Green and Sustainable Cloud of Things: Enabling Collaborative Edge Computing

Zhaolong Ning, Xiangjie Kong, Feng Xia, Weigang Hou, and Xiaojie Wang

The authors discuss some distinct research directions in CoTs with respect to the edge computing. Given the significance of energy efficiency and sustainability of edge deployment in CoTs, they put forward a green and sustainable virtual network embedding framework for cooperative edge computing in wireless-optical broadband access networks.

Abstract

The proliferation of IoTs beside the emergence of various cloud services push the horizon of edge computing. By offering cloud capabilities at the network edge closer to mobile devices, edge computing is a promising paradigm to resolve several vital challenges in IoTs, such as bandwidth saturation, energy constraints, low latency transmission, and data security and privacy. To provide a comprehensive understanding of edge computing supported by the integration of IoTs and cloud computing, that is, CoTs, this article first discusses some distinct research directions in CoTs with respect to edge computing. Given the significance of energy efficiency and sustainability of edge deployment in CoTs, we put forward a green and sustainable virtual network embedding framework for cooperative edge computing in wireless-optical broadband access networks. Specifically, we leverage a reliability function to confirm the number of backup edge devices, and embed virtual networks onto the suitable edge devices in CoTs. Finally, several research challenges and open issues are discussed.

Introduction

How to make the development of smart cities sustained is challenging due to the huge evolved investments. Owing to the superiorities of high computing capacity, competitive system scalability and low service cost, cloud computing has become a high-efficiency alternative in various fields of smart cities, such as social networks, commerce, transportation, education and health care services [1]. However, centralized cloud frameworks commonly face several challenges, such as network overload, end-to-end latency, single point of failure and high power consumption.

According to a report released by Cisco, mobile data traffic will increase by 53 percent annually from now until 2020 [2]. The number of network users provided by cloud services will be approximately 3.6 billion in 2018 (http://www.statista.com). With the advance triumphantly of Internet of Things (IoTs), the post-cloud era, that is, Cloud of Things (CoTs) is emerging. Thanks to the ever-increasing abilities of sensing, computing and communication of smart devices, nearly 50 billion things are likely to be connected through the Internet, where 45 percent of the generated data by IoTs will be handled by edge devices in 2019 [3]. However, it is overwhelmingly challenging for smart cities to manage or program such a huge system with massive mobile devices due to the high heterogeneity of IoT-based applications [4].

Currently, Wireless-Optical Broadband Access Networks (WOBAN) have become a key component in smart cities. As shown in Fig. 1, Wireless Mesh Networks (WMNs) and IoTs can be the front-end of WOBANs. The generated data are forwarded by routers via wireless channels and received by the neighbor optical network unit (ONU). Then, an optical line terminal (OLT) can receive the sensed data through a Passive Optical Network (PON). According to the obtained information flows, a decision can be made in the central office connecting to the OLT for industrial or commercial circumstances. With an ever-growing volume of data generated by smart devices and the saturating PON bandwidths, the central office becomes incapable of sustaining the continuous information flows. For the sake of alleviating the resource overprovisioning of a PON, deployment of edge devices (EDs) is advocated at both ONUs and wireless routers for collaborative edge computing, so that EDs can switch roles from data consumer to both data producer and data consumer.

Network virtualization technologies, such as Network Function Virtualization (NFV) and Software Defined Networking (SDN), are helpful to achieve highly-efficient edge deployment in CoTs by virtualizing device functions on top of standard hardware [5]. For example, a user's computational capability can be abstracted to the Virtual Network Embedding (VNE) on the WMNs or IoTs infrastructure in WOBANs. Because of the inter-channel interference caused by EDs, VNE-based collaborative edge computing generally cannot guarantee network survivability. However, network survivability is important for sustainable edge computing in CoTs.

Although rapidly growing interest has been drawn to edge deployment, the related research of sustainable edge computing in CoTs is still in its infancy. To our best knowledge, our presented Green and Sustainable VNE (GSVNE) framework is the first effort to study the available number of backup EDs, and embed virtual networks (VNs) onto the suitable EDs for collaborative edge computing in CoTs. The main contributions of this article can be summarized as:

- We overview the current research studies of cooperative edge computing in CoTs, and classify the related studies into four categories.

Digital Object Identifier: 10.1109/MCOM.2018.1700895

Zhaolong Ning, Xiangjie Kong, Feng Xia (Corresponding author) and Xiaojie Wang are with Dalian University of Technology, and the Key Laboratory for Ubiquitous Network and Service Software of Liaoning Province; Weigang Hou is with Northeastern University, Shenyang, China.

This article has been accepted for inclusion in a future issue of this magazine. Content is final as presented, with the exception of pagination.

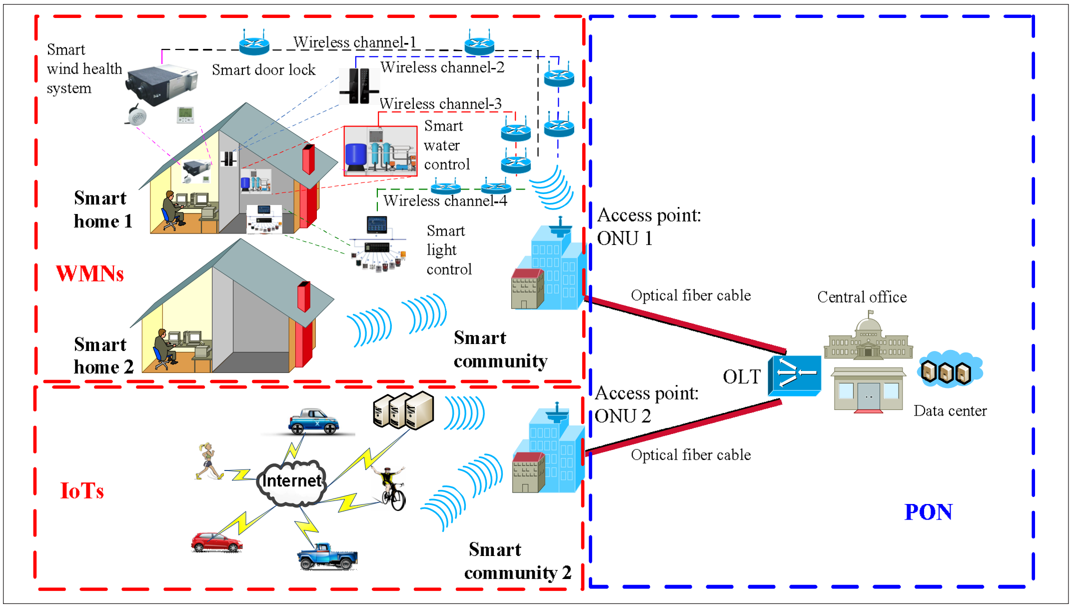

Figure 1. A WOBAN sketch for smart cities.

Due to the heterogeneity and dense deployment of devices, edge caching is beneficial for boosting spectral efficiency and reducing energy consumption in CoTs. Content placement and delivery are two core issues in edge caching. The former considers the location, size, and downloaded content of network caches, while the latter focuses on delivering content to the requested terminals.

- We put forward a GSVNE framework for cooperative edge computing in CoTs, and determine the number of EDs according to network reliability.
- We select the qualified network by our GSVNE. Our cooperative scheme can both serve as many virtual networks as possible and provide a satisfied sharing degree of network backup resources.

The rest of this article is organized as follows. Some promising research directions for edge computing in CoTs are introduced in the next section. After that, the GSVNE framework is put forwarded. We then discuss research challenges and emerging solutions for collaborative edge computing in CoTs before concluding this article.

Research Directions for Edge Computing in CoTs

Edge computing includes various technologies, such as cloudlet, local fog computing, sensing networks, cloud-centric IoTs and distributed peer-to-peer networks. In this section, we discuss the following four promising research directions for edge computing in CoTs.

Edge Computing over WOBANs

The integration of fiber and wireless access networks can offer the advantages of reliability in optical fiber networks and ubiquity in wireless networks. WOBANs can offer broadband services for both fixed subscribers and mobile clients. Therefore, it has become a powerful framework to provide widespread connectivity for services and applications in smart cities. Future smart cities not only require compatible radio interfaces, but also need the integration of heterogeneous networks to manage rapid data growth. Therefore, distributed edge computing over WOBANs is advocated. Furthermore, edge computing is also leveraged to satisfy the tactile-level latency requirements in future networks. This is because the total latency for mobile cloud computing is generally between 30 ms and 100 ms, which is intolerable for latency-critical applications, such as autonomous driving and online gaming [6]. Edge computing over WOBANs can be fulfilled by offering cloud computing capability to the ONUs or access points located at the front-end of WOBANs directly. It is worth noting that places with close proximity to subscribers are the preferred choices for edge deployment to reduce latency. Since the integration of edge computing and WOBANs can provide highly-efficient resource management from a technical perspective, sharing network infrastructures in WOBANs is economical and can decrease operational expenditures by edge computing [7].

Cache-Enabled Edge Computing

Due to the heterogeneity and dense deployment of devices, edge caching is beneficial for boosting spectral efficiency and reducing energy consumption in CoTs. Content placement and delivery are two core issues in edge caching. The former considers the location, size, and downloaded content of network caches, while the latter focuses on delivering content to the requested terminals. Service caching for resource allocation and data caching are two important challenges in this context. Although contents can be cached at WMNs or IoTs in WOBANs, the backhaul links are still constrained. The traffic in wired Internet, backhaul and mobile core networks can be reduced by deploying caches in base stations, relays and user devices [8]. Some trade-offs for edge caching deserve to be investigated, for example, the trade-offs between backhaul capacity and cache size, between throughput and outage in cache-enabled device-to-device (D2D) networks, and between energy consumption and offloading ratio.

Collaborative Edge Computing

Due to a local bottleneck, it is impossible for a single ED to cope with all computing tasks. Therefore, collaborative edge computing based on ED cooperation in the PON backhaul is promising to handle mission-critical applications. Also, collaborative edge computing is compatible with

Since computation merely happens in the data facility of participants, data integrity and privacy can be well guaranteed. Some other promising topics include: role definition, resource mapping, computation offloading, data storage, and cooperation stimulation between cloud and edge entities.

Definition	Description
n	Total number of WMNs and IoTs
N	Total number of VNs
N_1	The number of embedded VNs processed by front-end WMNs and IoTs
N_2	The number of VNs processed by the PON backhaul
Γ_w (Γ_b)	Working (backup) graph of the qualified WMN and IoT
$\eta^i_{\Gamma_\omega}$ ($\eta^i_{\Gamma_b}$)	Binary variable to identify whether the ith VN is embedded onto Γ_w (Γ_b)

Table 1. Main notations.

the Infrastructure as a Service (IaaS) model, making the device's computing ability customized. A context-aware collaborative edge computing system has been investigated for 5G networks, including three utilization circumstances: mobile edge orchestration, collaborative video caching, and interference cancellation between layers [9]. However, the opportunity for collaboration is limited by privacy concerns, thus effective incentives and privacy-preserving schemes are necessary for data sharing. One promising utilization of collaborative edge computing is connecting health information with location information to fuse geographically distributed and multi-domain data for healthcare [10]. Since computation merely happens in the data facility of participants, data integrity and privacy can be well guaranteed. Some other promising topics include: role definition, resource mapping, computation offloading, data storage, and cooperation stimulation between cloud and edge entities.

Virtualized Edge Computing

With the objective of enabling collaborative edge computing in CoTs for smart cities, flexible management of computing resources needs to be coordinated. Network virtualization is advocated, which is economical to customize computing tasks for mobile device users. The current network structures largely benefit from the integration of NFV, SDN and edge computing. One ED can be enabled by NFV and offer computing services to multiple devices through generating virtual machines, so that various tasks and network functions can be simultaneously performed. On one hand, computation burdens on the SDN controller and transmission backbone can possibly be alleviated by edge computing. On the other hand, the configuration and management costs generated by edge computing can be reduced by NFV and SDN [11]. An SDN-enabled edge computing architecture normally contains three layers: edge layer, SDN infrastructure, and northbound applications. Some beneficial fields in edge computing supported by SDN can be summarized as: high-efficiency resolution and control, flexibility and unrestrained innovation, service-oriented implementation, virtual machine deployment, interoperability, adaptability, economical solution and multiplicity of scope [12]. Furthermore, the integration of SDN and edge computing can be regarded as a distributed SDN, in which the corresponding network requirements of edge computing (such as energy consumption, latency and storage) should also be reached.

The Designed Framework

Deployed EDs are likely to lose efficiency due to inter-channel interference. Also, the process of information flow would cause large delays because of the failure of EDs. With the purpose of guaranteeing network sustainability, we put forward a green and sustainable VNE framework for collaborative edge computing in CoTs. Table 1 summarizes the main notations in this article.

An Overview of the GSVNE Framework

Our two-stage framework is illustrated in Fig. 2. In the first stage, we divide WMNs and IoTs into working and backup sub-parts. For the sake of reducing the consumed resources on the backup sub-graphs, we select k EDs for backup-resource provisioning. After obtaining the value of k by the network reliability function, we derive the number of backup EDs for network resource division. Furthermore, the backup graph is constructed before confirming the geographical locations of the backup EDs. By dividing the resources of WMNs and IoTs, the second stage embeds VNs onto the most appropriate working graph by power control before selecting a suitable backup ED. Next, we conduct the green and sustainable embedding operation for VNs on the most qualified WMN or IoT to guarantee the maximum sharing degree (SD) of backup network resources, representing the maximal amount of virtual nodes embedded onto the same backup ED. Because a constant path exists between the bottleneck and another WMN or IoT with the backup ED, link mapping is not necessary. Therefore, the target of the second phase becomes minimizing the quantity of the consumed backup carriers along fixed paths.

Problem Formulation

The WOBAN in our framework has n WMNs and IoTs. The ith ($i \in [1, n]$) WMN or IoT has M_i EDs and P_i wireless routers, where M and P are the sets of EDs connected to the same ONU and deployed at wireless routers, respectively. For one WMN or IoT, the current VN can be expressed by a four-tuple model $vn(s, \phi, wb, c)$. The sets of virtual nodes mapped on the working ED connecting to ONU and at wireless routers are illustrated by s and ϕ, respectively. The sets of consumed radio bandwidth and computing resources for the working EDs are denoted by wb and c, respectively. The following steps describe how to embed the virtual nodes onto the working graph.

We first choose a working ONU-level ED to find the mapping solution. After that, we select a collection of working EDs at wireless routers to hold the other virtual nodes, so that all the mapped EDs are with more computing resources than the required resources for VN serving. In addition, the required bandwidth of wireless channels for link mapping should be larger than the consumed bandwidth for mapping by network virtualization. Then, the VN model becomes a two-tuple model denoted by $vn(\mathrm{Ew}(s, \phi), c)$, containing the collection of working EDs mapped by the virtual nodes and the consumed computing resources by EDs, respectively. Once the locations of k backup EDs are confirmed, the VN model is embedded onto the backup graph Γ_b by selecting the backup ED with the highest *SD* among the remaining candidates, whose computing resources are no less than *c*. After the current network is mapped onto Γ_b, the maximal *SD* equals the number of the virtual nodes mapped onto the ED.

We denote the number of embedded VNs processed by WMNs or IoTs by N_1. The binary variables $\eta^i_{\Gamma_\omega}$ and $\eta^i_{\Gamma_b}$ identify whether the *i*th VN is embedded onto the working or backup graph of the qualified WMN or IoT (Γ_w or Γ_b), respectively. For $i \in [1, N_1]$ and $\Gamma_b \in [1, n]$, the total assigned transmission power (*TP*) to EDs is the summation of $\eta^i_{\Gamma_\omega}$ multiple the required *TP* on the working graph TP_{Γ_ω}. Similarly, the maximum *SD* of the backup resource can be calculated by the sum product of $\eta^i_{\Gamma_b}$ and the *SD* of the backup ED on Γ_b.

If there exists no local bottleneck, the objective function of our framework is to minimize the allocated *TP* to EDs for each backup resource *SD*. Otherwise, we assign another backup ONU-level ED from WMNs or IoTs, so that $(1 + |\phi|)$ virtual nodes are within the VN.

We denote N_2 by the number of VNs successfully processed by the PON backhaul. The number of consumed optical fiber cables can be obtained by rounding the fraction between N_2 and the number of VNs served by optical fiber cables. The number of consumed backup carriers is rounding the fraction between the number of VNs served by the optical fiber cables multiples *wb* and the carrier's initial bandwidth provisioning on the optical fiber cable (*ba*). Therefore, the total number of consumed backup carriers to serve VNs is rounding the fraction between N_2 multiple *wb* and *ba*. The overall optimization objective is to minimize the integration of two parts, that is, the costs of local WMN or IoT embedding and PON-backhaul embedding.

Heuristic Algorithm

Due to the high computational complexity of the integral linear programming problem, we design a heuristic algorithm to solve the formulated optimization problem iteratively. The process of the GSVNE for collaborative edge computing can be expounded by the following steps.

Step 1: Initialize the corresponding information of WOBANs and VNs according to the network status and information.

Step 2: Confirm the amount of backup EDs at each WMN or IoT according to the reliability function:

$$P_r(k) = (1 - P_r)^k \times P_r^{(|M \cup P| - k)} \quad (1)$$

where P_r denotes the failure probability of ED, and $(|M \cup P| - k)$ is the residual EDs of available wireless routers for working resource units. Deploying more backup EDs can increase the value of *k*, resulting in a higher network reliability. If $P_r < 0.5$, the value of *k* should be increased until the ED failure probability decreases below a threshold.

Step 3: For $\forall u \in M$, $\forall v \in P$, we deploy the residual $(k - |M|)$ backup EDs; satisfying the minimum hop of *u* and *v* is no more than the max-hop count between the locations of the ONU-level ED and other wireless routers.

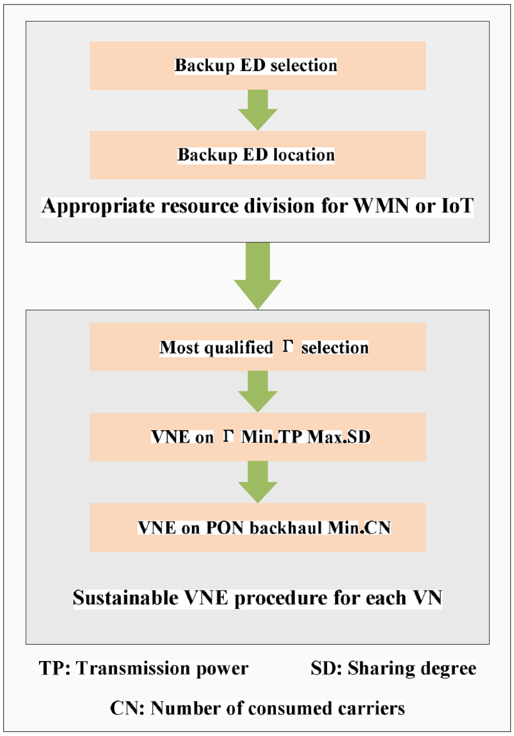

Figure 2. Framework of GSVNE.

Step 4: The resources of the initial and residual backup EDs are divided into two parts according to a scaling factor *f*. If the ED is not deployed at the wireless router, the initial backup resource provisioning is 0.

Step 5: Select a qualified WMN or IoT for VN serving. For each WMN or IoT, the vertexes with lower computing capacity than *c* on Γ_w and Γ_b are removed together with the corresponding wireless channels. Then, the weight factor of wireless channel (u, v) is updated by decreasing the required weighting factor among all the outgoing channels of vertex *u* iteratively. Afterward, we remove the wireless channels whose weighting factors are 0 to lower the assigned *TP* further. If the *SD* of the available backup graph reaches the maximum value, the most qualified network can be decided under the condition that the corresponding path is also available.

Step 6: If the most qualified WMN or IoT is available in the current VN, the amount of N_1 is added by 1. Otherwise, we choose the backup ED from different WMNs or IoTs by iteratively decreasing the consumed computing resources of backup EDs, and the quantity of N_2 is added by 1. The VNE process fails if the backup ED is unavailable; then the quantity of unembedded VNs is increased by 1. The VNE process continues until the last VN can be processed successfully. Otherwise, go back to Step 5.

The flow chart of the designed algorithm is illustrated in Fig. 3.

We first choose a working ONU-level ED to find the mapping solution. After that, we select a collection of working EDs at wireless routers to hold the other virtual nodes, so that all the mapped EDs are with more computing resources than the required resources for VN serving.

Performance Evaluation

We select the map of the UC Davis campus [13] for smart university community evaluation. In this simulation, each WMN or IoT has the same number of EDs for simplification. As illus-

Providing secured connections to the massive devices in CoTs with various processing abilities is also significant. On one hand, the decentralized feature of EDs in WMNs or IoTs is vulnerable; on the other hand, D2D-enabled communication is advocated to reduce latency and balance network loading.

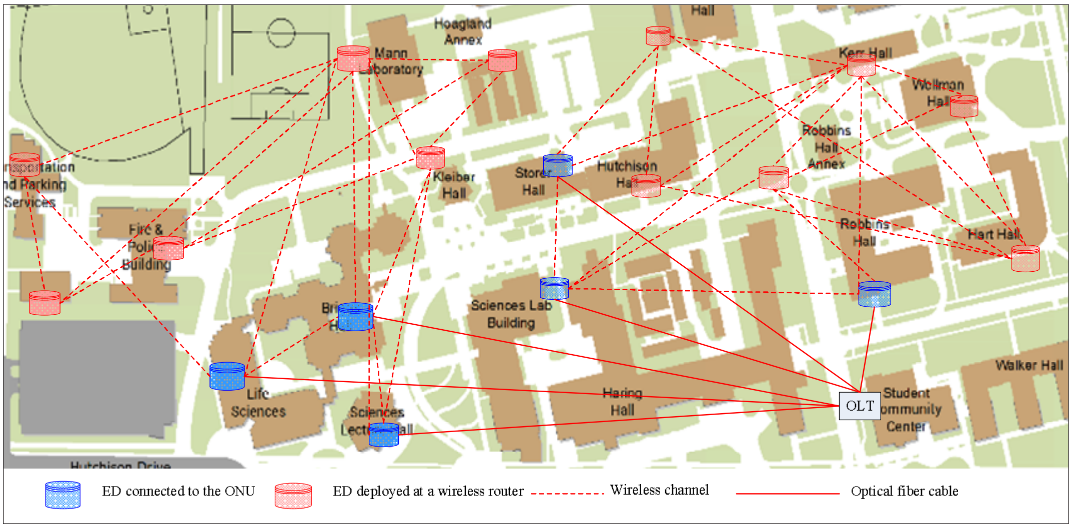

Figure 4. The map of UC Davis campus.

trated in Fig. 4, local EDs together with three EDs and six wireless routers are deployed in the test WOBAN. For the VNs, we set $|\phi| = 2$, $c = 1$, $wb = 1$, and $ba = 6$.

The consumed *TP* between GSVNE and the benchmark without graph cutting is demonstrated in Fig. 5a. As the horizontal axis of Fig. 5a shows, different *f* values correspond to various numbers of N_1. We can observe that the total *TP* enhances as the number of embedded VNs increases, and the total consumed power of GSVNE is decreased by around 10 percent. The main reason is that the working graph for each WMN or IoT can be simplified by reserving wireless channels with low weights in GSVNE.

Next, we fix *f* to 0.5, and evaluate the performance of the total number of embedded VNs. When the circumstance $N > N_1 = 60$ happens, the VNs cannot be embedded by the WMNs or IoTs. As shown in Fig. 5b, it is obvious that the number of the served VNs by the benchmark without PON backhaul is always along the border line caused by a local bottleneck. However, the sustainability of our method can be well guaranteed with more VNs because the backup EDs can still be selected via the PON backhaul. The increase trend remains non-linear after $N = 110$, because network resources of the ONU-level backup EDs are also constrained.

Figure 5c compares the maximum *SD* of the backup resources among the theoretical upper bound, GSVNE, and the benchmark without PON backhaul. Since a portion of backup resources offered by EDs is wasted without PON backhaul, the *SD* of the benchmark is constant and equals 30, well corresponding to the number of served VNs in Fig. 5b. The maximum *SD* in our method does not rapidly approach the border line at first, since the PON backhaul is not indispensable when the number of VNs is less. The maximal *SD* of GSVNE remains stable after $N = 80$, well illustrating why the performance of GSNVE remains stable after $N = 110$ in Fig. 5b. Finally, Fig. 5d illustrates the consumed amount of backup carriers between our method (*W_GSVNE*) and the optimal bound (*W_bound*). The utilized number of backup carriers by GSVNE increases as the total number of VNs increases. We notice that the gained performances between GSVNE and the optimal bound overlap when $N = [60, 100]$, which demonstrates the effectiveness of our method to guarantee network sustainability. After that, W_GSVNE is lower than *W_bound* since the network resources of PON backhaul are constrained.

Research Challenges

There are still many challenges and open issues to enable collaborative edge computing in CoTs. In this section, we discuss some important aspects of them as follows.

Security and Privacy

Although emerging services are enabled by edge computing, its novel features challenge the current security and privacy frameworks. The main reasons are:

- Heterogeneous network structures and deployment of edge computing in WOBANs cause conventional authentication and trust schemes to lose efficiency.
- Although novel communication and software technologies are designed to support edge computing, it is challenging to judge the role of an edge server, that is, whether it is an attacker or an eavesdropper.
- Virtualized infrastructure at the edge is vulnerable to privacy invasion or denial of service attacks caused by virtual machine manipulation.

Furthermore, providing secured connections to the massive devices in CoTs with various processing abilities is also significant. On one hand, the decentralized feature of EDs in WMNs or IoTs is vulnerable; on the other hand, D2D-enabled communication is advocated to reduce latency and balance network loading. Some other research challenges still exist, such as updating software and security credentials in large-scale CoTs, protecting resource-constrained devices, acquiring the security status of devices in CoTs, and coping with security threats by safe and continuous operations [6].

Scalability and Energy Efficiency

Extensive connections and communications between cloud services and devices generate huge incoming/outgoing network traffic, challenging the scalability and energy efficiency of CoTs significantly. The joint optimization of radio-and-computa-

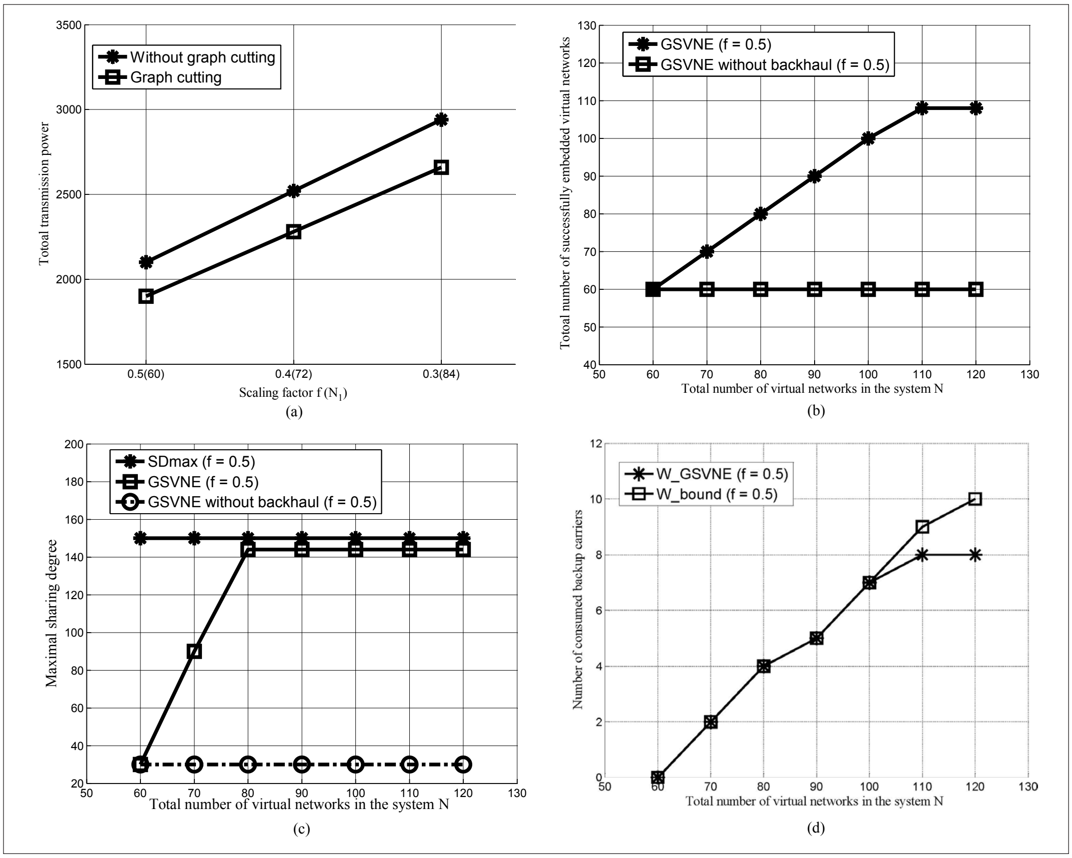

Figure 5. Performance evaluation: a) total TP; b) total number of successfully embedded VNs; c) maximal SD; d) number of consumed backup carriers.

tional resource management is promising to fulfill energy-efficiency edge computing. Different from on-demand elastic cloud resources, many devices in WMNs or IoTs are supplied by battery. The massive interaction in CoTs without network optimization would exhaust the device's energy quickly. Computational offloading from devices in WMNs or IoTs to edge devices is promising, and D2D-enabled mobility-aware offloading has become a hot research topic. One possible offloading solution is to jointly consider a user's mobility information, computational ability and channel state. For example, computation-intensive data can be offloaded to the edge servers to reduce latency, and large-size data can be handled by D2D communications to improve energy efficiency [14]. However, D2D communications would generate severe interference, which challenges network scalability. Therefore, interference cancelation and adaptive power control should also be involved in edge computing.

Data Abstraction

Although data abstraction has been widely discussed in cloud computing, it is still challenging in edge computing. On one hand, a huge number of things communicates with the EDs at network edge; on the other hand, the deployed EDs report data to the gateway periodically, such as smart water control, light control and door locks in smart home 1 in Fig. 1. Then the processed data can be sent to the upper layer for service providing. However, challenges exist:

- The formats of the delivered data from devices are various.
- A trade-off has to be made between the degree of data abstraction and data storage.
- How to extract useful information reported by the various and unreliable edge devices.
- Different applicable operations coexist on edge devices [10].

Caching at EDs

Cache placement and content delivery are two main processes for caching. Although various research works have focused on cache updating during the placement process, online caching is still very challenging for EDs. Recently, cache deployment in mobile core networks has been studied, but how to alleviate the constrained bandwidth of the backhaul links is still challenging. For the content delivery process, it would be beneficial if a user's preferences can be predicted, so that the content that the user is interested in can be cached beforehand. Although the techniques of data mining and machine learning have developed rapidly, how to predict content popularity accurately in a scalable manner is still quite open. Furthermore, it is important to study the cooperation among caches by integrating social networks and D2D communications. Also, how to design a highly-efficient incentive scheme for cooperative caching among EDs is also challenging [15].

Conclusion

In this article, we first emphasized the significance of sustainable edge computing in CoTs, and discussed research directions for edge computing. Next, a novel GSVNE framework was presented

Cache placement and content delivery are two main processes for caching. Although various research works have focused on cache updating during the placement process, online caching is still very challenging for EDs. Recently, cache deployment in mobile core networks has been studied, but how to alleviate the constrained bandwidth of the backhaul links is still challenging.

The performance analysis showed that our framework can embed more VNs and effectively approach the maximal SD of backup resources compared with the existing schemes. Although our method is an alternative for high energy-efficiency and sustainable edge deployment in CoTs, the corresponding avenues are not flat.

for collaborative edge computing in WOBANs, which can confirm the number of backup EDs and embed the VNs onto the suitable EDs. The performance analysis showed that our framework can embed more VNs and effectively approach the maximal SD of backup resources compared with the existing schemes. Although our method is an alternative for high energy-efficiency and sustainable edge deployment in CoTs, the corresponding avenues are not flat. Finally, we listed future research challenges and open issues.

Acknowledgment

This work is partially supported by the National Natural Science Foundation of China (61401082, 61572106 and 61502075); the China Postdoctoral Science Foundation (2018T110210); and the Fundamental Research Funds for the Central Universities under grant (DUT17RC(4)49).

References

[1] N. Luong *et al.*, "Resource Management in Cloud Networking Using Economic Analysis and Pricing Models: A Survey," *IEEE Commun. Surveys & Tutorials*, vol. 19, no. 2, 2017, pp. 954–1001.
[2] Cisco visual networking index: Global mobile data traffic forecast update, 2015–2020, Cisco White paper, 2016.
[3] Z. Ning *et al.*, "A Cooperative Quality Aware Service Access System for Social Internet of Vehicles," *IEEE Internet of Things J.*, vol. 5, no. 4, 2018, pp. 2506-17.
[4] Z. Ning *et al.*, "Vehicular Social Networks: Enabling Smart Mobility," *IEEE Commun. Mag.*, vol. 55, no. 5, 2017, pp. 49–55.
[5] W. Hou *et al.*, "Green Survivable Collaborative Edge Computing in Smart Cities," *IEEE Trans. Industrial Informatics*, vol. 14, no. 4, 2018, pp. 1594-1605.
[6] M. Chiang and T. Zhang, "Fog and IoT: An Overview of Research Opportunities," *IEEE Internet of Things J.*, vol. 3, no. 6, 2016, pp. 854–64.
[7] B. Rimal *et al.*, "Mobile Edge Computing Empowered Fiber-Wireless Access Networks in the 5G Era," *IEEE Commun. Mag.*, vol. 55, no. 2, 2017, pp. 192–200.
[8] Z. Ning *et al.*, A Social-Aware Group Formation Framework for Information Diffusion in Narrowband Internet of Things, *IEEE Internet of Things J.*, vol. 5, no. 3, 2018, pp. 1527-38.
[9] T. Tran *et al.*, "Collaborative Mobile Edge Computing in 5G Networks: New Paradigms, Scenarios, and Challenges," *IEEE Commun. Mag.*, vol. 55, no. 4, 2017, pp. 54–61.
[10] W. Shi *et al.*, "Edge Computing: Vision and Challenges," *IEEE Internet of Things J.*, vol. 3, no. 5, 2016, pp. 637–46.
[11] K. Liang *et al.*, "An Integrated Architecture for Software Defined and Virtualized Radio Access Networks with Fog Computing," *IEEE Network*, vol. 31, no. 1, pp. 80–87.
[12] A. Baktir *et al.*, "How Can Edge Computing Benefit from Software-Defined Networking: A Survey, Use Cases & Future Directions," *IEEE Commun. Surveys & Tutorials*, vol. 19, no. 4, 2017, pp. 2359–91.
[13] R. Rofoee *et al.*, "Hardware Virtualized Flexible Network for Wireless Data-Center Optical Interconnects," *IEEE/OSA J. Optical Commun. and Networking*, vol. 7, no. 3, 2015, pp. A526–36.
[14] Y. Mao *et al.*, "A Survey on Mobile Edge Computing: The Communication Perspective," *IEEE Commun. Surveys & Tutorials*, vol. 19, no. 4, 2017, pp. 2322–58.
[15] T. Tran *et al.*, Cooperative Hierarchical Caching in 5G Cloud Radio Access Networks, *IEEE Network*, vol. 31, 2017, no. 4, pp. 35-41.

Biographies

Zhaolong Ning received the M.S. and Ph.D. degrees from Northeastern University, Shenyang, China. He was a research fellow at Kyushu University, Fukuoka, Japan. He is an associate professor in the School of Software, Dalian University of Technology, China. His research interests include edge computing, vehicular networks and network optimization.

Xiangjie Kong [M'13, SM'17] received the B.Sc. and Ph.D. degrees from Zhejiang University, Hangzhou, China. He is currently an associate professor in the School of Software, Dalian University of Technology, China. His research interests include computational social science, urban big data, and mobile computing.

Feng Xia [M'07, SM'12] received the B.Sc. and Ph.D. degrees from Zhejiang University, Hangzhou, China. He is currently a full professor in the School of Software, Dalian University of Technology, China. His research interests include computational social science, network science, data science, and mobile social networks.

Weigang Hou received the Ph.D. degree in communication and information systems from Northeastern University, Shenyang, China, in 2013. He is currently an associate professor with the School of Computer Science and Engineering, Northeastern University, Shenyang, China. His research interests include Internet of Things, optical network+traffic grooming, and SDN.

Xiaojie Wang received the Master degree from Northeastern University, China, in 2011. From 2011 to 2015, she was a software engineer atn NeuSoft Corporation. Currently, she is pursuing a Ph.D. degree with Dalian University of Technology, China. Her research interests are edge computing, vehicular networks, and social computing.