Crowd V-IoE: Visual Internet of Everything Architecture in AI-Driven Fog Computing

Wen Ji, Bing Liang, Yuqin Wang, Rui Qiu, and Zheming Yang

Abstract

Fog computing has emerged as a unifying platform to provide computing, communication, and storage for a variety of mobile applications. That helps achieve high bandwidth, high intelligence, low latency, and low energy consumption in handling massive networking devices and emerging rich multimedia services in 5G networks. Current prominence and future promises are changing from the Internet of Things (IoT) to the Internet of Everything (IoE), which is a union of people, process, data, and things. However, the development of fog radio access networks (F-RANs) is challenged by the diversity of IoE, ultra-high-definition videos on demand from users, and low-latency requirement of heterogeneous IoT devices. In this article, we present an architecture of visual IoE (V-IoE) in F-RANs. We systemically analyze the key challenges of V-IoE from the perspective of F-RANs, and propose a crowd V-IoE architecture. Through experimental results, we demonstrate that our proposed architecture exhibits better performance with lower bandwidth requirement, lower energy consumption, and lower latency in F-RANs. Finally, we conclude with a discussion of potential directions.

Introduction

With the rapid development of communications and networks, billions of mobile users and devices receive seamless and stable wireless services supported by the communication infrastructures. In 2012, fog computing was introduced by Cisco to exploit a wide breed of IoT applications and services at the edge of the network [1]. Fog computing presents a highly virtualized platform that integrates compute, storage, and networking services between end devices and cloud data centers. Compared to cloud radio access networks, F-RANs, based on fog computing architectures, exhibit better utilization of edge devices due to the collaborative signal processing and computing [2, 3]. F-RANs are becoming a popular technology for mobile networks [4].

However, recent reports have shown that F-RANs faced three challenges. First, it is hard for the distributed architecture that uses edge devices of F-RANs to cope with the heterogeneity of massive IoT devices. IoT has become a prevalent system in which devices and machines connect to the Internet and each other. According to Cisco's forecast, there will be approximately 14.6 billion machine-to-machine (M2M) connections and 28.5 billion networked devices by 2022. On the other hand, about three billion users worldwide surf on the Internet through mobile devices each day, which forms an "Internet of Everyone" [5]. A broader concept, "Internet of Everything" (IoE), denoted as the intelligent connection of people, devices, data, and things, is replacing conventional IoT architecture [6]. IoE brings together people, process, data, and things to make networked connections more relevant and valuable [7]. Current 5G promotes the whole world to the era of IoE through technological innovations on connectivity, rates, latency, energy, and utility [8]. How to integrate communication-computation-storage resources to support increasing IoE requirements challenges the future directions of F-RANs.

Second, multimedia data is becoming the dominant traffic in communication networks. According to Cisco's forecast, by 2022, video traffic will account for 82 percent of all IP traffic and will reach up to 79 percent of mobile data traffic. F-RANs have deployed a fog-cloud structure to improve the scalability. This structure consists of a vast number of large-scale IoT devices and highly capable end-user devices, namely, real-time multimedia transmission in IoE systems. However, IoE video communication over F-RANs is difficult as it requires supporting stringent requirements in limited capability (e.g., low power and insufficient bandwidth) of end systems, and it requires managing stable and reliable transmission among both variable numbers of devices and diverse numbers of user requirements [9]. How to support efficient distributed computation, storage, and delivery of large-volume multimedia over F-RANs will highlight cooperative performance for high definition, virtual reality (VR), augmented reality (AR), and 360° multimedia processing applications in the future fifth generation (5G).

Third, artificial intelligence (AI) is becoming one of the major factors in the information and communication technologies. AI impacts F-RANs from architectures to algorithms. In view of F-RAN applications, AI provides advanced and intelligent management among massive homogeneous and heterogeneous devices, which are connected through F-RAN controllers. The capacities of transmitting large-volume data, optimizing the diverse resources, offloading computation, and caching

This work is supported by the National Key R&D Program of China (2017YFB1400100), the National Natural Science Foundation of China (61572466), and the Beijing Natural Science Foundation (4202072).

Wen Ji (corresponding author) is with Institute of Computing Technology, Chinese Academy of Sciences and Peng Cheng Laboratory; Bing Liang, Yuqin Wang, Rui Qiu, and Zheming Yang are with the Institute of Computing Technology, Chinese Academy of Sciences, and the University of Chinese Academy of Sciences.

Digital Object Identifier: 10.1109/MWC.001.1900349

1536-1284/20/$25.00 © 2020 IEEE

The diverse performance of massive IoT devices differ in terms of capacity, computation, energy-supply, storage, sizes, speed, and visual sensors. Hence, it is urgent to design flexible, distributed, and cooperative schemes for managing massive and heterogeneous IoT-centric communications efficiently under the constraint of ultra-low latency.

frequent access contents require more intelligent methods. AI allows IoE to perceive and learn the human world through pattern recognition and machine learning methods, and further, harvest energy from ambient radio frequency signals [10]. Thus, machines can respond, communicate, and control the environment as humans do. To understand the intrinsic features of complex multimedia data, transmit various types of data efficiently, predict available resources accurately, respond to the environment proactively, and decide tasks smartly still challenge the design for the architecture of F-RANs.

In this article, we put forward an AI perspective on visual IoE designs for F-RANs, and propose a communication-computing-caching integrated framework together with deep learning. In most applications where the requirements of IoE devices are highly heterogeneous in terms of bandwidth, latency, and power consumption, existing F-RANs equipped with baseband unit (BBU), high power node (HPN), remote radio head (RRH), and core networks cannot provide a satisfactory solution. Thus, F-RANs with an intelligent design are desired. To deal with these applications, we address the video data and the massive IoE devices to design a comprehensive AI-enabled F-RAN. Our objective is to design a flexible architecture with efficient utilization among communication, computing, and storage resources. Particularly, we exploit the nature of AI. We present an IoE-profile representation method, where both complex human-centric transmission and diverse IoT devices access F-RANs through IoE-profiles are examined. We change conventional multimedia streaming to triple-metric feature-type transmission, where the compressed video is replaced by hybrid semantic, feature, and small compressed videos. The bit rate of transmitted data is sharply reduced. Then massive IoE devices are intelligently managed through crowd control.

The remainder of this article is organized as follows. Current technologies and service platforms for F-RANs are first reviewed. Key features and challenges are analyzed. The architecture with four core components of V-IoE in F-RANs is proposed. We give the numerical results. Finally, we unveil the open issues and challenges in future research.

Key Features and Challenges of V-IoE in Fog Computing

According to the latest reports by Cisco and Gardener, an unprecedented number of IoE devices as well as a large amount of demands for rich videos over 5G networks will dominate the future technical tendency of communications and networks. To cope with these challenges, we explore an AI-driven V-IoE architecture over F-RANs. V-IoE is aimed at managing and dispatching complex multimedia contents closer to end users intelligently and quickly by exploring the capabilities of intelligence, storage, computation, and communication. Future V-IoE communication includes hybrid human-centric and IoT-centric cooperative transmission systems. Consequently, the V-IoE architecture over F-RAN is expected to provide the following advantages for the end-IoE, service providers, and content providers.

For human-centric communications: The current communication network has evolved into human-centric rich multimedia access services. Next-generation content distribution networks with many more considerations of video traffic have developed an "edge-cloud" structure for improvement of scalability. Ultra-high-definition (UHD), VR/AR, 360° videos, as well as intensive access challenge current seamless user experience through both F-RANs and core networks. It is necessary to reduce the network backhaul bandwidth requirements through service localization, to accelerate caching and local offload, to develop on-demand user-based content distributions, and to alleviate the pressure of data transmission in core networks.

For IoT-centric communications: High-reliability and low-latency scenarios are the main typical scenes in F-RANs, where data transmission is sensitive to delay and on-demand reliability, such as intelligent transportation and telemedicine. For example, the end-to-end delay requirement in recent autonomous unmanned aerial vehicles (UAVs)/unmanned ground vehicles (UGVs) applications is less than 1 ms, which poses a challenge to existing F-RAN architectures. Consequently, it is necessary to design a network architecture from the viewpoints of network layer technology (e.g., caching and slicing) and cross-layer technology (e.g., joint control in transmission and application layers), which will be easily deployed on fog access points (F-APs), and fog user equipments (F-UEs) before accessing the Internet or performing remote cloud computing. The traditional computation and transmission burdens are migrated from cloud sides to edge sides [11].

Content caching and prediction benefit to reduce network end-to-end delay among massive IoT devices and to improve both quality of service (QoS) of IoT and quality of experience (QoE) of users. Meanwhile, V-IoT over heterogeneous networks is becoming a new trend in future networks. However, for the V-IoT, those problems in bandwidth and power consumption could become even worse. Moreover, the diverse performance of massive IoT devices differ in terms of capacity, computation, energy supply, storage, sizes, speed, and visual sensors. Hence, it is urgent to design flexible, distributed, and cooperative schemes for managing massive and heterogeneous IoT-centric communications efficiently under the constraint of ultra-low latency.

For AI-driven IoE-centric communications: In F-RANs, a large amount of signal processing and computing is performed in a distributed manner through efficiently managing a fronthaul centralized BBU pool, F-APs, and F-UEs. However, UHD/VR/AR video streaming in current IoE designs still have to conquer the following four problems.

- Concerning bandwidth, complex video sources are diverse in terms of rates, resolutions, frame rates, color information, content, depth, and so on. Not only available large-volume bandwidth requirements, but also complex characteristics of video sources should be considered. AI technologies including content analysis, data mining through machine learning, accurate prediction, and optimal decision need to assist F-APs and RRH to manage all video tasks smartly and automatically.

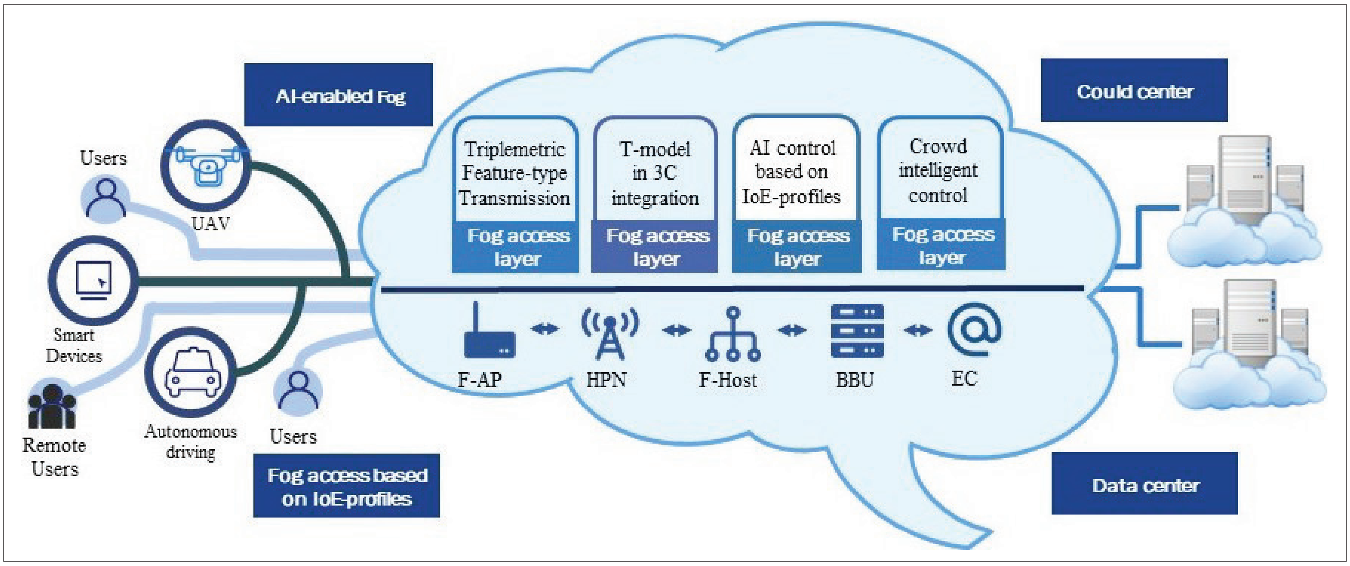

FIGURE 1. Architecture of the proposed AI enables F-RANs.

- Concerning delay, large-volume data results in high delay inevitably. The most difficult but key method to resolve the delay problem is to reduce data size. AI methods benefit data transmission by transmitting features instead of original compressed multimedia data. Since the data volume of features is only 1–70 percent of conventional compressed video data, this means that feature-based architectures might expand the utilization of current networks up to 10 times compared to current data-based networking. However, how to extract suitable features to express the video data in F-RANs depends on AI technologies.
- Concerning heterogeneity, transmitting video data among IoEs through F-RANs is influenced by three aspects. From the source side, video data is affected by acquisition sensors, access capability, available bandwidth, and computation capacity. From the relay side, the processing depends on caching sizes, jitter frequencies, priority, services reservation, scheduling schemes, multiplexing, and fairness. From the end side, the final effect has a close relation with objects observed by human or IoT devices, displays, frozen time, operation modes, quality, interaction modes, energy saving modes, and batteries. As a consequence, to understand how a human processes complex information and interacts appropriately with multiple IoEs plays an important role in processing video data in F-RANs automatically and intelligently.
- Concerning environments, ambient intelligence is tightly integrated with the application scenarios. When all the above technologies are applied to homes, offices, stores, hospitals, and transportation services, the infrastructure and density of deployment in F-RANs directly affect the performance of services. AI could assist IoE in understanding the presence and context through recognition, change the response according to the identified activities, and implement approximate reasoning processes to achieve adaptations. AI based on machine learning could be effective by extracting the knowledge from data and learning the actual and desired requirements from environments. However, how to exploit the profiles of IoEs in complex F-RANs is still open.

System Architecture and Key Techniques

Crowd V-IoE is a hieratical AI-driven fog computing architecture, which is composed of V-IoE nodes that access F-RANs through IoE profiles, fog equipments, cloud centers, and data centers. Through four fundamental components, F-RANs evolve into an AI-enabled fog, as shown in Fig. 1. The fog computing access point (F-AP) usually uses both WiFi and 5G connections to provide low energy consumption, low latency, and a high-bandwidth wireless connection to V-IoE devices. The edge computing (EC) includes edge storage and edge servers. Edge servers are used to respond to the requests from V-IoE nodes, to predict the next behavior and operation of users or IoT devices, to preprocess complex analysis for frequent request contents, to implement optimal caching and scheduling strategies, to balance both fronthaul and backhaul transmission channels, and to extract and analyze features through online machine learning. The above operations are virtualized into fog access layers. The objective of a fog access layer is to manage V-IoE nodes, available resources, and heterogeneous connections efficiently.

V-IoE nodes represent geographical distributed users and IoT devices. The difference between those two types of nodes is that the receiver of the former is humans for user ends, whereas the terminal of the latter is IoT devices. Although users access the Internet through smartphones, laptops, and so on, the videos transmitted to users require high QoE, high bandwidth, and low latency. However, when multimedia data is transmitted to IoT devices, the IoT collects the necessary information to implement analysis, response, and control. It implies that the specific feature instead of the whole multimedia data is the true useful data for IoT devices. Consequently, feature data plays an active role in providing high-quality ser-

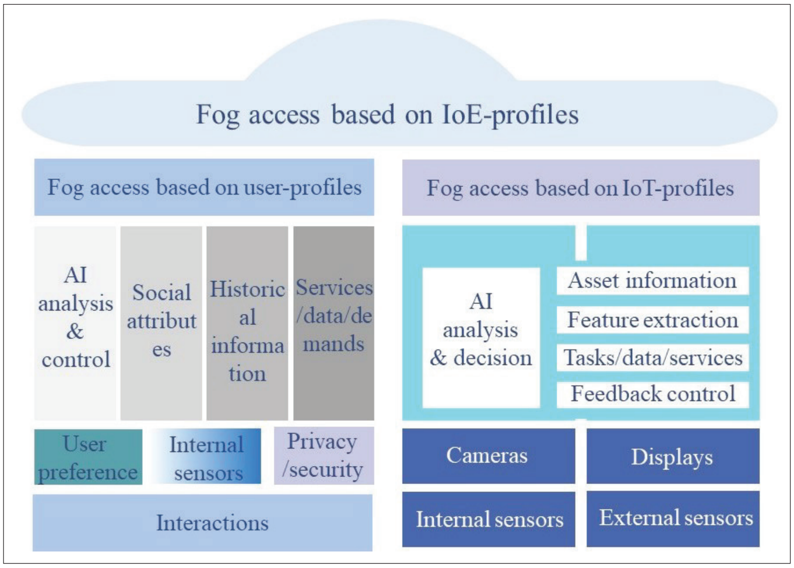

FIGURE 2. Fog access based on IoE-profiles.

vices for IoT devices, such as autonomous vehicles, UAVs, video surveillance, and event alert in smart cities.

Heterogeneous networks represent a typical F-RAN architecture, where IoE is connected to specific F-APs, remote radio heads (RRHs), or high power nodes (HPNs) through fog access management. The role of fog access management is a highly virtualized layer that provides intelligent computing, storage, and networking services between IoE devices and conventional network infrastructures. Moreover, the F-AP is a characteristic of F-RANs because it implements collaborative radio signal processing by integrating both fronthaul and backhaul functionalities [2].

Cloud centers and data centers provide large-scale data mining and analysis through machine learning. In particular, the cloud side provides a complex visual data management platform, where the multimedia data is compactly analyzed in terms of data and features. In traditional architectures, multimedia data from UE is directly transmitted to cloud servers to perform complex data analysis. Relays, HPNs, and RRHs emphasize signal transmission and enhancement, but seldom analyze and compute the inner hidden information. And the data analysis is usually implemented on the cloud server side. Instead, in the proposed architecture, the cloud in the form of mobile distributed clouds or conventional centralized clouds provides advanced multimedia analytical services.

The definition and characteristics of the proposed architecture are:
- Flexible fog computing access layers through profile representation
- Optimal integration among communication-computing-caching resources
- Multistreams including compressed multimedia data, feature data, and semantic data
- Crowd intelligent control among massive IoE

These characteristics are elaborated to make the F-RAN a more intelligent architecture for a large number of IoEs in big visual data services and applications.

Flexible Fog Computing Access Layer through Profile Representation

The proposed crowd V-IoE in F-RANs introduces profiling and a comprehensive adaptation mechanism among V-IoE nodes, fog infrastructures, cloud, and data centers. Complex multimedia contents provided for the IoE could be cross-controlled and adjusted by parameters of streams and service modifications in physical network components. In F-RANs, we present an IoE profile-based strategy to improve the intelligence of data processing and interaction responses. An IoE profile specifies the set of feature representation to both a user's characteristics and IoT devices. IoE profiles address the mobile fog access needs in F-RANs.

Based on IoE profiles, the fog access among crowd devices has the ability to understand the complex relationship in complex devices and datasets through learning and analysis. In detail, IoE profiles include two parts: user-profiles and IoT-profiles, as shown in Fig. 2.

Fog access based on user-profiles responds to the interactions between a user and corresponding devices. Through analyzing the social attributes of users and historical information, the device could learn users and evolve itself to understand the served users. IoT-profiles describe the adaptable entity, current contexts, and diverse circumstances. The difference of proposed IoT-profiles is that first, feature extraction is introduced to provide extensibility on semantic representation. Visual features including objects, texture, structures, focused regions, and other bitstreams, are extracted at IoT devices for preliminary analysis, and then encoded and transmitted with original compressed multimedia bitstreams to improve the discriminative capability. Second, AI analysis and decision are introduced to implement feature extraction and improve the automotive interoperability.

Triple-Efficiency Model (T-Model) in Communication-Computing-Caching Integration

To facilitate massive multimedia data delivery, and to fit complicated network conditions and various mobile devices, the integration among communication-computing-caching is a promising and feasible technique to realize efficient content transmission in F-RANs. Conventional methods mainly discuss the integration through separate evaluation [12]. However, what is the role of the integration, if any, in measuring the efficiency? We propose a triple-efficiency model (T-model) to answer this question.

We notice that the common characteristic of communication-computing-caching is data. Crucially, the only difference between tripartite is processing. Each party with similar productivity has indistinguishable efficiency, such as data processing capacity per unit time. Efficiency varies greatly among communication-computing-caching. Assume that the existence of a statistical function $U^{comm}(r)$ that modulates the efficiency of communication component, which has a unique value in terms of data processing capacity per unit time for a file of length r. That is $c^{comm} = U^{comm}(r)$. Similarly, we have the statistical representations of computing and caching components; there are

$c^{comp} = U^{comp}(r)$ and $c^{cach} = U^{cach}(r)$. The current architecture design in both computer science and communication disciplines implicates the methodology of parallel pipeline processing. The nature indicates the presence of multiplicative processes, which motives us to write the efficiency metric by the integration of communication-computing-caching as $C^{Tri} = U^{comm}(r) \cdot U^{comp}(r) \cdot U^{cach}(r)$. The metric C^{Tri} captures the efficiency of integration to take advantage of communication, computing, and caching in parallel, speeding up the process.

Triple-Metric Feature-Type Transmission: Compressed Multimedia, Features, and Semantic Streams

Current IoE devices are gradually evolving to be highly autonomous and intelligible, providing convenience for upper layer processing. For example, the data prediction and decision modules analyze the data only from features instead of the original whole multimedia data. This characteristic subverts traditional transmission architectures. As shown in Fig. 3, we propose a triple-metric feature-type transmission framework to replace the conventional multimedia streaming.

Triple-metric feature-type transmission provides three types of streams:

1. Semantic data. It describes semantic and information-aware notifications of events, interests, focuses, extraction labels, and other IoE-based preprocessing.
2. Feature data. It represents local neighborhood operations applied to image data and provides a simplified but discriminant form for the information extraction and postprocessing of recognition, prediction, or other AI operations.
3. Compressed data. They are the original compressed multimedia data.

The advantages of triple-metric feature-type transmission are described as follows. First, the volume of semantic and feature data occupies only less than 1 to 30 percent of compressed multimedia data, which sharply reduces transmission bandwidth. Second, conventional feature extraction methods for video data are computed in pixel levels. This means that if the recognition or prediction operation is implemented on remote server sides, the operations could be processed if and only if the system could decode all the compressed videos in pixels. This redundant processing wastes a large amount of computation and caching resources.

Third, current video streams employ lossy compression, which reduces the accuracy of data mining. Consequently, triple-metric feature-type transmission shows better prospects in view of communication, computing, and caching aspects.

Crowd Intelligent Control among Massive IoE Devices

Current F-RANs are composed of massive IoE devices and 5G/Internet infrastructures for a large amount of applications in smart environments. The fog access virtual layer is introduced to support hybrid control by incorporating IoE devices in a crowd manner.

We capture the ability among communication-computing-caching and extract the crowd intelligence derived from empirical measurements. IoE devices are tasked to seamlessly carry out operations. They exhibit a level of integrated intelligence in specific tasks, such as large-scale transportation management in smart cities. An innovative crowd-intelligent detector is introduced in our F-RANs, where the detector considers the dynamics of context and exploits tangible F-RANs to infer the real control results that could realistically affect visual data transmission among F-RANs. In detail, we transfer the intelligent analysis in conventional mobile clouds to edge and IoE sides. The intelligence is decentralized to distribute crowd intelligence. The decisions could be immediately processed on IoE nodes, instead of the reactions from remote clouds. The latency is roughly reduced because of the unnecessary redundant communication and computation are shaved off.

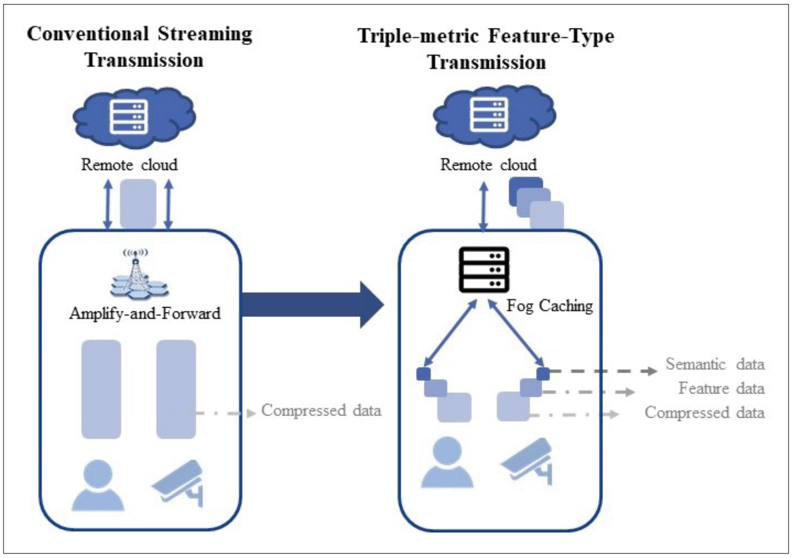

FIGURE 3. Triple-metric feature-type transmission.

Numerical Results

This section evaluates the performance of video transmissions in the proposed AI-enable F-RAN system. The experiment includes three parts: transmission bit rates in different scenarios, energy consumption of terminal equipment, and the evaluation of transmission on a typical F-RAN.

Triple-Metric Feature-Type Transmission in Different Scenarios

As mentioned in the previous section, we design a triple-metric freature-type transmission model to further reduce the transmitted data. We examine the performance of the proposed transmission method in an intelligent transportation system. Experiments are carried out based on the open Cityscapes dataset [13]. In detail, 19 classes of features are extracted from the original 30 classes of features in the dataset. We introduce deep learning, namely DeeplabV3+ [14], to perform feature extraction. Traffic scenes are analyzed to extract five types of features, including vehicle, pedestrian, traffic sign, traffic lights, and road scenes. As shown in Fig. 4, experimental results indicate that the bit rates of the proposed fea-

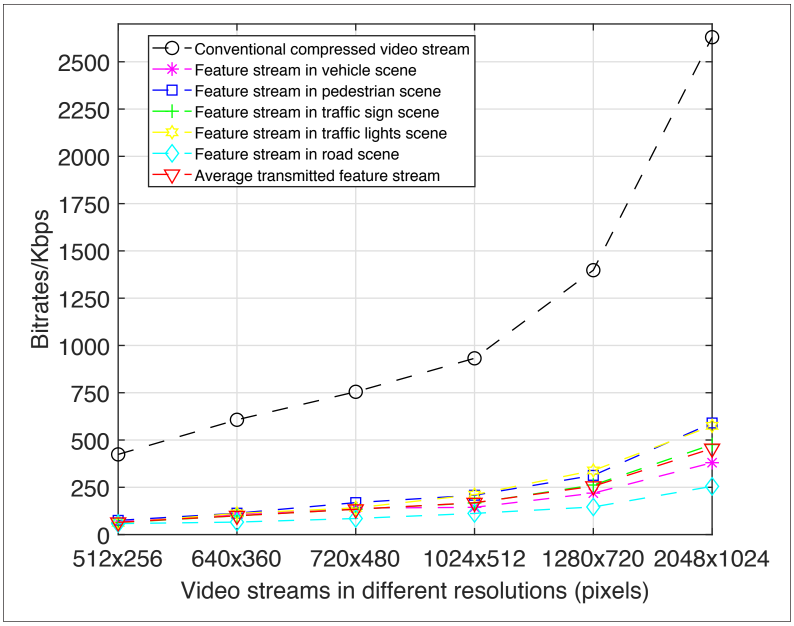

FIGURE 4. Transmitted bit rates between triple-metric feature-type transmission and conventional compressed video transmission in different scenes.

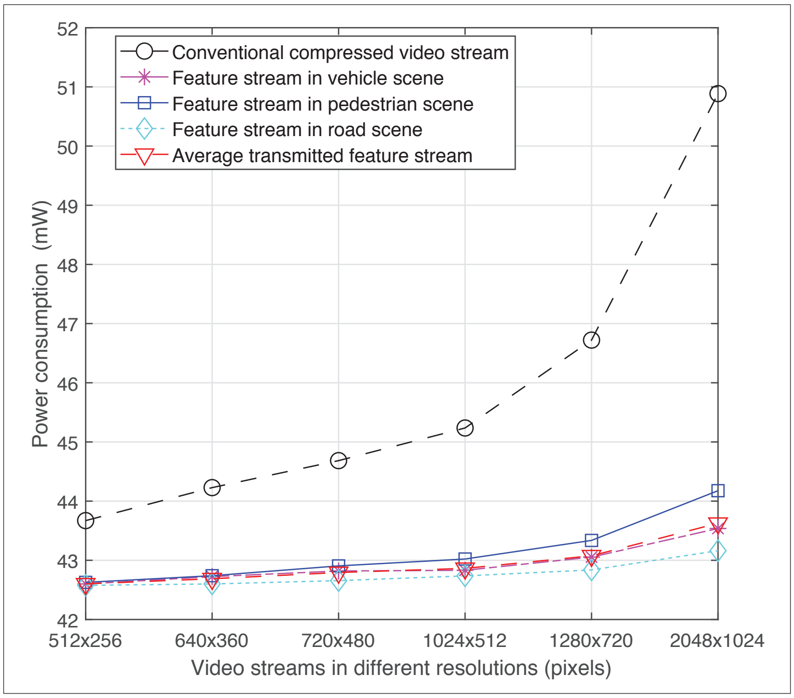

FIGURE 5. Energy consumption of IoE devices during transmitted video streams.

ture-type method are decreased by 70 percent on average compared to those of the conventional compressed video stream.

Energy Consumption of IoE Devices

This experiment evaluates the energy consumption when the proposed method is applied to video transmission. The experimental test is conducted in a wireless communication environment. This test assumes that the distribution of devices follows Poisson point process, and the short-term fading channel follows Rayleigh distribution.

Figure 5 shows the experimental results of energy consumption when video streams in different resolutions are applied. Observations show that compared to typical transmission, the proposed method can reduce power consumption of IoE devices. Furthermore, when video streams are in higher resolutions, the energy consumption by the proposed method does not grow as much as the typical method does. The results demonstrate that feature-type transmission is not affected by resolutions. The proposed method can save energy effectively and continuously. Moreover, the proposed triple-metric feature-type transmission exhibits better performance, particularly in UHD video applications.

Evaluation of Integration in a F-RAN

We validate the effectiveness of the proposed crowd V-IoE architecture presented in previous sections in a typical F-RAN. The Fog environmental setting is based on [15]. We evaluate the performance of integration among communication, computing, and caching by using the proposed T-model mentioned earlier. As shown in Fig. 6, the ability of data processing in terms of log C^{Tri} is the highest by using the proposed triple-metric feature-based transmission, such that the speeds of computation, transmission, and caching in a fog node are fastest. In contrast to conventional multimedia transmission, the proposed crowd V-IoE architecture effectively enhances the ability of data processing in an F-RAN.

Conclusions and Future Work

In this article, we propose a visual IoE architecture for AI-driven F-RANs that leverages the massive IoE devices and large-volume multimedia data. The advantages of this architecture include building an IoE-profile representation to manage fog access, changing conventional multimedia streaming to feature-type transmission, and presenting a triple-metric model to speed up the efficiency of communication-computing-caching integration. The proposed crowd V-IoE architecture can lower the bit rates of transmitted data and accelerate the ability of data processing of a typical F-RAN.

The V-IoE opens multiple future research directions related to future F-RANs. First, video data has become the dominant traffic. However, the exploding growth of data traffic is far beyond the growth of capacity. Some questions are how multimedia data should be employed in communication systems, what kind of quantitative analysis for delays, complexity, and transmitted power, and what metrics the designers can use in practical applications. All the questions should be conducted under realistic constraints, such as those found in high-speed F-RANs. Second, today's multimedia services and communication practices are tightly coupled with economic considerations. Economic models have the potential to generate benefits for both users and operators. Monetizing video traffic efficiently is leading the research tendency in both the academic and industrial fields in future F-RANs. Third, the rapid surge of users' generated contents and online social networks results in a large amount of social multimedia. Social multimedia is different from traditional multimedia in the sense that social relationships and user preferences mainly affect transmission routes, distributions, and storage. When users' social characteristics are considered in transmission design, latency for the social IoT can be further reduced. Consequently, more deep research on social-aware multimedia processing, particularly in transmission, distributions, and storage designs in F-RANs, is also highlighted.

References

[1] F. Bonomi *et al.*, "Fog Computing and Its Role in the Internet of Things," *Proc. ACM MCC.*, Aug. 2012, pp. 13–16.

[2] H. Zhang *et al.*, "Fog Radio Access Networks: Mobility Management, Interference Mitigation, and Resource Optimization," *IEEE Wireless Commun.*, vol. 24, no. 6, Dec. 2017, pp. 120–27.

[3] M. Peng, S. Yan, and H. Poor, "Ergodic Capacity Analysis of Remote Radio Head Associations in Cloud Radio Access Networks," *IEEE Commun. Lett.*, vol. 3, no. 4, Apr. 2014, pp. 365–68.

[4] M. Peng *et al.*, "Fog-Computing-Based Radio Access Networks: Issues and Challenges," *IEEE Network*, vol. 27, no. 5, July–Aug. 2016, pp. 46–53.

[5] G. Cerf and N. Quaynor, "The Internet of Everyone," *IEEE Internet Computing*, vol. 18, no. 3, May/June 2014, pp. 96–96.

[6] A. Banafa, "The Internet of Everything (IoE)"; https://www.bbvaopenmind.com/en/technology/digital-world/the-internet-of-everything-ioe/, 2016.

[7] D. Evans, "The Internet of Everything: How More Relevant and Valuable Connections Will Change the World"; https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.html, 2012.

[8] P. Szilágyi, "5G for People and Things: Key to the Programmable World," *Proc. UCAAT*, Oct. 2016, pp. 13–16.

[9] W. Ji *et al.*, "Visual IoT: Enabling Internet of Things Visualization in Smart Cities," *IEEE Network*, vol. 33, no. 2, Mar./Apr. 2019, pp. 102–10.

[10] M. Min *et al.*, "Learning-Based Computation Offloading for IoT Devices with Energy Harvesting," *IEEE Trans. Vehic. Tech.*, vol. 68, no. 2, Jan. 2019, pp. 1930–41.

[11] M. Peng and K. Zhang, "Recent Advances in Fog Radio Access Networks: Performance Analysis and Radio Resource Allocation," *IEEE Access*, vol. 4, Aug. 2016, pp. 5003–09.

[12] Y. Zhou *et al.*, "Communications, Caching, and Computing for Next Generation HetNets," *IEEE Wireless Commun.*, vol. 25, no. 4, Aug. 2018, pp. 104–11.

[13] M. Cordts *et al.*, "The Cityscapes Dataset for Semantic Urban Scene Understanding," *Proc. CVPR*, 2016, pp. 3213–23.

[14] L.-C. Chen *et al.*, "Encoderdecoder with Atrous Separable Convolution for Semantic Image Segmentation," *Proc. ECCV*, 2018, pp. 801–18.

[15] T. Dang and M. Peng, "Joint Radio Communication, Caching and Computing Design for Mobile Virtual Reality Delivery in Fog Radio Access Networks," *IEEE JSAC*, vol. 37, no. 7, July 2019, pp. 1594–1607.

Biographies

Wen Ji [M'09, SM'16] (jiwen@ict.ac.cn) is currently a professor with the Institute of Computing Technology (ICT), Chinese Academy of Sciences (CAS), Beijing, and the Peng Cheng Laboratory, Shenzhen, China. She received her M.S. and Ph.D. degrees in communication and information systems from Northwestern Polytechnical University, China, in 2003 and 2006, respectively. From 2014 to 2015, and in 2018, he was a visiting scholar at the Department of Electrical Engineering, Princeton University, New Jersey. Her research interests include multimedia communication and networking, video coding, channel coding, information theory, optimization, network economics, and ubiquitous computing. She holds over 20 patents and has more than 80 refereed publications in journals, conference proceedings, and book chapters related to her areas of interest. She served or serves on various Editorial Boards, including as a Guest Editor for *ACM Transactions on Embedded Computing Systems* and the *IEEE Journal on Selected Areas in Communications*.

Bing Liang (liangbing@ict.ac.cn) received his B.Sc. degree from Heilongjiang University, China, in 2014 and his M.S. degree in electronics and communications engineering from North China Electric Power University, China, in 2017. Currently, he is pursuing a Ph.D. degree at the Institute of Computing Technology (ICT), Chinese Academy of Sciences (CAS), Beijing. His current research interests include multimedia communication and networking, video transmission, game theory, and machine learning.

Yuqin Wang (wangyuqin@ict.ac.cn) received her B.Sc. degree in computer science from Wuhan University, China, in 2017. Currently, she is pursuing an M.Sc. degree in computer application technology at the Institute of Computing Technology (ICT), Chinese Academy of Sciences (CAS). Her current research interests include multimedia communication and networking, video transmission, and edge computing.

Rui Qiu (qiurui@ict.ac.cn) recevied his B.Sc. degree in radio and television engineering from Communication University of China in 2019. He is currently working toward an M.Sc. degree in computer application technology at the Institute of Computing Technology(ICT), Chinese Academy of Sciences(CAS). His research interests include point cloud compression, machine learning, and computational advertising.

Zheming Yang (yangzheming@ict.ac.cn) received his B.Sc. degree in electronic engineering from North China University of Science and Technology in 2019. Currently, he is pursuing a Ph.D. degree at the Institute of Computing Technology (ICT), Chinese Academy of Sciences (CAS). His current research interests include intelligence measurement, video transmission, and network economics.

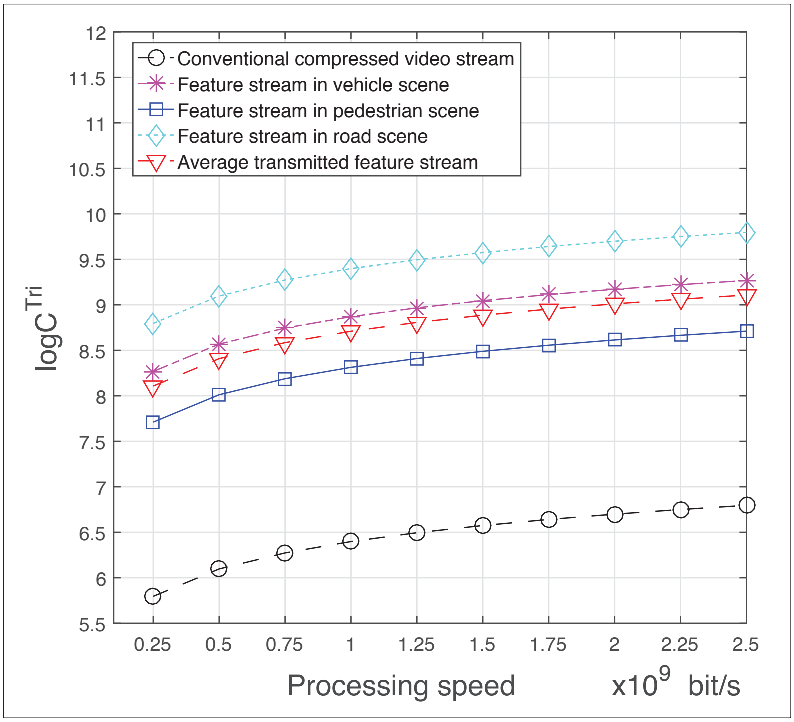

FIGURE 6. Ability of data processing comparisons in a Fog circumstance.