Zhu, Chao; Pastor, Giancarlo; Xiao, Yu; Ylajaaski, Antti

**Vehicular Fog Computing for Video Crowdsourcing**

*Published in:*
IEEE Communications Magazine

*DOI:*
10.1109/MCOM.2018.1800116

Published: 01/10/2018

*Document Version*
Peer reviewed version

*Please cite the original version:*

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.
Vehicular Fog Computing for Video Crowdsourcing: Applications, Feasibility and Challenges

Chao Zhu, Giancarlo Pastor, Yu Xiao*, and Antti Ylä-jääski

Abstract—With the growing adoption of dash cameras, we are seeing great potential for innovations by analyzing the video collected from vehicles. On the other hand, transmitting and analyzing a large amount of video, especially high-resolution video in real time, requires a lot of communications and computing resources. In this work, we investigate the feasibility and challenges of applying vehicular fog computing for real-time analytic of crowdsourced dash camera video. Instead of forwarding all the video to the cloud, we propose to turn commercial fleets (e.g. buses and taxis) into vehicular fog nodes, and to utilize these nodes to gather and process the video from the vehicles within communication ranges. We assess the feasibility of our proposal in two steps. Firstly, we analyze the availability of vehicular fog nodes based on a real world traffic dataset. Secondly, we explore the serviceability of vehicular fog nodes by evaluating the networking performance of fog-enabled video crowdsourcing over two mainstream access technologies, DSRC and LTE. Based on our findings, we also summarize the challenges to large-scale real-time analytic of crowdsourced videos over vehicular networks.

1 INTRODUCTION

Modern vehicles are becoming more intelligent and fully connected by being equipped with onboard sensors and wireless communication devices [1]. Among onboard sensors, dash cameras can provide details of the surrounding environment outside the vehicle and are used as a way of capturing evidence. Besides, researchers have proposed to utilize the video data collected from vehicles for parking space detection, traffic monitoring [2], and assisted driving [3]. For many applications that require real-time analytic of high-resolution video streams, moving data to the cloud for processing is not feasible, due to the high latency caused by the roundtrip between vehicles and the cloud. Meanwhile, few vehicles today have sufficient computing power to conduct complex video analytic locally. Therefore, gathering and processing the video data from vehicles remains challenging.

Fog computing, shares the same principle with mobile edge computing [4], has been proposed to complement cloud computing by bringing computing resources and application services to the edge where the data is being generated and acted upon [5]. Different architectures of vehicular fog computing (VFC) that include vehicular cloudlets (i.e. cloudlets carried by vehicles) have been proposed in the literature. Hou et al. [6] proposed to turn vehicles, especially slow-moving and parked vehicles, into cloudlets, and to form a local cloud called JamCloud by gathering the computing resources available on nearby vehicles. Satyanarayanan [7] proposed to turn every vehicle (e.g. autonomous cars) into a cloudlet that has substantial processing capability and storage for processing local sensor data, and to share the processing results with the zone cloudlet in each coverage zone. Note that both architectures have not been evaluated in the real world.

To provide more reliable computing and communication services, we proposed in our previous work a different VFC architecture [8]. In our architecture, only the commercial fleets (e.g. buses and taxis) with predictable driving routes are turned into vehicular fog nodes and are utilized for providing cost-effective and on-demand fog computing for vehicular applications. These vehicular fog nodes, collaborating with the computing nodes co-located with cellular base stations (cellular fog nodes), serve the vehicles within the range for single-hop communication. Additionally, due to the mobility of commercial fleets, the service from vehicular fog nodes can be dynamically managed and dispatched on demand and the over-subscription of resources can be reduced.

In this paper, we focus on the feasibility analysis of real-time video crowdsourcing under our VFC architecture. We answer the following questions through vehicular traffic analysis and network simulation.

1) What would be the availability of vehicular fog nodes in urban areas?
2) What would be the serviceability of vehicular fog nodes, in terms of network performance of video crowdsourcing using current vehicular networking technologies (i.e. DSRC and LTE)?
3) What are the other challenges regarding the system reliability and resource efficiency?

In this article, we first give an overview of the architecture of VFC. Then, we describe example applications based on
video crowdsourcing. Next, we analyze the availability of vehicular fog nodes, in terms of the spatial-temporal distribution of buses, using Luxembourg SUMO Traffic (LuST) [9], which is built based on real traffic data in the city of Luxembourg. Furthermore, we simulate the scenario of VFC-based video crowdsourcing using SUMO [10] and VeinsLTE [11], and evaluate the networking performance with varying time points in 24 hours. Last, we summarize the challenges of deploying VFC in urban areas, such as interference and service interruption, and propose potential solutions.

2 VEHICULAR FOG COMPUTING

In this section, we define the related terms and give an overview of the process of video crowdsourcing in VFC.

2.1 Related Terms

Fog nodes We consider two types of fog nodes.

1) Cellular fog nodes: The computing nodes co-located with cellular base stations.
2) Vehicular fog nodes: The computing nodes carried by moving vehicles with on-board DSRC and LTE communication modules.

Client vehicles Vehicles that generate dash camera video are defined as Client vehicles.

Service zones Today, modern cities are fully covered by cellular networks. We divide an urban area into Service zones, and select a cellular fog node (Zone head) within the zone to manage and coordinate all the fog nodes. Vehicular fog nodes and client vehicles always inform the zone head when they enter or leave the zone, utilizing the existing cellular registration mechanisms.

2.2 Process of Video Crowdsourcing

Figure 1 illustrates the process of video crowdsourcing in VFC. The whole process consists of four operations.

2.2.1 Discovering fog nodes

In the initial stage, a client vehicle needs to figure out which vehicular fog nodes are located within its communication range. It broadcasts one-hop probe messages over DSRC, and collects responses from vehicular fog nodes. Any vehicular fog nodes that respond are included in the list of fog candidates.

2.2.2 Sending requests

After discovering fog candidates, the client vehicle sends a request to the zone head over LTE. The request contains information about the dash camera video to be transmitted to fog candidates.

- Dash camera video profiles: The description of the generated dash camera video, such as video length and supported video resolutions.
- Video generator: The client vehicle which generates dash camera video and sends the request.
- Fog candidates: The vehicular fog nodes within the client vehicle’s communication range.

2.2.3 Collecting dash camera video

When receiving a request from any client vehicle, the zone head assigns the vehicular fog node which is the closest to the client vehicle to collect the dash camera video.

2.2.4 Conducting vehicular application service

As illustrated in Figure 1, after the collection of the dash camera video, specific vehicular application (e.g. road construction detection and parking navigation) service would be conducted in the fog nodes.

3 VIDEO CROWDSOURCING APPLICATIONS

Modern vehicles are equipped with cameras, GPS, radar and other sensors. Compared with other sensors data, video captured by dash cameras can provide more details of the environment outside the vehicle. By collecting and sharing video from vehicles, VFC-based video crowdsourcing is emerging as a new paradigm for measuring and mapping phenomena of common interest [2]. However, crowdsourcing video from vehicles would generate a big volume of data, which can be better handled by fog computing from its new functionality. In this section, we introduce the promising video-crowdsourcing-based vehicular applications that can benefit from VFC and classify them into three categories based on their update cycles. We list the requirements of mentioned applications in Table 1.

3.1 Applications with Short Update Cycle

- Driving assistance: Emerging assist-driving applications include real-time situational awareness [12], cooperative lane changing, and see-through for passing [1]. As listed in Table 1, this group of applications demand for intensive computation with ultra-low latency.
- Local 3D map generation: Today autonomous vehicles rely extensively on high-definition 3D maps to navigate the environment. In practice, crowdsourced dash camera video can be utilized for creating and updating a 3D local traffic map that indicates the real-time locations and moving directions of neighboring vehicles. As listed in Table 1, this kind of application requires video from large amount of participants for being trained with limited time. Fog computing deployment is ideally suited for local maps use case due to the real-time and local nature of the information needed for accurate and augmented situational awareness of the road users.

These applications involve data-intensive and latency-sensitive computing tasks and have an extremely strict requirement for the validity period of crowdsourced video. Compared with pushing crowdsourced video to cloud, VFC can benefit from new architecture to better handle the big volume of video coming from vehicles.

3.2 Applications with Moderate Update Cycle

- Parking navigation: Circulating detour for seeking empty parking spots in a congested area may cause extra traffic
congestion and increase the risk of car accident [2]. Real-time parking information can save the parking time for drivers and improve quality of life in metropolitan areas. However, the information about vacant on-street parking slots is hard to be collected and published in real time. Compared with burying sensors under the pavement beneath of the slotted parking spaces, recognizing free parking spaces from crowdsourced video by object recognition technology is a more economic solution.

- **Road construction detection:** In some scenarios, such as terrible weather or serious car accident, specific segments of roads may need to be constructed. As illustrated in Figure 1, signs are set up to close lanes when temporary construction work is ongoing in a snowy day. In VFC-based video crowdsourcing paradigm, vehicular fog nodes can detect and publish the information about the location of temporary road construction in the local area to avoid traffic congestion.

Compared with applications described in Section 3.1, the above-described applications allow longer validity period of crowdsourced video. For example, the validity period of crowdsourced video for parking navigation may last for several minutes and that of for detecting temporary road construction may last for few hours. Furthermore, in these applications, local relevance is one of the important features and the utility lifetime of crowdsourced video is explicit. Compared with cloud computing, VFC can provide a more efficient and timely service with location-aware and location-specific functionality.

### 3.3 Applications with Long Update Cycle

- **Infrastructure improvement recommendation:** Road map and traffic regulation in a city may change over time. For example, a one-way street may be constructed to a two-way street and speed limit for a road may be reduced from 40km/h to 20km/h. Detecting and re-annotating these changes by road surveyors would consume huge amount of manpower and resources. However, the images and video recorded by dash cameras contain equivalent and detailed information of road infrastructures. By collecting and training the dash camera video across a range of computer vision tasks, the road attribute inference problem becomes amenable to a deep-learning approach. The trained models can handle infrastructure modification over time without re-annotation and can be employed to recommend improvements of city infrastructures.

- **Crime scene reconstruction:** Reconstruction of a crime as the basis of evidence is significantly important for law enforcement. However, criminals are likely to destroy the closed-circuit TV cameras in the crime scene. In such a scenario, a dash camera approach should be preferred because completely distributed opportunistic cooperation would make it very hard for potential attackers to disable surveillance.

Even though some applications allow long update cycle, they require a large amount of crowdsourced video as input. Thus, compared with cloud computing, processing data locally can reduce the ingress traffic to the cloud.

### 4 Feasibility

In this section, we explore the feasibility of VFC-based vehicular video crowdsourcing. First, we study the availability of vehicular fog nodes in urban areas by using Luxembourg SUMO Traffic scenario (LuST) [9]. Then, we simulate the scenarios of VFC-based real-time video crowdsourcing using VeinsLTE [11].

#### 4.1 Availability Vehicular Fog Nodes

The LuST scenario simulates the real traffic in the city of Luxembourg using SUMO [10]. The LuST scenario covers an area of 156 km², including 932 km of roads, 38 bus routes and 563 bus stops. To take a close look, we choose an area covering 0.25 km² acreage in the center of Luxembourg city. The latitude of that area ranges from 46°36′52″ to 49°36′53″, while the longitude ranges from 6°7′53″ to 6°7′54″. In our simulation, buses running in the selected area are configured as vehicular fog nodes that provide fog computing services for the surrounding vehicles. As shown in Figure 2, we run the simulation in 3 time windows, which are 6 : 00 ~ 6 : 10, 8 : 00 ~ 8 : 10 and 13 : 00 ~ 13 : 10. As shown in Figure 2, these points represent morning idle hour, morning rush hour and noon rush hour, respectively. We use a tuple (number of client vehicles, number of vehicular fog nodes) to denote the proportion between the number of client vehicles and that of vehicular fog nodes. The proportion is (85, 27), (200, 28) and (164, 25), respectively.

Dedicated Short-range Communication (DSRC) and LTE are the most popular vehicular networking technologies. DSRC is designed based on IEEE 802.11p. The data rate of DSRC can reach up to 27Mb/s with hundreds of meters coverage [13]. Compared with DSRC, LTE has a much wider coverage and more deterministic quality of service (QoS) guarantees. According to the reports in [14], LTE can support user equipment (UE) with high mobility at speed of 350 km/h.
In this paper, DSRC is responsible for data transmission between client vehicles and vehicular fog nodes. Additionally, we place a base station (cellular fog node) in the center of the selected area. In our simulation, each vehicle is equipped with both DSRC module and LTE stack. Each vehicle can stream video either to vehicular fog nodes via DSRC, or to a cellular fog node over LTE.

Figure 3 illustrates the distribution of distance from each client vehicle to its nearest vehicular fog node. 90 percentage of vehicles are located less than 200 meters away from the closest vehicular fog node, which means they can stream videos in real time to the vehicular fog node within short transmission distance. When the traffic becomes crowded (i.e. more than 200 client vehicles running in the selected area during 8:00 ~ 8:10), over 80 percentage of client vehicles can reach a vehicular fog node within 100 meters.

In our simulation, the client vehicles can be served by surrounding vehicular fog nodes during travel period when their distance is within one-hop DSRC range. We calculate the distribution of the proportion of served period in Figure 4. Over 90 percentage of client vehicles could receive fog computing services for more than 85 percentage of the traveling time. When traffic becomes crowded, more than 87.5 percentage of vehicles can be served for almost 100 percentage of the traveling time. Obviously, the served period increases with the congestion level of urban traffic.

4.2 Serviceability

In this section, we evaluate the performance of VFC-based real-time video crowdsourcing with simulation, and analyze the network performance of traffic conditions (morning idle hour, morning rush hour and noon rush hour) while using DSRC and LTE, respectively.

4.2.1 Simulation Configuration

We simulate the scenarios of VFC-based real-time video crowdsourcing with VeinsLTE [11]. VeinsLTE is an open source framework for running vehicle network simulation. VeinsLTE connects a microscope road traffic simulator SUMO [10] with a network simulation engine called OMNET++ through traffic control interface (TraCI). VeinsLTE includes a WAVE DSRC communication stack and a LTE-based vehicular communication stack. With VeinsLTE, vehicles in the simulation can either exchange data with each other through DSRC, or connect to the cellular fog node over LTE. Furthermore, we add an application module to VeinsLTE to implement the scenario of video streaming from vehicles to vehicular fog nodes and from vehicles to the cellular fog node. The client application on each client vehicle continuously stream video to nearby vehicular fog nodes or the cellular fog node which is located in the center of selected area. The application servers running on vehicular fog nodes and the cellular fog node take care of the processing of the received video. In this paper, we focus on the data transmission part, leaving the data processing out of scope.

To figure out whether VFC can enable real-time crowdsourcing of high-resolution videos, we compare the network performance when transmitting videos with various resolutions. We set the video resolution between 240p and 1080p, which is compressed with the Youtube-HD standard, and the frame rate as 24fps.

We measure the performance of video transmission using three metrics, which are end-to-end delay, packet loss ratio and throughput. We calculate the average end-to-end delay and packet loss ratio over the total number of flows in the network and compute the throughput on the server side at the end of the simulation.

4.2.2 Simulation Results

Table 2 lists the end-to-end delay of video transmission over DSRC and LTE. The end-to-end delay over DSRC remains low (several milliseconds) when the video resolution is not higher than 480p (640 * 480). However, the transmission latency over LTE stays at a high level, which is several seconds, with the same video resolution. When the video resolution is 720p (1280 * 720), the end-to-end transmission delay over DSRC increases to 5 seconds and that over LTE is more than 50 seconds. Both DSRC and LTE cannot support video streaming in VFC when the video resolution is higher than 720p.

Table 2 also illustrates the packet loss ratio of these two access technologies. We can see the packet loss ratio of DSRC is extremely high (more than 70 percentage) when the video resolution is 240p. This is because the number of DSRC channels is limited and the communication among large number of connected vehicles causes serious interference. Compared with DSRC, LTE has a better performance when the video resolution is less than 360p (480 * 360). However, when the video resolution is higher than 720p, both access technologies have a serious packet loss issue (more than 90 percentage).

We list the network throughput of two access technologies in Table 2. At 6 am, the network throughputs of DSRC and LTE increase with the video resolution as far as the video resolution is lower than 720p. However, when the video resolution is higher than 720p, both access technologies' network throughput drop dramatically due to the serious packet loss issue. At 8 am, the falling inflection point of network throughput is 480p. This is because at that time, compared with early morning, the traffic is more crowded and the communication interference is more serious.

According to the simulation results, DSRC performs well when the video resolution is low. But the packet loss ratio would rise up when the interference becomes serious. LTE achieves higher network throughput whereas the end-to-end transmission delay gets high. Furthermore, as illustrated in Section 4.1, vehicles in the urban area can get service from surrounding vehicular fog nodes for the most traveling time in a typical working day. In conclusion, most of vehicles can get guaranteed video crowdsourcing services over heterogeneous networks under the architecture of VFC.

5 CHALLENGES AND OPEN ISSUES

To provide VFC-based real-time video crowdsourcing with high performance and reliability, there are still several tech-
technical challenges to be solved. We discuss these challenges and potential solutions below.

5.1 Interference in V2V Communication Networks
Due to the interference in V2V communication networks, the packet loss rate of the DSRC connections increases dramatically with the number of connected vehicles, as shown in Table 2. The vehicles running the WAVE stack are supposed to be equipped with a single radio. However, by theory, they can communicate on all the 7 channels in the DSRC band. To eliminate channel competition, one solution is that each vehicular fog node dynamically assigns orthogonal channels to client vehicles. Since the number of available channels is limited, these channels will be spatially reused among vehicular fog nodes. In case some fog nodes are closely located, the cellular fog node (zone head) will help assign channels in a coordinated manner (through dynamic point selection, coordinated beamforming or scheduling) [15] to prevent from severe inter-fog interference.

5.2 Service Interruption
Vehicles on the street are moving with varying speed. As a result, the distance between vehicles is changing all the time. When the client vehicles leave the communication range of a connected vehicular fog node, the ongoing fog computing services such as video streaming will be interrupted. As the driving routes of vehicular fog nodes are predictable, it is possible to predict the service availability and to plan the migration of the service to another vehicular fog node. For example, videos can be forwarded to zone heads when the vehicles are driving away from a vehicular fog node.

5.3 Resource Management
In practice of VFC deployment, it is important to coordinate and manage all the vehicular fog nodes in an efficient manner. The resource management strategies must handle two issues. Firstly, one vehicle may be located within the communication ranges of several vehicular fog nodes. In this case, it is a question about how to select the most suitable vehicular fog node for the vehicle. Secondly, vehicular fog nodes are distributed in a distributed manner. They may not be able to communicate directly with each other due to a long distance. It would be more feasible to deploy resource management modules on cellular fog nodes, and utilize the wired connections between cellular fog nodes and the cellular connectivity of vehicular fog nodes to schedule the service deployment.

6 CONCLUSION
In this paper, we study the applications, feasibility and challenges of implementing real-time video crowdsourcing under the VFC framework in urban areas. We assess the availability of potential vehicular fog nodes based on a set of real-world-based SUMO traffic scenario [9]. Then, we test VFC-based video crowdsourcing in real-world-based vehicular traffic network scenarios, in terms of the network latency, packet loss ratio and throughput. The findings are summarized below:

1) Our experimental results prove the feasibility of enabling real-time video crowdsourcing with VFC. Concretely, approximately 90 percentage of the vehicle clients can reach the VFC network in a typical working day.
2) The impact of varying traffic conditions and high-quality video transmission affects severely the system performance and reliability of LTE networks, while the proposed VFC network represents a viable solution.
3) Our study reveals and reports relevant technical challenges to be solved, namely interference in V2V communications, computing service interruption and management.

In the future, we will explore the resource allocation in VFC with the consideration of service QoS and fog node mobility. Furthermore, the business model and pricing mechanism of VFC service also need to be explored. Additionally, the D2D and WLAN-based approaches co-exist and co-scheduling of various access technologies must take into account. We will implement a VFC testbed and test the network performance in a real 5G-enabled vehicle networking environment.

REFERENCES
Chao Zhu (chao.1.zhu@aalto.fi) received the BE degree in Computer Science and Technology from BUPT, China, in 2012 and MS degree in Computer Technology from Tsinghua University, China in 2016. He is currently a Ph.D. candidate in both Department of Computer Science and Department of Communications and Networking in Aalto University, Finland. His research interests include mobile edge computing and vehicular networking.

Giancarlo Pastor (giancarlo.pastor@aalto.fi) is a Postdoctoral Researcher at Aalto University and Adjunct Professor at IE Business School. He received his Ph.D. degrees in Electrical Engineering from Aalto University, Carlos III University of Madrid and King Juan Carlos University in 2016, B.Sc. degree in Mathematics from Autonomous University of Madrid in 2010, and B.Sc. degree in Electronics Engineering from Pontifical University Catholic of Peru in 2006. His research interests include tensor completion and technology-based entrepreneurship.

Yu Xiao (yu.xiao@aalto.fi) received the doctoral degree in computer science from Aalto University in 2012. She is currently an assistant professor at the Department of Communications and Networking, Aalto University. Her current research interests include mobile crowdsensing, augmented reality and edge computing.

Antti Yli-Jääski (antti.yla-jaaski@aalto.fi) received the PhD degree from ETH Zurich in 1993. He has been a professor for telecommunications software, Department of Computer Science, Aalto University since 2004. His current research interests include green ICT, mobile computing, and service architecture.
Fig. 1: Video crowdsourcing based vehicular applications
Fig. 2: Traffic demand in the LuST scenario
Fig. 3: CDF of distance from nearest vehicular fog node
Fig. 4: CDF of served period proportion