An Air-Ground Integration Approach for Mobile Edge Computing in IoT

Zhenyu Zhou, Junhao Feng, Lu Tan, Yejun He, and Jie Gong

The growing demands for massive connectivity, ultra-low latency, and high reliability in various emerging resource-hungry and computation-intensive applications pose new challenges in network access capacity. This has motivated the authors to conceive a new MEC framework from an air-ground integration perspective.

ABSTRACT

Mobile edge computing (MEC), which enables delay-sensitive computing tasks to be executed at network edges, has been proposed to accommodate the explosive growth of the Internet of Things. However, the growing demands for massive connectivity, ultra-low latency, and high reliability in various emerging resource-hungry and computation-intensive applications pose new challenges in network access capacity. This motivates us to conceive a new MEC framework from an air-ground integration perspective. First, we present a review of the state-of-the-art literature. Then the architecture and technological benefits of the proposed framework are elaborated. Next, four typical use cases are introduced, and a case study is conducted to demonstrate the significant performance improvements in computation capability and communication connectivity based on real-world road topology. Finally, we present challenges and research directions, and conclude this article.

INTRODUCTION

The Internet of Things (IoT), which connects a great variety of physical things such as smartphones, sensors, home appliances, vehicles, and wearable devices through a network, enables a new paradigm shift for ubiquitous information exchange and communication [1, 2]. However, the long transmission distance between IoT devices and remote data centers and the capacity-constrained backhaul links in the conventional cloud computing paradigm impose new challenges on reliable quality of service (QoS) and quality of experience (QoE) provisioning. To this end, mobile edge computing (MEC), in which computing and storage resources are placed at network edges, has emerged as a promising solution. In MEC, delay-sensitive computing tasks can be executed in close proximity to IoT devices to reduce response time and alleviate traffic congestion at core networks [3].

However, the explosive demands for massive connectivity, ultra-low latency, and high reliability in delay-sensitive and multimedia-rich IoT applications pose new challenges in network access capacity (i.e., the number of connections that can be simultaneously accommodated by MEC nodes). Since physical infrastructures of MEC are generally deployed in a fixed fashion, this lack of mobility results in inability to meet the computation and connectivity demands with the characteristics of fast temporal, spatial, and spectral variations. Considering a hot zone where a large number of IoT devices try to connect to MEC nodes simultaneously, this immense number of connection requests and computing demands will severely overwhelm the network access and computing capacities of the MEC nodes, which eventually incurs acute performance degradation. In the opposite circumstances, a large quantity of resources in the MEC node remain underutilized when the loads within coverage shift from peak to valley. Therefore, the mismatch between irregular user demands and rigid network capacity will seriously impede the QoS and QoE guarantees and decrease resource utilization efficiency.

To address the above challenges, we propose an air-ground integrated MEC framework, which actively explores the systematic and complementary integration of the communication, computing, and storage resources from both air and ground segments to provide on-demand deployment densification of edge servers. The redundant resources in unmanned aerial vehicles (UAVs) and ground vehicles can be envisaged as supplementary edge computing servers for efficient resource utilization and reliable service provisioning [4]. On one hand, it is illustrated that UAVs not only provide vast coverage over wide geographical areas, but also possess unique characteristics like fast deployment, easy programmability, and high scalability. For instance, the outage probability of cell edge users can be effectively reduced by the UAV-enabled light of sight (LoS) connectivity and wide ground coverage. On the other hand, vehicular edge computing, which distinguishes itself from static MEC with its high geographical distribution density and proximity to users, can be deployed with UAVs in a complementary way to provide service differentiation or to enhance computing capability and communication connectivity in urban areas. Specifically, the large volume of vehicles in hotspots can provide additional computing capability and multihop data delivery capacity. Furthermore, users who suffer from intermittent connection and service interruption when UAVs are running out of battery power can be served by nearby vehicles in order to enable seamless connection. Therefore, instead of relying on a single technology, this new paradigm

Digital Object Identifier: 10.1109/MCOM.2018.1701111

Zhenyu Zhou, Junhao Feng, and Lu Tan are with North China Electric Power University; Yejun He is with Shenzhen University; Jie Gong is with Sun Yat-sen University. Jie Gong is the corresponding author.

Category	Literature	Feature	Merit	Limit
MEC	• [5]	• Provide MEC services for FiWi networks • Consider both MEC and conventional cloud scenarios	• Achieve the performance gains in FiWi broadband access network	• UAV or vehicle enabled MEC is not considered • The air-ground integrated resources are neglected
	• [6]	• Provide MEC services for distributed applications • Develop an effective middleware	• Enhance current MEC architecture by deploying and developing distributed applications more flexibly	• UAV or vehicle enabled MEC is not considered • The air-ground integrated resources are neglected
	• [7]	• Provide MEC services for renewable energy • Study energy harvesting MEC systems	• Increase the energy efficiency	• UAV or vehicle enabled MEC is not considered • The air-ground integrated resources are neglected
	• [8]	• Provide MEC services for multi-user systems • Develop an online joint computing and radio resource management algorithm	• Minimize the power consumption and long-term average weight of devices	• UAV or vehicle enabled MEC is not considered • The air-ground integrated resources are neglected
UAV or vehicle enabled MEC	• [9]	• Communication and computing-based on vehicular resources	• Enhance the quality of applications and services	• The air-ground integrated resources are neglected
	• [10]	• Service distribution among vehicular fog nodes	• Promote the performance of computing and storage for fog computing	• The air-ground integrated resources are neglected
	• [11]	• UAV-based MEC platform for IoT	• Increase the task offloading efficiency	• The air-ground integrated resources are neglected
Air-ground integration	• [12]	• Multi-UAV-aided vehicular networks	• Improve the performance of vehicular networks	• The unique characteristics of MEC are ignored
	• [13]	• Exploit the advantages of space, air, and ground segments to promote vehicular services	• Support diverse vehicular services efficiently and cost-effectively	• The unique characteristics of MEC are ignored

Table 1. A comprehensive summary of related works.

combines the advantages of different segments to improve connection quality and capacity between MEC nodes and IoT devices.

In this article, first, we present a literature review. Then we narrate the proposed architecture from three aspects, the cloud computing layer, edge computing layer, and device layer, with a particular emphasis on its technological advancements. Next, four typical use cases are elaborated to illustrate potential values of the air-ground integrated MEC framework. A case study is presented to demonstrate the significant performance improvements in computing capability and communication connectivity based on real-world road topology. Finally, we discuss several open research challenges and draw conclusions for this article.

RELATED WORKS

In this part, we present an exhaustive commentary on the development of MEC, vehicle or UAV assisted MEC, and air-ground integration. Table 1 provides a brief classification of the related literature.

MEC has attracted a growing amount of attention in the research community. In [5], Bhaskar *et al.* employed conventional cloud computing and MEC to achieve performance gains in the fiber-wireless (FiWi) broadband access network. An enhanced architecture of MEC was proposed in [6] to facilitate distributed application development and deployment. In [7], Xu *et al.* provided an efficient resource management algorithm and addressed the challenges of integrating renewable energy into MEC. An online joint computing and radio resource management algorithm was developed for multi-user MEC systems in [8], where the power consumption and long-term average weight of MEC servers and devices were minimized.

There are other works that have already provided some effective strategies to shift the functionalities of MEC to vehicles or UAVs. In [9], Hou *et al.* proposed a vehicular fog computing (VFC) architecture to enhance the quality of applications and services. In [10], Sookhak *et al.* presented a fog vehicular computing (FVC) concept that expanded fog computing capacity by utilizing idle resources of vehicles and alleviated the acute performance degradation during peak hours. In [11], Motlagh *et al.* compared the performance of local processing and UAV-assisted MEC processing, demonstrating that the performance of system responsiveness and energy harvesting can be greatly improved by the latter. However, these works ignore the specific research problems in air-ground integrated MEC such as the interoperation and integration of air-ground resources, the respective network reconfiguration, and the mobility management.

A few works considered the research problems in vehicular networks based on space-air-ground integration. In [12], Zhou *et al.* developed an architecture of air-ground integrated cooperative vehicular networks to facilitate vehicular applications. In [13], Zhang *et al.* proposed a space-air-ground integrated vehicular network architecture, which exploits resources from both space and air segments to enhance the connectivity capacity for vehicular networks. However, these works mainly target the research problems in vehicular networks, while the use cases and research challenges in the air-ground integrated MEC have been neglected.

The air-ground integrated MEC framework is conceived by exploring the advantages of software-defined networking (SDN), which provides centralized network control and flexible resource management through the separation of data and control functions and the advanced abstraction of underlying infrastructures.

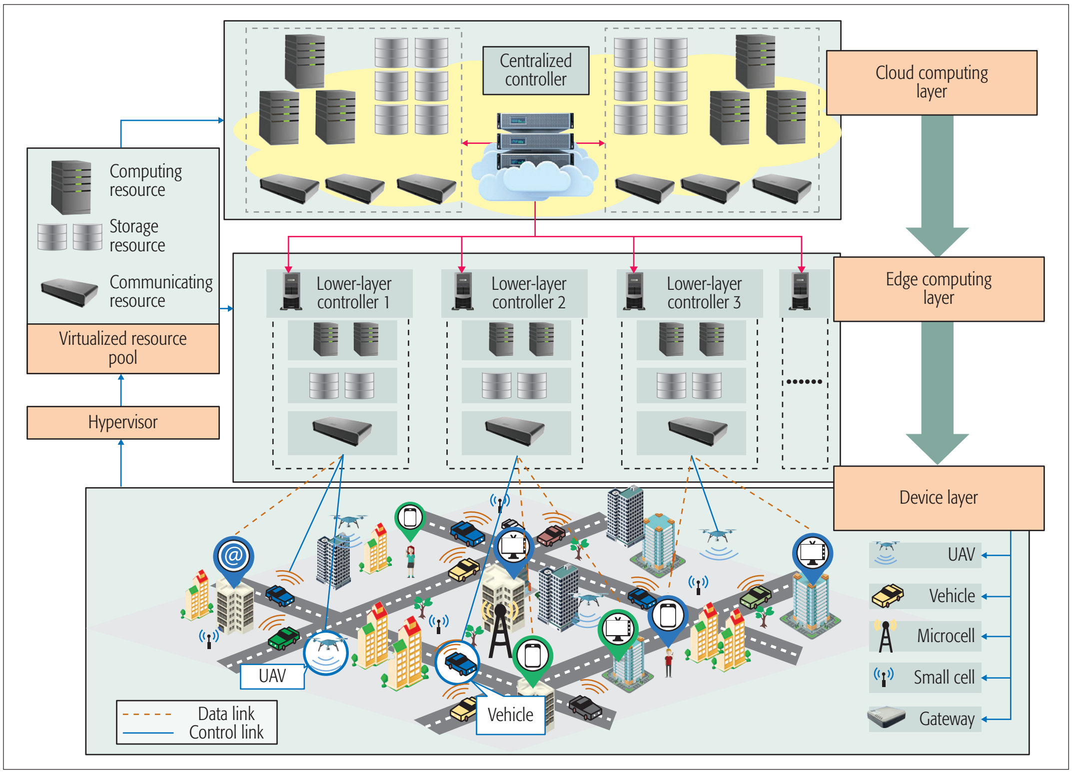

Figure 1. The architecture of the air-ground integrated MEC.

THE AIR-GROUND INTEGRATED MEC FRAMEWORK

In this section, we introduce the conceptual architecture of the proposed air-ground integrated MEC framework and present the corresponding technological benefits.

ARCHITECTURE OVERVIEW

The air-ground integrated MEC framework is conceived by exploring the advantages of software-defined networking (SDN), which provides centralized network control and flexible resource management through the separation of data and control functions and the advanced abstraction of underlying infrastructures. Figure 1 illustrates the conceptual architecture of the proposed air-ground integrated MEC framework, which relies on the deep convergence of three layers: the cloud computing layer, edge computing layer, and device layer.

The device layer is mainly composed of different programmable and reconfigurable equipments and elements spanning both the air and ground network segments, including UAVs, vehicles, base stations, routers, gateways, and various IoT devices. The device layer resources as well as the upper-layer computing and storage resources are extracted and virtualized by a hypervisor, and are eventually consolidated into a virtual resource pool. The hypervisor takes over the driving of underlying infrastructures, and provides an abstraction to the controllers [14]. Through a proper mapping between physical infrastructures and virtualized network resources, multiple virtual MEC nodes can be established and operated on the same underlying infrastructures. Each MEC node only occupies a slice of the overall virtualized resources, and is independently controlled by the upper-layer controllers.

The cloud and edge computing layers together provide basic computing functionalities for IoT applications. A hierarchical computing architecture is adopted to handle computing task allocation and resource management at different scales. The edge computing layer, which acts as a beneficial complement to the cloud computing layer, is more suitable for local processing of small-scale delay-sensitive computing tasks in a distributed fashion. In the edge computing layer, the communication, computing, and storage resources are deployed close to users and distributed across the network edges. Lower-layer edge controllers can only see and manage their own virtual networks, and are controlled by the upper-layer cloud controllers.

Compared to the edge computing layer, the scale of resources in the cloud computing layer is several orders of magnitudes larger; the resources are uniformly controlled and managed by the cloud controller with centralized intelligence and global network knowledge. Vendor-independent and fine-grained control of the virtualized resources becomes feasible via the data-control decoupling and the high-level abstraction of physical infrastructures. To accommodate the different QoS requirements of IoT applications, the upper-layer controllers can reassemble the virtual resources and allocate them to each lower-layer controller, which then dynamically adjusts the resource assignment based on traffic demands.

BENEFITS OF AIR-GROUND INTEGRATED ARCHITECTURE

Connectivity Enhancement: The redundant resources in both air and ground segments are integrated in a complementary way to enhance the connectivity capability between MEC nodes

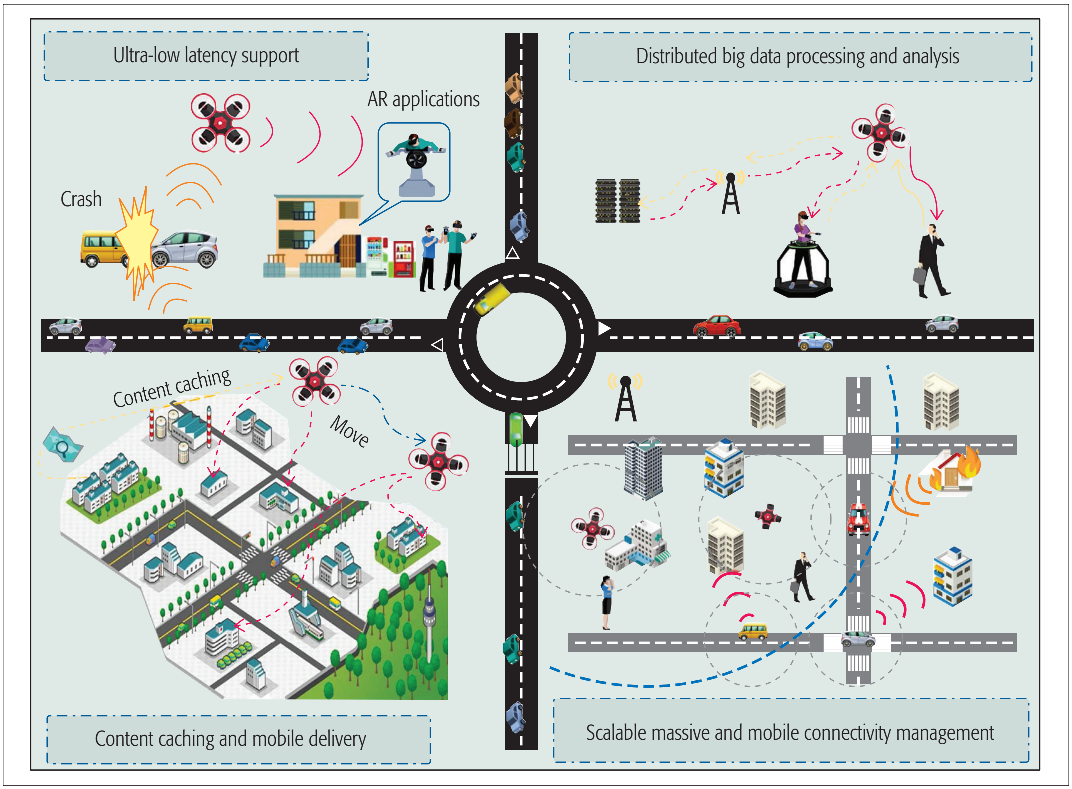

Figure 2. Four typical use cases of the air-ground integrated MEC.

The proposed air-ground integrated MEC framework allows effective mobility management via software-defined air interface separation. UAVs with high altitudes, large-scale coverage, and LoS connections, can effectively eliminate the unnecessary handover and provide uninterrupted computing service for high-mobility users.

and IoT devices. Considering the scenario of a hot pedestrian zone, the number of connections that can be simultaneously afforded is severely limited by the inferior non-LoS (NLoS) channel conditions due to surrounding blockages. In comparison, the utilization of UAVs can provide on-demand augmentation of connectivity capacity due to the wide coverage areas, fast deployment, and availability of LoS connections. Furthermore, vehicles that are close to pedestrians can act as relays to provide multihop data delivery, which expands the effective coverage of MEC services.

Adaptive Resource Allocation: To match resource provision with the time and space varying computing demands, the scale of resource virtualization and the granularity of resource allocation can be adjusted dynamically at different levels. For instance, a number of interconnected UAVs and vehicles, which contain both ground-to-ground and ground-to-air links, and involve a number of resources including bandwidth, power, buffer, queue, and so on, can be mapped to a single virtual link from a higher degree of virtualization. The virtualized link resources are uniformly managed by the centralized controller for load balancing and connectivity improvement. The hierarchical control architecture also enables flexible adjustment of resource allocation granularity. Particularly, the micro-scale management of load distribution, power control, spectrum allocation, and user association is taken care by the lower-layer edge controllers based on instantaneous computing demands, while the macro-scale management of edge server placement, UAV/vehicle dispatch, and inter-segment resource coordination are controlled by the upper-layer cloud controllers based on long-term observations and predictions.

Differentiated QoS Guarantees: In MEC, a number of IoT applications with diverse requirements on throughput, latency, reliability, and connectivity coexist in the same network. To cope with these diverse QoS requirements and provide differentiated services, fine-grained resource allocation is performed under the guidance of application requests, which are first translated into explicit instructions and delivered to the respective controllers via standardized application programming interfaces (APIs). With effective data-control separation and an advanced level of virtualization, the high-level QoS requirements can be enforced by the hierarchical controllers that dynamically adapt resource provisioning in accordance with service demands. In addition to adaptive resource orchestration, the outage probability of users located far away from MEC nodes can be significantly reduced by exploring the connectivity enhancement capability provided by air-ground integration. UAVs and vehicles that enable on-demand mobile connectivity can provide effective data offloading for heavy-load cells, which significantly enhances network access capacity and QoS provisioning reliability.

Effective Mobility Management: The proposed air-ground integrated MEC framework allows effective mobility management via software-defined air interface separation. UAVs with high altitudes, large-scale coverage, and LoS connections, can effectively eliminate unnecessary handover and provide uninterrupted computing service for high-mobility users. Furthermore, based on the knowledge gained from the historical data (e.g., the distribution of computing demands), the macro-scale UAV/vehicle dispatch and the micro-scale resource allocation can be jointly optimized by the controllers. Since the virtualization of

Figure 3. The real-world road topology of the Xidan area.

underlying air and ground infrastructures is implemented by the hypervisor in a transparent way, the edge controllers can manage the respective virtual resources in a highly abstracted manner instead of directly interacting with hardware/software heterogeneities, which significantly increases the efficiency of inter-segment resource management and coordination. Hence, the controllers are able to reconfigure the network structures and reassemble the virtualized resources with strict timeliness in response to dynamically varying network topologies and computing demands.

USE CASES

The three basic application scenarios of the incoming fifth generation (5G) era are enhanced mobile broadband (eMBB), massive machine-type communications (mMTC), and ultra-reliable low-latency communications (URLLC). In this section, we present four use cases to demonstrate how to support a grand set of 5G applications with high throughput, low latency, and high mobility requirements with the proposed air-ground integrated MEC framework. Particularly, the use case "ultra-low latency support" corresponds to the URLLC scenario, the use cases "distributed big data processing and analysis" and "content caching and mobile delivery" correspond to the eMBB scenario, and the use case "scalable massive and mobile connectivity management" corresponds to the mMTC scenario. The use cases are shown in Fig. 2, and are detailed as follows.

Ultra-Low Latency Support: In the proposed framework, the infrastructures of air-ground integration rely on underlying MEC devices such as vehicular networks or UAVs instead of remote cloud computing servers; thus, latency issues can be improved significantly and user experience can be enhanced effectively. In particular, the LoS connections provided by UAVs or vehicles can dramatically improve connection quality, which is essential to support delay-sensitive applications. Furthermore, with the implementation of air-ground integration, timely reactions and adaptive solutions can be achieved in disaster relief and emergency management. For instance, when a traffic accident happens, the accident vehicle directly reports to the surrounding vehicles and UAVs, which then provide real-time audio/video monitoring for further inspection and management. In practice, a secret Google project, SkyBender, has already demonstrated that data can be delivered 40 times faster compared to existing 4G LTE systems.

Distributed Big Data Processing and Analysis: Different from conventional cloud computing or static edge computing with fixed location and computing capability, the computing capacity in the air-ground integration framework can be augmented dynamically in accordance with traffic demands. For example, a group of UAVs that are interconnected with each other via LoS air-to-air links can jointly process a large amount of traffic data by exchanging some key data with each other. In other words, the onboard data processing capacity can be shared among UAVs through the air-to-air connections that can be approximated as free space communication. With such distributed data processing capability, the complicated network behaviors that involve large volumes of data can be analyzed in real time to flexibly adjust resource allocation and reconfigure network structures according to dynamic traffic profiles.

Scalable Massive and Mobile Connectivity Management: By combining the technological advantages of each segment, static base stations with constant power supply, fixed backhaul connection, and large transmission power can be utilized to provide access service, while UAVs and vehicles provide on-demand content delivery and computing service by exploiting their mobility capability. In this sense, scalable massive connectivity can be supported via opportunistic scheduling and dynamic connectivity management. Taking vehicular networks as an example, the MEC based on mobile multihop connections can be regarded as a supportive mechanism for realizing emerging intelligent transport system (ITS) applications such as self-driving, which not only collects real-time information from roadside sensors and vehicular applications directly, but also enables local message analysis and warning (e.g., accidents) broadcasting to nearby vehicles. In addition, UAVs can also provide on-demand mobile connectivity for public safety surveillance. For instance, when a fire occurs, a group of UAVs can be dispatched immediately to the fire site to provide live video footage and static images, allowing proactive reconnaissance and decision making during the search and rescue operation.

Content Caching and Mobile Delivery: The proposed framework can reduce the content distribution latency and eliminate network overload via mobile content caching and delivery from the following three aspects. First, the data that are frequently requested can be selectively cached in the air-ground integrated edge servers without any further replicated transmission in the backhaul. Second, vehicles that are proximate to pedestrians and UAVs with flexible deployment and reconfigurability are combined to optimize network access and wireless coverage, which in turn allows more users to fetch the contents cached in the mobile edge servers previously, and hence, further minimizes the duplicate content

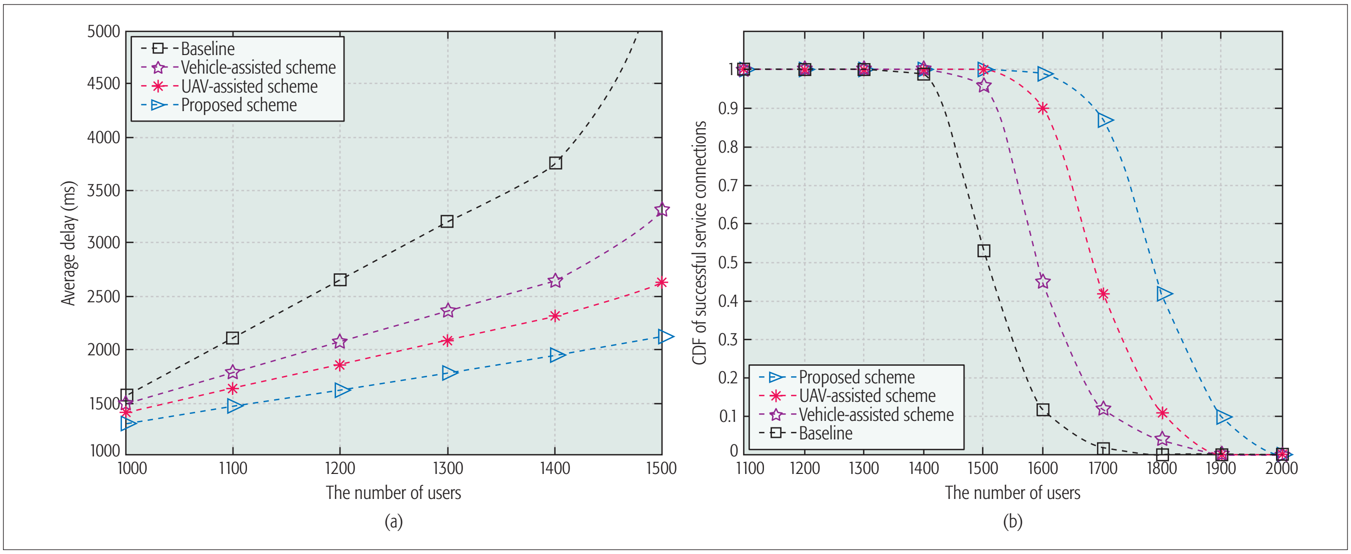

Figure 4. Performance of the air-ground integrated MEC framework: a) average delay vs. the number of users.; b) CDF of successful service connections vs. the number of users.

transmission. Last but not least, with the mobility support of UAVs and ground vehicles, the duplicate content caching in different locations can be greatly reduced, as they can roam across the network to move the popular contents along with user trajectories.

CASE STUDY

In this section, a case study is conducted to evaluate the performance of the proposed air-ground integrated MEC.

EXPERIMENTAL SETUP

The simulation is implemented based on two main platforms: SUMO and MATLAB. SUMO is adopted to generate vehicular traffic in real-world road topologies. As a representative business and financial district in Beijing, China, the Xidan area is selected as the evaluation scenario, which is characterized with large-scale office buildings and commercial districts as well as pedestrian zones. Figure 3 shows a snapshot of the Xidan area obtained from Google Maps. The digital map information is available in OpenStreetMap, and the corresponding data are imported to SUMO by using the NETCONVERT tool. Then, based on the road topologies including nodes, edges, connection links, traffic signs, and so on, vehicular traffic is generated by using the RANDOMTRIPS tool with a step of 10^{-1} s. The total duration of the simulation is set as 3.6×10^3 s. The coordinates and velocity of each vehicle can be saved into an Excel file.

The flying trajectory of UAVs is developed based on [15], with an altitude around 60~150 m. The battery capacity of a small-type UAV is 20,000 mAh. The flight duration is around 30 min [11]. The computing capability of any UAV is set as 0.7~1.0 GHz. The transmission power of any UAV or vehicle is set as 23 dBm. We assume that there are 1000~2000 mobile users who are randomly distributed along the roadside based on the real-road topology. Each user might carry a couple of IoT devices, such as smartphones, smart watches, and smart glasses. The computing tasks generated for each user are assumed to follow a Poisson process, which are offloaded to MEC nodes (UAVs or vehicles). The total delay of the offloaded portion arises from task uploading, queuing, processing, and result feedback. The traffic model of any MEC node is considered as an $M/M/c$ queue by supposing that the service time obeys an exponential distribution, where c denotes the number of co-located edge servers.

Afterward, the simulation of task offloading is conducted in MATLAB. Some simulation parameters such as user location, number of computing tasks, and channel models are generated based on theoretical models rather than real-world measurements. The reason is that these theoretical models (e.g., the free space propagation path loss model) have been intensively adopted in many previous works and provide good tractability. The real-world implementation and experiment will be left as future work.

NUMERICAL RESULTS

The proposed framework is compared to three heuristic schemes: the baseline scheme, the vehicle-assisted scheme, and the UAV-assisted scheme. For the purpose of fair comparison, the numbers of static edge servers in all of the four schemes are set the same to remove the impacts of static edge servers and fairly compare the performance gain achieved by integrating vehicles or UAVs. Meanwhile, the total numbers of vehicles and UAVs are kept as the same for the vehicle-assisted scheme, the UAV-assisted scheme, and the proposed scheme.

Figure 4a shows the relationship between the number of users and the average delay performance. It is observed that the average delay performances of all four schemes increase monotonically with the number of users. The reason is that the time spent on task uploading, queuing, and processing will increase accordingly as more users try to connect to the same edge nodes and offload computing tasks. Nevertheless, numerical results demonstrate that the average delay can be reduced by a significant margin via efficient utilization of the air-ground integrated resourc-

The limited onboard battery capacity is another key constraint of UAVs as edge servers. Besides the energy consumed in hovering, and accelerating/de-accelerating for climbing up/down and speed up/slow down, additional communication and computing energy consumption, as well as the additional weight gained by mounting the communication and computing modules, have to be taken into consideration.

es. The average delay achieved by the proposed scheme is 48 percent lower than that of the baseline scheme when the number of users reaches 1400. The reason is that UAVs and vehicles with LoS communication links provide effective countermeasures to relieve the negative impacts of multi-path fading. It is also observed that the proposed scheme outperforms the UAV-assisted scheme. The reason is that vehicles with flexible deployment and reconfigurability are conceived in a mutually beneficial way to enhance network access capacity. For example, the large volume of vehicles in hotspots that are close to pedestrians can significantly reduce transmission delay due to the shorter transmission distance.

Figure 4b shows the cumulative distribution function (CDF) of successful service connections vs. the number of users. Here, the successful service connections represent the ratio of users whose computing requests are successfully accepted by MEC nodes. It is noted that when the number of users increases to 1600, only 12 percent of service requests can be successfully connected in the baseline scheme due to the access and backhaul related limitations. In comparison, the proposed scheme outperforms the baseline scheme by 88 percent due to the fact that the integrated air-ground resources with LoS connectivity and large-scale coverage can effectively enhance QoS performances of cell edge users and provide more reliable communication links than the traditional static connections. Furthermore, from both Fig. 1 and Fig. 2, we observe that the performance of the UAV-assisted scheme always outperforms the vehicle-assisted scheme. The reason is that connection interruption occurs when the associated vehicle moves away and there is no alternate vehicle in proximity to continue service provisioning, which is less likely to happen for UAVs with larger coverage.

Challenges and Research Directions

Despite the tremendous benefits brought by the concept and architecture for providing lower latency and high reliability services to IoT devices, there are still several challenges.

Coordination among UAVs and Ground Vehicles

To establish a network by UAVs and ground vehicles, inter-UAV-vehicle communication and coordination are required. Due to the mobility of UAVs and ground vehicles, the coordination among them is not as simple as the fixed infrastructures with fiber-based backhaul. These mobile edge servers have to be interconnected via wireless links. One way is that each individual UAV coordinates with its neighbors distributively, which requires little controlling overhead but leads to higher inaccuracy. Such a strategy might also result in an error spreading problem. The other way is to nominate a cluster head to collect the information of all the UAVs to make a centralized decision. In this case, the cluster head may be overwhelmed by the immense overheads so that additional computing cannot be afforded. In this case, the round-robin protocol in which each UAV with fewer computing tasks acts as a cluster head for a short duration may be considered as a candidate solution.

Joint Trajectory Design and MEC Optimization

One of the key design challenges is to adaptively adjust UAV trajectory to meet the computing task requirements. From a single-user perspective, research efforts should focus on how to predict the user movement and track the trajectory so that the computing tasks can be offloaded immediately, and the computation results can be fed back to users on time. For multi-user networks, the design of UAV trajectory becomes even more challenging, where the distribution and mobility of users, task priority and delay requirements, user locations and distance to UAVs, as well as energy consumption issues need to be jointly considered. Intuitively, users can be scheduled in a time-division multiplexing way, that is, only one user is served per channel by the UAV when they are close enough to each other. Then the serving order of users becomes critical, where not only the transmission delay, but also the queuing and computation delays should be taken into account. In fact, computing and communication can be scheduled in different time slots to fully utilize the resources of UAVs.

Energy-Efficient Management

The limited onboard battery capacity is another key constraint of UAVs as edge servers. Besides the energy consumed in hovering, and accelerating/de-accelerating for climbing up/down and speed up/slow down, additional communication and computing energy consumption, as well as the additional weight gained by mounting the communication and computing modules, have to be taken into consideration. There have been some efforts to model the propulsion energy consumption of UAVs, while the onboard energy consumption of communication and computing requires further investigation. On the other hand, UAV station replacement, recharging scheduling, as well as energy harvesting can be considered as alternate directions to compensate the battery capacity limitation. Compared to existing works that solely rely on UAV-assisted communications, the proposed air-ground integration approach provides a promising way to address the energy consumption problem. For example, users in hotspots with high vehicle density can be served by nearby vehicles to reduce the energy consumption of UAVs. Furthermore, since the transmission distance from vehicles to pedestrians is much shorter compared to that of UAVs, the system energy efficiency can also be increased significantly.

Conclusions

In this article, we have proposed an air-ground integrated MEC framework by exploiting the benefits of high mobility and flexible computing resource allocation of UAVs and vehicles. By smartly managing the aviation trajectory of UAVs and seeking the LoS air-ground channels, the proposed framework can support numerous IoT use cases including content caching and mobile delivery, distributed big data processing and analysis, ultra-low latency support, and so on. Numerical results based on the case study for a real-world road environment show that the overall delay can be greatly reduced with the proposed air-ground integrated MEC framework.

ACKNOWLEDGMENTS

This work was partially supported by the Beijing Natural Science Foundation under Grant Number 4174104, Beijing Outstanding Young Talent under Grant Number 2016000020124G081, the National Natural Science Foundation of China (NSFC) under Grant Numbers 61601181 and 61771495, and the Fundamental Research Funds for the Central Universities under Grant Number 2017MS001.

REFERENCES

[1] Y. Zhang *et al.*, “Cognitive Machine-to-Machine Communications: Visions and Potentials for the Smart Grid ,” *IEEE Network*, vol. 26, no. 3, May 2012, pp. 6–13.

[2] Y. Wu *et al.*, “Energy-Aware Cooperative Traffic Offloading via Device-to-Device Cooperations: An Analytical Approach,” *IEEE Trans. Mobile Comp.*, vol. 16, no. 1, Jan. 2017, pp. 97–114 .

[3] X. Huang *et al.*, “Exploring Mobile Edge Computing for 5G-Enabled Software Defined Vehicular Networks,” *IEEE Wireless Commun.*, vol. 24, no. 6, Dec. 2017, pp. 55–63.

[4] C. Wu *et al.*, “A Reinforcement Learning-Based Data Storage Scheme for Vehicular Ad Hoc Networks,” *IEEE Trans. Vehic. Tech.*,vol. 66, no. 7, July 2017, pp. 6336–48.

[5] B. P. Rimal *et al.*, “Mobile-Edge Computing Versus Centralized Cloud Computing Over a Converged FiWi Access Network,” *IEEE Trans. Net. Service Mgmt.*, vol. 14, no. 3, Sept. 2017, pp. 498–513.

[6] A. Carrega *et al.*, “A Middleware for Mobile Edge Computing,” *IEEE Cloud Comp.*, vol. 4, no. 4, July/Aug. 2017, pp. 26–37.

[7] J. Xu *et al.*, “Online Learning for Offloading and Autoscaling in Energy Harvesting Mobile Edge Computing,” *IEEE Trans. Cogn. Commun. Net.*, vol. 3, no. 3, Sept. 2017, pp. 361–73.

[8] Y. Mao *et al.*, “Stochastic Joint Radio and Computational Resource Management for Multi-User Mobile-Edge Computing Systems,” *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, Sept. 2017, pp. 5994–6009.

[9] X. Hou *et al.*, “Vehicular Fog Computing: A Viewpoint of Vehicles as the Infrastructures,” *IEEE Trans. Vehic. Tech.*, vol. 65, no. 6, June 2016, pp. 3860–73.

[10] M. Sookhak *et al.*, “Fog Vehicular Computing: Augmentation of Fog Computing Using Vehicular Cloud Computing,” *IEEE Vehic. Tech. Mag.*, vol. 12, no. 3, July 2017, pp. 55–64.

[11] N. H. Motlagh *et al.*, “UAV-Based IoT Platform: A Crowd Surveillance Use Case,” *IEEE Commun. Mag.*, vol. 55, no. 2, Feb. 2017, pp. 128–34.

[12] Y. Zhou *et al.*, “Multi-UAV-Aided Networks: Aerial-Ground Cooperative Vehicular Networking Architecture,” *IEEE Vehic. Tech. Mag.*, vol. 10, no. 4, Dec. 2015, pp. 36–44.

[13] N. Zhang *et al.*, “Software Defined Space-Air-Ground Integrated Vehicular Networks: Challenges and Solutions,” *IEEE Commun. Mag.*, vol. 55, no. 7, July 2017, pp. 101–09.

[14] Z. Zhou *et al.*, “Software Defined Machine-to-Machine Communication for Smart Energy Management,” *IEEE Commun. Mag.*, vol. 55, no. 10, Oct. 2017, pp. 52–60.

[15] V. Sharma *et al.*, “UAV-Assisted Heterogeneous Networks for Capacity Enhancement,” *IEEE Commun. Letters*, vol. 20, no. 6, Apr. 2016, pp. 1207–10.

BIOGRAPHIES

ZHENYU ZHOU [M’11, SM’17] (zhenyu_zhou@ncepu.edu.cn) received his Ph.D degree from Waseda University, Japan, in 2011. Since March 2013, he has been an associate professor at North China Electric Power University. He is an Editor for *IEEE Access* and *IEEE Communications Magazine*, and Workshop Co-Chair for IEEE ISADS ’15, GLOBECOM ’18, and others. He received the 2017 IET Premium Award, and the IEEE ComSoc GCCTC 2017 Best Paper Award. His research interests include green communications and vehicular communications.

JUNHAO FENG (fengjh@163.com) is currently working toward an M.S. degree at North China Electric Power University. His research interests include resource allocation, interference management, and energy management in D2D communications.

LU TAN (13717568897@163.com) is currently working toward an M.S. degree at North China Electric Power University. Her research interests include green communications and smart grid.

YEJUN HE (heyejun@126.com) received his Ph.D.degree in information and communication engineering from Huazhong University of Science and Technology in 2005. He is a full professor with the College of Information Engineering, Shenzhen University, China, where he is the director of the Guangdong Engineering Research Center of Base Station Antennas and Propagation, and the director of the Shenzhen Key Laboratory of Antennas and Propagation. His research interests include wireless mobile communication, antennas, and radio frequency. He is a Fellow of IET.

JIE GONG [S’09, M’13] (gongj26@mail.sysu.edu.cn) received his B.S. and Ph.D. degrees from the Department of Electronic Engineering at Tsinghua University in 2008 and 2013, respectively. He visited the University of Edinburgh in 2012. During 2013– 2015, he worked as a postdoctorial scholar at Tsinghua University. He is currently an associate professor in the School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China. His research interests include cloud RAN, energy harvesting technology, and green wireless communications.

By smartly managing the aviation trajectory of UAVs and seeking the LoS air-ground channels, the proposed framework can support numerous IoT use cases including content caching and mobile delivery, distributed big data processing and analysis, ultra-low latency support, and so on.