
  2) If $b_i(t).lic = c_j.lic$ and $b_i(t).time \leq c_j.time$, then x withdraws this one.
- **INSERT:**
  1) If $b_i(t).lic \neq c_j.lic$ for each $c_j \in C$ and $t - b_i(t).time \leq d$, then x create a new c_k in C where k did not use, and x sets $c_k.time$ as $b_i(t).time$, $c_k.type$ as $b_i(t).type$, $c_k.c$ as $b_i(t).c$, $c_k.lic$ as $b_j(t).lic$, and $c_k.loc$ as $b_i(t).loc$.
  2) If $b_i(t).lic \neq c_j.lic$ for each $c_j \in C$ and $t - b_i(t).time > d$, then x withdraws this one.

IV. V2X Data Fusion Scheme

We assume each vehicle has GPS receiver to provides its current location and magnetometer to provides its current speed. Each vehicle broadcasts its driving status (such as license plate, current location, current speed, etc.) through V2V communication. To map driving status to the image $I(t)$, we propose a V2V data fusion scheme in Fig. 2. It consists of 3 modules: (1) Data Preprocessing, (2) Weight Calculation, and (3) Mapping Decision. Those modules are running on the vehicle side. The approach by Lee et al. [9] is based on $B(t)$ and $I(t)$. They don't consider the effect of loss and delay of packets. We consider these situations and propose solutions. Below, we describe the proposed scheme in detail by showing how the three modules operate in a vehicle at current time t.

A. Data Preprocessing Module

When vehicle x gets the current front image $I(t)$, it calculates all vehicles' information identified in $I(t)$ including type, color, license plate, distance, and angle. All the information from $I(t)$ is put in $V(t) = \{v_1(t), v_2(t), \ldots\}$. And x obtains its current location $x(t).loc$ from its GPS, and its orientation

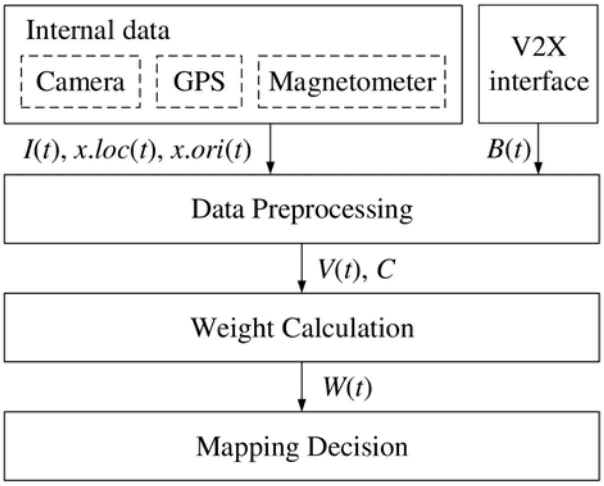

Fig. 1. Procedure of data fusion.

$x(t).ori$ from its magnetometer. Note that we set the value of orientation being 0 if the sensor's orientation is pointing north. Then for each $c_i \in C$, it updates $c_i.dist$ as $|x(t).loc - c_i.loc|$ and $c_i.angle$ as $\arccos \frac{|c_i.loc.y - x(t).loc.y|}{c_i.dist} + x(t).ot \mod 360$. Note that how x maintain the set C is described in Section 3.

Now that we have the data obtained from the photo analysis and the data obtained from V2V, the next step we will calculate the similarity matrix between $W(t)$ and C.

B. Weight Calculation Module

This module is responsible for calculating weight matrix $W(t) = \{w_{v_i(t),c_j} \mid v_i(t) \in V(t) \text{ and } c_j \in C\}$ by comparing the similarity between elements $v_i(t) \in V(t)$ and $c_j \in C$. We first calculate the weight of each element pair and then add them up. Higher total weight means higher similarity. The difference from the approach by Lee et al. [9] is that we consider the effect of loss and delay of packets when we calculate the weight of distance and angle. At the same time, even if some vehicles' V2V packets are not received at the moment, we still use the latest V2V packets to calculate its weight. For each $c_j \in C$, we set $\alpha_j = (\frac{d - \text{current time} + c_j.time}{d})^2$ being the weight of the effect of loss or delay of packet, where d is the time to measure the data freshness. According to data types of $V(t)$ and C, we define the following 5 weights.

1) Weight of type: We set the weight of type $w_{type}(v_i(t), c_j) = 1$ if $v_i(t).type = c_j.type$; otherwise, $w_{type}(v_i(t), c_j) = 0$.
2) Weight of color: We set $w_{color}(v_i(t), c_j) = \frac{M_{RGB} - d_{color}}{M_{RGB}}$ where $M_{RGB} = \sqrt{3 \times 255^2}$ is the maximum difference of RGB, and d_{color} is the difference of $v_i(t).c$ and $c_j.c$ in R, G, and B parameters.
3) Weight of license plate: We compare the license plate numbers (or characters) between $v_i(t).lic$ and $c_j.lic$ one by one. Let s_{lic} denote the same number of words between them, and let len_{lic} being the number of characters of license plate. We set the weight of license plate $w_{lic}(v_i(t), c_j) = \frac{s_{lic}}{len_{lic}}$.
4) Weight of distance: We set $w_{dist}(v_i(t), c.j) = \alpha_j \times \frac{M_{dist} - \min\{M_{dist}, |v_i(t).dist - c_j.dist|\}}{M_{dist}}$, where M_{dist} is the maximum distance between any GPS measure and any camera measure.
5) Weight of angle: We set $w_{angle}(v_i(t), c_j) = \alpha_j \times \frac{360 - |v_i(t).angle - c_j.angle|}{360}$.

Then we sum up all weights to derive $w_{v_i(t),c_j}(t) = w_{type}(v_i(t), c_j) + w_{color}(v_i(t), c_j) + w_{lic}(v_i(t), c_j) + w_{dist}(v_i(t), c_j) + w_{angle}(v_i(t), c_j)$. Note that here we give each weight equal importance. In face, according to the uniqueness of each element, giving different importance may further improve the performance.

C. Mapping Decision Module

We pair a vehicle in C to a vehicle in $V(t)$ as follows.

S1. We set C^* as C, V^* as $V(t)$, and S as an empty set.

S2. If both C^* and V^* are not empty set, then we compute the confidence of c_j in C^* as $d_j = \frac{\max_{v_i(t) \in C^*}\{w_{v_i(t),c_j}(t)\}}{\sum_{v_i(t) \in V^*} w_{v_i(t),c_j}(t)}$; otherwise, return S and end the process.

S3. We select c_j from C^* with the highest confidence d_j. Then we select $v_i(t)$ from V^* such that $w_{v_i(t),c_j} = \max_{v_k(t) \in V^*}\{w_{v_k(t),c_j}\}$.

S4. We remove c_j from C^*, $v_i(t)$ from V^*, and add $(c_j, v_i(t))$ to S. Then goto **S2**.

The final list of S is the pairing result. Fig. 3 show the difference between our scheme and the approach by Lee et al. [9]. We assume that Car 4 left the communication range since time t_2, and we assume that $t_4 - t_1 > d$ and $t_5 - t_3 < d$, where d is the time to measure the data freshness. At time t_5, Car 1, Car 2, Car 3, and Car 5 are in V2X communication range; Car 1 and Car 2 broadcast its driving status; Car 3 and Car 5 do not broadcast their driving status; the packet from Car 2 is losing. We map driving status to the image $I(t)$ according to $C = \{b_1(t_5), b_2(t_3), b_3(t_4), b_5(t_4)\}$, and Lee et al. are based on $B(t_5) = \{b_1(t_5)\}$. The question is that they are missing the information of Car 2, Car 3 and Car 5 since the packet of Car 2 is losing, and the packets of Car 3 and Car 5 do not broadcast in this moment. Our method uses the most recently data to solve this problem since the status of a vehicle will not change much in a very short time. On the other hand, we withdraw the data $b_4(t_1)$ form C since this data is expired.

V. Experimental Result

We conducted our experiment in an Inter i7-8700k server with 64GB RAM and two NVIDIA Geforce GTX 1080ti graphic cards. The software on it included Ubuntu 16.04.6, CARLA 0.9.7, and YOLOv3. CARLA simulator supports experiment of autonomous driving systems, and YOLO provides real-time object detection. Since CARLA simulator did not support V2V communication, it was done by post-simulation. We only allow vehicles to receive V2V packets from other vehicles within 30 meters.

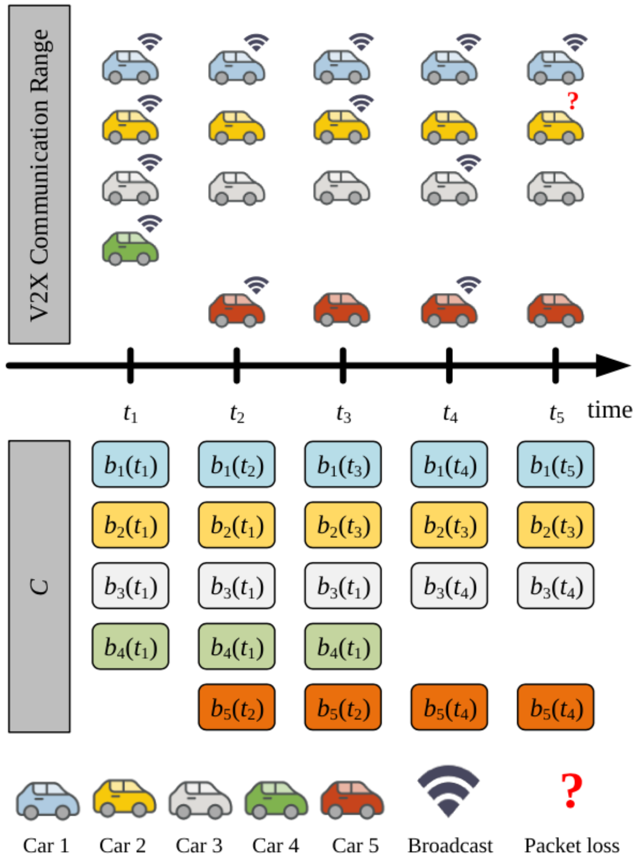

Fig. 2. The effect of packet loss.

In our scenario, we consider the rate of packet loss 5%, and we consider that V2V communication allows vehicles to broadcast and receive messages 1 time pre second. Fig. 4 points out the impact of packet loss on data fusion. After the system has been running for a while, the case without packet loss has about 82% accuracy. When packet loss occurs, its accuracy is only about 52%, while our proposed method can improve the accuracy up to about 72%. Therefore, it is possible to improve the accuracy of the results of data fusion in a real an lossy environment by predicting the changes of data or using historical data.

VI. Conclusion

Packet loss and delay of V2X communication in a real mobile environment are unavoidable. In this work, we measure the effect of packet loss and delay on the V2X data fusion and see that the performance significantly drops in a lossy network. To alleviate the negative effect of packet loss and delay, we enhance the previous method by maintaining a memory to memorize the most recently V2X message and considering the information freshness in the proposed schemer. Experimental result shows that the new method performs much better than previous method in a lossy environment. In the future, we will try to predict the car status by using historical data to further enhancing the accuracy.

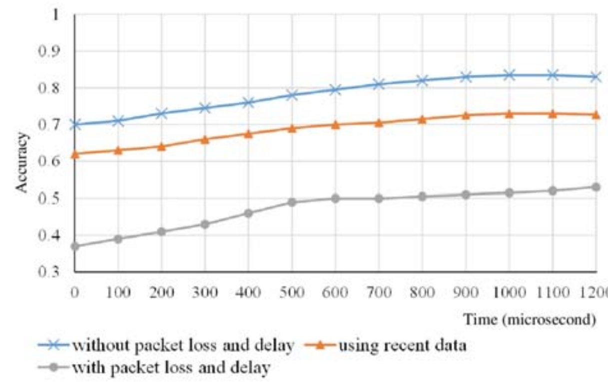

Fig. 3. Simulation result.

Acknowledgements

Y.-C. Tseng's research is co-sponsored by Pervasive Artificial Intelligence Research (PAIR) Labs, MoST 105-2218-E-009-003, and MoE project "Center for Open Intelligent Connectivity".

References

[1] 3GPP TR 21.914, "Release 14 Description; Summary of Rel-14 Work Items," V14.0.0, May 2018.
[2] 3GPP TR 21.915, "Release 15 Description; Summary of Rel-15 Work Items," V15.0.0, September 2019.
[3] 3GPP TR 21.916, "Release 16 Description; Summary of Rel-16 Work Items," V0.4.0, March 2020.
[4] 3GPP TR 38.885, "Study on NR Vehicle-to-Everything (V2X)", 2019-03, V16.0.0.
[5] L. Caltagirone, M. Bellone, L. Svensson, M. Wahde, "LIDAR–camera fusion for road detection using fully convolutional neural networks", Robotics and Autonomous Systems 111 (2019) 125-131.
[6] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: an open urban driving simulator, in Proceedings of the 1st Annual Conference on Robot Learning, 2017, 1–16.
[7] J. Dickmann, N. Appenrodt, J. Klappstein, H.-L. Bloecher, M. Muntzinger, A. Sailer, M. Hahn, and C. Brenk, "Making Bertha see even more: radar contribution", IEEE Access, 3 (2015) 1233 - 1247.
[8] N. M. Drawil and O. Basir, "Intervehicle-communication-assisted localization", IEEE Transactions on Intelligent Transportation Systems, 11 (2010) 678–691.
[9] T.-K. Lee, Y.-C. Kuo, S.-H. Huang, G.-S. Wang, C.-Y. Lin, and Y.-C. Tseng, "Augmenting car surrounding information by inter-vechicle data fusion", in Proceedings of the 2019 IEEE Wireless Communications and Networking Conference (WCNC).
[10] F. Meinl, M. Stolz, M. Kunert, and H. Blume, " An experimental high performance radar system for highly automated driving", in Proceedings of the 2017 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM).
[11] K. Sakaguchi and R. Fukatsu, "Cooperative perception realized by millimeter-wave V2V for safe automated driving", in Proceedings of the 2018 Asia–Pacific Microwave Conference.
[12] R O. Chavez-Garcia and O. Aycard, "Multiple Sensor Fusion and Classification for Moving Object Detection and Tracking", IEEE Transactions on Intelligent Transportation Systems, 17 (2016) 525-534.
[13] K. Siddiqi, A. D. Raza, and S. S. Muhammad, "Visible light communication for V2V intelligent transport system", in Proceedings of the 2016 International Conference on Broadband Communications for Next Generation Networks and Multimedia Applications (CoBCom).
[14] N. Varga, L. Bokor, A. Takács, J. Kovács, and L. Virág, "An architecture proposal for V2X communication-centric traffic light controller systems", in Proceedings of the 2017 15th International Conference on ITS Telecommunications (ITST).