Where Resources meet at the Edge

Klervie Toczé and Simin Nadjm-Tehrani
Dept. of Computer and Information Science
Linköping University, Linköping, Sweden
Email: [klervie.tocze, simin.nadjm-tehrani]@liu.se

***Abstract*—Edge computing is a recent paradigm where network nodes are placed close to the end users, at the edge of the network. Efficient management of resources within this configuration is crucial due to scarcity and geographical spreading of edge resources. We begin by a brief description of the edge paradigm, the most generic edge architecture, and the terminology associated to it. Then, we propose and elaborate on a preliminary taxonomy for edge resource management, together with a substantial review of works in the area. Finally, we identify some research challenges.**

I. Introduction

Edge computing, a paradigm promoting network nodes with computational and storage resources close to the devices (mobile phones, sensors), at the edge of the current network, promises a new communication era in which industry can meet the rising performance needs of future applications.

Indeed, with a forecast of 8.9 billion mobile subscriptions in the world by 2022, of which 90% will include mobile broadband, coupled to an eightfold increase in mobile traffic and 18.1 billion of Internet of Things (IoT) devices also sending data [1], there will be a considerable strain put on the network. The current network technologies need to undergo a paradigm shift in order to respond to that [2]. The idea of edge (or fog) computing is to avoid overwhelming the network up to the cloud with data by, if possible, moving some computing and data analysis closer to the users, to enable better scalability [3].

Moving computing and storage to the edge of the network has other benefits [3] such as reducing the latency and jitter [4], which is especially important for real-time applications such as self-driving cars. In addition, it enables more privacy for the users by making it possible to keep private data at the edge and to enforce privacy policies for the data sent to the cloud (such as blurring sensitive info on a video [2]). Finally, it makes the network more resilient by being able to still process requests at the edge even if the central cloud is down.

In order to achieve this and to make edge computing a reality and a success, there is a need for efficient resource management at the edge. Indeed mobile devices or IoT devices are resource-constrained devices, whereas the cloud has almost unlimited but far away resources. Providing and/or managing the resources at the edge will enable the end device to spare resources (e.g. stored energy in batteries), speed up computation and make it able to use resources it does not possess. Moreover, keeping data close to where it was generated enables better control, especially for privacy related issues. Finally, being located close to the user, edge computing makes it possible to increase the quality of provided services through the use of profiling within a local context, without compromising the privacy or having to handle a large number of users. This is known as *context adaptation*.

Even though this is still an emerging research area, there is a lot of work ongoing under different denominations (mobile cloud computing [5], fog computing [6], edge computing [3], mobile edge computing [7], mobile edge cloud [8], and distributed cloud [9]). A common concept here is an intermediate level between the device and the traditional cloud. It is possible to find in the literature numerous surveys about those paradigms in general [5], [10], [11], [12], [8] or specific aspects of them such as security [7]. However, those most often do not consider the resource aspect and the existing surveys about resources either consider it at a high level [13], consider only resource/service provisioning metrics [14] or focus on offloading [15], [16], [17], i.e. executing a task in another device than the requesting device, usually one having more powerful computational capacities.

However, there are more areas in resource management than just offloading and this paper aims at providing a brief overview of the current work done in those areas, focusing on work where the resources are at the edge or where the resource management is performed at the edge. Therefore, work considering direct interactions from a device to a cloud [18] or from the cloud to the edge [19] are not considered. This paper is a preliminary version of a larger survey on edge computing with resource management as a focus [20].

In the remaining parts of this paper, we will explain the terminology used in Section II, and then propose and describe a taxonomy of resource management at the edge in Section III. In Section IV we will discuss some research gaps and Section V will conclude this paper.

II. Terminology

Edge computing is an innovative area that brings together previously separate parties such as telecommunication actors, vehicle vendors, cloud providers, and actors from the Internet domain (for example Google with the Google glasses). Therefore, the terminology used in research work is varying and still evolving. In this section, we present the relevant concepts associated to edge computing which will be used in the rest of the paper.

Following the development of the IoT, it is nowadays not only computers or smartphones which can be connected to the

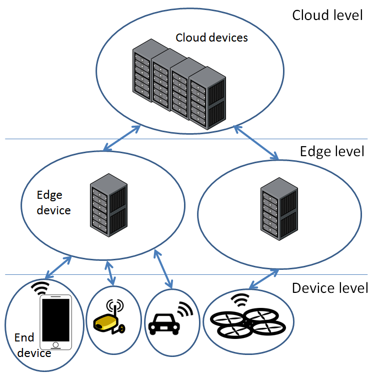

Fig. 1. Three-level architecture used in edge computing.

network, but a large variety of things such as cars, sensors, or home appliances. In this paper, all those objects located at the user end of the network, which are producing data or are in need of cloud/edge resources will be called *end devices*. Devices installed at the edge specifically for edge computing purposes are called *edge devices*. Finally physical components of the cloud are designed by the term *cloud devices*. Resources which are managed are used to perform *tasks*, which can be composed to provide a *service* to the user.

There is currently no standard architecture for edge computing, although there exists industry and research initiatives such as the Open Edge Computing[1] community, and a European Telecommunications Standards Institute (ETSI) standardization group working on Multi-access Edge Computing[2]. The generic architecture used for our review is a three-level architecture as presented in Figure 1. In this architecture, end devices (forming the first level, called the *device level*) are connected to an edge device (located within the second level, the *edge level*), which itself is connected to the rest of the network, including the cloud (forming the third level, the *cloud level*). The edge device usually has relatively high computational power, though it remains less powerful than a conventional datacenter used in the cloud computing paradigm. In the literature, such edge devices are named for example cloudlets[21], [22] or micro datacenters [23] and they can be located for example in shops, companies or co-located with the base stations of the telecom access network. Indeed, in the ongoing work being done on what the fifth generation (5G) of telecommunication networks will look like, a cloud radio access network (C-RAN) is envisaged [24], [25], with possible interwork with other edge computing areas such as mobile cloud computing [26].

III. Taxonomy of Edge Resource Management

In this section, we present the core elements of a taxonomy of resource management at the edge with the aim to get an overview of research in this area. This taxonomy, illustrated in Figure 2, presents three main aspects: resource type, objective of resource management, and resource use. In the coming subsections, we will describe the different parts of the taxonomy, and how they have been addressed in existing work. The full survey [20] substantially extends this taxonomy (which is avoided here due to space restrictions).

A. Resource type

The first aspect of resource management at the edge is the identification of the resources being managed. From the surveyed works, resources can be classified into two main types: physical and virtual.

1) Physical: The majority of the surveyed articles consider the management of physical resources. Commonly considered resources are computational or storage resources (such as CPU), needed to execute the task. Other common resources are communication resources, such as bandwidth or spectrum, since edge or end devices are not isolated. For example, Liu et al [27] consider mostly wireless bandwidth and computing resource when deciding to handle a request at the edge or the cloud level.

Other works consider the storage resources that can be needed for applications with more data. Confais et al. [28] present how a storage service can be provided for fog/edge infrastructure, based on the InterPlanetary file system and scale-out network attached systems. The aim is to propose a service similar to the Amazon Simple Storage Service solution[3] for the edge.

The next resource which is considered in the reviewed articles is energy. Efficient use of energy is important for two main reasons. First, mobile devices (and especially smartphones) are constrained by their limited battery life. Second, there is a rising awareness about the need to use less energy and, when possible, to use more green energy (i.e. renewable energy). Fan et al. [29] thus present a virtual machine migration scheme which aims at using as much green energy as possible in the context of green edge devices (in this case, cloudlets) networks. A comprehensive study of energy apportionment policies in several computing and communication domains has been presented by Vergara et al. [30] and has a natural bearing on energy accounting and apportionment in the edge.

Finally, Yousaf et al. [25] emphasize the fact that different resources should not be considered in isolation as there are interactions between them. Thus, they describe and use the concept of resource affinity in their scheme.

2) Virtual: The second resource type found in the literature is virtual resources, classified into two categories.

First, virtual resources can be represented through the use of virtual machines (VMs). This is the case of Gu et al. [31], Tärneberg et al. [32], and Plachy et al. [33].

Another approach is to abstract the physical resources into virtual resources and manage those instead. Works using this approach define new units, e.g. Aazam et al. [23], who utilize a Virtual Resource Value as the unit for resources, thereafter

[1] http://openedgecomputing.org/

[2] http://www.etsi.org/technologies-clusters/technologies/multi-access-edge-computing

[3] https://aws.amazon.com/fr/s3/

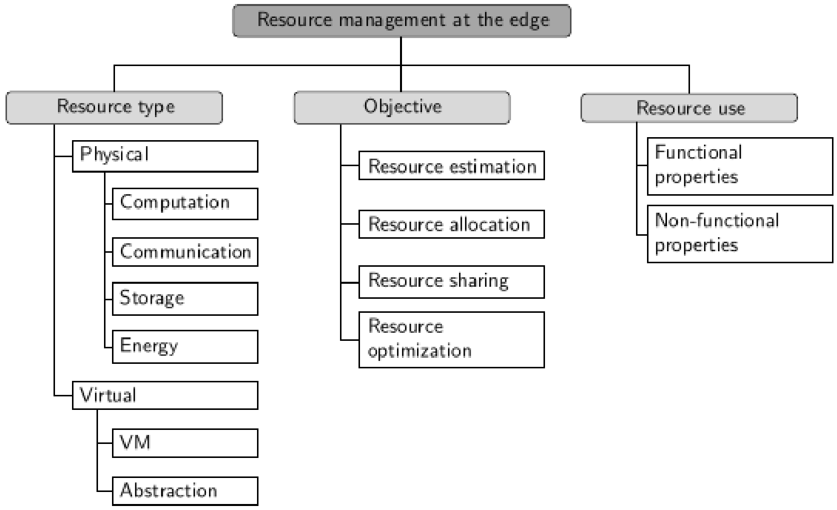

Fig. 2. A taxonomy of resource management at the edge.

mapped to physical resources according to the type of service and current policies of the cloud service provider. In their work, they provide examples on how the mapping could be done.

Finally, the border between physical and virtual resources is definitely not a strict one and several articles are considering both of them. For example, Liu et al. [27] consider VMs at the same time as physical resources such as wireless bandwidth and computing resources, and Yousaf et al. [25] also mention VMs when describing the concept of resource affinity.

B. Objective

The second main element presented in this taxonomy is the objective of resource management. Indeed, resource management at the edge can be decomposed into several areas addressing different problems. One research work can address several of the areas.

1) Resource estimation: This is the ability to estimate how much resource will be needed to complete a task, or carry a load. This is important, especially for being able to handle fluctuations in resource demand while maintaining a good quality of service (QoS) for the user. Indeed, resources at the edge can be mobile, and thus become unreachable, which makes them not as reliable as resources in a datacenter. Moreover, user mobility implies that there can be sudden user churn, with the corresponding service needs having to be handled by the edge.

In their work, Aazam et al. [23] argue that the fog can estimate the future resource consumption so that resources can be allocated in advance. They formulate an estimation of the required resources which takes into account the reliability of the customer, into what they call the relinquish probability. In another article [34], they present the same idea but with an emphasis on how different customers can be charged for the service.

2) Resource allocation: The next area is to actually allocate resources so that the task can be executed. This problem can be tackled from different perspectives: where to allocate, when to allocate, and how much to allocate. Different solutions are provided in the surveyed work depending on how the considered system looks like and which constraints it has.

Some works tackle all the three different perspectives, as the work by Liu et al. [27]. It presents a multi-resource allocation system which first decides whether the request should be served or rejected, then where to run it (edge or cloud level), and finally how much bandwidth and computing resources should be allocated for this task. Their aim is to maximize the benefits of the system while guaranteeing QoS for the users. However, most of the surveyed works emphasize the first perspective, i.e. where should the task be executed and the resource allocated for the best possible execution. The definition of best execution varies depending on the considered system and the focus of the research.

Oueis [35] tackles the issue of load distribution and resource allocation in small cell clusters. She formulates a joint computational and communication resource allocation and optimization problem in a multi-user case with focus on latency and power efficiency. Borylo et al. [36] investigate dynamic resource provisioning. They present a policy in which the edge can use the cloud in compliance with the latency requirements of the edge but enables a better energy efficiency by using resources in datacenters powered by green energy.

There is also research done in specific aspects of the edge. For example Wang et al. [26] propose a joint cost-effective resource allocation between the mobile cloud computing infrastructures and the cloud radio access network infrastructure.

Several researchers approach the resource allocation problem by studying where different virtual entities such as applications or VMs should be created and executed and how they can be moved during execution if the new location is better. For example, Tärneberg et al. [32] study application placement and present a system model for mobile cloud network and an application placement algorithm with the aim of guaranteeing application performance, minimizing cost, and tackling the resource asymmetry problem. Yousaf et al. [25] propose a VM migration system that takes into account the relationship between resource units when taking migration decisions. They present this work in the context of 5G but it should be applicable to all physical machines hosting VMs. Gu et al. [31] and Plachy et al. [33] study VM placement. In the former, it is done from a cost perspective (both VM deployment cost and inter base station communication cost) in medical cyber physical systems, while in the latter they include user mobility predictions in their algorithm.

Finally, resource allocation can also be tackled from other perspectives, such as creating a storage service for edge applications [28] or as a part of creating an edge-based data analytics scheme [37].

3) Resource sharing: Resources on end devices are heterogeneous and most of the time scarce, and edge devices also have limited resources compared to cloud devices. Sharing resources between devices or between end and edge devices aims at tackling three different issues: not having the needed resource in the device where the task is initiated, not having enough of it or using other devices' resources in order to get a faster completion of the task.

However, even if resource sharing can bring benefits for a group of end devices, it is not obvious that users will agree to share their resources, especially if they are always on the providing side. Therefore, there is a need to develop incentives for resource sharing such as the work by Tang et al. [38]. They propose a double bidding mechanism for demander and supplier of resource where the focus is on how to encourage mobiles with resources to share them, taking into consideration two cases: price-taking and price-anticipating users.

4) Resource optimization: Finally, a fourth objective pursued in the surveyed works is to optimize the resources used at the edge. This is usually a joint objective together with one of the previously described objectives. Which aspect should be optimized and the associated constraints varies among the works but the four main aspects are latency, energy, communication cost and financial cost. The way the optimization problem is formulated and solved also varies.

One of the aims of edge computing being to decrease latency compared to cloud solutions, Oueis [35] tackles the issue of load distribution and resource allocation with the aim of being both latency- and energy-efficient.

Energy is indeed an important aspect of computing at the edge, and some works focus on optimizing with regards to this aspect, for example by optimizing the use of green energy: Do et al. [39] want to minimize the carbon footprint of video streaming service and Fan et al. [29] want to optimize grid consumption in order to use more green energy.

Finally, computing at the edge also has a communication cost and a financial cost. For example, if some service provider wants to use resources from the edge devices but is not the owner of those devices, they will have to pay to use the resources, in a similar way to what is done today with cloud resources. This is tackled by Wang et al. [26] who propose optimization with regards to expenditure cost for using resources in a mobile cloud computing solution. Finally, Gu et al. [31] and Arkian et al. [37] consider deployment cost as well as communication cost between the different actors.

C. Resource use

The third main aspect considered in this article is the purpose for which the resources will be used.

1) Functional properties: Indeed, in most of the surveyed articles, resources are used in order to get the requested service, i.e. for satisfying functional properties. For example, resources can be used for video stream services [39] or crowd sensing applications [37].

2) Non-functional properties: However, when using the edge computing paradigm, you also can get desirable additional properties not directly related to the service to obtain, as well as properties which are consequences of how the edge is organized. All those properties will require additional work to be performed.

For example, Fan et al. [29] consider the cost of VM migration, where the migration is done in order to be more energy-efficient. This is not part of providing the service but is improving the way it is provided. Another non-functional property is the fact that we can't have devices sharing resources without incentives. This is tackled by Tang et al. [38].

It is interesting to note that none of the surveyed papers analysed the overhead of the mechanisms that provide the properties enforced by edge computing and which are advocated as good reasons to go towards this paradigm, such as privacy [40] or context adaptation [41].

IV. Discussion

Resource management at the edge is still a new area, and it is possible to identify areas with research challenges that can be taken care of by new researchers coming into this field.

From the surveyed works, it appears that edge computing is a paradigm which brings together diverse domains and actors. Therefore, techniques which were previously applied in only one of those domains may be applicable to edge computing with the required adaptations. Research has to be done in order to provide those adaptations and evaluate if the techniques are performing well within the edge paradigm. For example, network function virtualization (NFV), which is studied in the telecommunication domain [25] and seen as an enabler for 5G networks [42], could be considered in other edge computing areas as well. Software-defined networking [36], [43] and techniques used in distributed cloud computing [9] are other examples of those techniques.

When looking at the different objectives for resource management at the edge, we can notice that resource allocation dominates most of the research work, whereas other areas such as resource estimation or resource sharing are less researched. However, they are very important in order to make computing at the edge flexible and scalable. Regarding resource sharing, it can be considered at each level (between end devices but also between edge devices) but also across levels.

Another aspect which is important in order to evaluate the proposed solution is the end device mobility. Indeed, a lot of the scenarios considered for edge computing are large-scale and with very mobile users. It is therefore essential to have efficient and realistic mobility models reflecting the end devices' mobility in the case of edge computing. Due to the large diversity of end devices which can be used and their different characteristics, there is a need to have different models according to different scenarios. Potentially edge solution intended to serve networks of cars moving on a road network will have different characteristics from an edge solution intended to serve persons within a shopping mall.

Moreover, and as already mentioned by other researchers [14], [44], more tools need to be developed to test the new proposals in relevant conditions and preferably in realistic setups. Indeed, the surveyed works are using different methods for testing the contributions. The most common method used to validate the model or algorithm proposed and/or to solve the optimization problem formulated is the use of an analytical tool (e.g. an optimisation engine). Quite a lot of papers do not state the (solver) tool that they use, but when they do, Matlab[4] is the most common one [33], sometimes associated with other tools or APIs such as lp_solve[5] [27] or CVX[6] [26]. Other tools mentioned are IBM CPLEX[7] [29] and Gurobi[8] [31].

Another common approach is to use a simulator. It can be a simulator designed for regular cloud environments, such as CloudSim [45] used by Aazam et al. [34], but there also exists a dedicated simulator designed for fog computing, called iFogSim [46], which extends CloudSim. Another simulator used by Borylo et al. [36] is the OMNeT++[9] simulator. Custom simulators are also present among the platforms e.g. Tärneberg et al. [32] implemented an event-driven simulator based on SimPy[10].

Finally, there are works that evaluate their approach using physical testbeds. Confais et al. [28] use the testbed Grid'5000[11], which is a large-scale testbed focusing on parallel and distributed computing. In the literature reviewed, we also found Cumulus [47], a platform for computational offloading at the edge proposed by Gedawy et al., and which represents a testbed implementation. Lastly, Yousaf et al. [25] mentioned that they plan to create a testbed in their future work. However, most of the simulators or testbeds used in the edge computing area are recent, so there is a need for more work on evaluating them, improving them, and extending them. More specific application domains may require completely new tools if the current ones do not have the possibility to cover those domains as well.

Last but not least, we could observe that although there is a lot of work about resource management for resources which will be used to provide functional properties, there is not a lot of work on the resource footprint of edge algorithms dealing with aspects such as security, privacy or context adaptation. Research about those aspects can be found for the mobile cloud paradigm for example [48], [49], but they consider scenarios where the edge level is absent. Since edge computing cannot become widely used without strong security and privacy properties, it is especially important to research on the resource overhead for providing those properties, since a too high overhead can signify a technology which is not usable in practice.

V. Conclusion

In this paper we briefly presented the terminology associated with resource management at the edge together with a preliminary taxonomy and a short overview of some of the research in the area. The work is currently extended towards a major survey with additional taxonomy concepts and classifications [20]. We briefly alluded to some important findings and identified challenges which need to be researched on.

In addition to finalising the major survey, in our own future work, we plan to look deeper into the resource overhead of non-functional properties in edge computing.

Acknowledgment

This work was supported by the Swedish national graduate school in computer science (CUGS).

References

[1] Ericsson. (2016, Nov) Ericsson mobility report. [Online]. Available: https://www.ericsson.com/assets/local/mobility-report/documents/2016/ericsson-mobility-report-november-2016.pdf

[2] M. Satyanarayanan, P. Simoens, Y. Xiao, P. Pillai, Z. Chen, K. Ha, W. Hu, and B. Amos, "Edge analytics in the internet of things," *IEEE Pervasive Computing*, vol. 14, no. 2, pp. 24–31, Apr 2015.

[3] M. Satyanarayanan, "The emergence of edge computing," *Computer*, vol. 50, no. 1, pp. 30–39, Jan 2017.

[4] M. Etemad, M. Aazam, and M. St-Hilaire, "Using DEVS for modeling and simulating a fog computing environment," in *2017 International Conference on Computing, Networking and Communications (ICNC)*, Jan 2017, pp. 849–854.

[5] N. Fernando, S. W. Loke, and W. Rahayu, "Mobile cloud computing: A survey," *Future Generation Computer Systems*, vol. 29, no. 1, pp. 84 – 106, 2013.

[6] F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, "Fog computing and its role in the internet of things," in *Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing (MCC '12)*. New York, NY, USA: ACM, 2012, pp. 13–16.

[7] R. Roman, J. Lopez, and M. Mambo, "Mobile edge computing, fog et al.: A survey and analysis of security threats and challenges," *Future Generation Computer Systems*, 2016, http://dx.doi.org/10.1016/j.future.2016.11.009.

[8] H. Liu, F. Eldarrat, H. Alqahtani, A. Reznik, X. de Foy, and Y. Zhang, "Mobile edge cloud system: Architectures, challenges, and approaches," *IEEE Systems Journal*, 2017, preprint, https://doi.org/10.1109/JSYST.2017.2654119.

[9] P. T. Endo, A. V. de Almeida Palhares, N. N. Pereira, G. E. Goncalves, D. Sadok, J. Kelner, B. Melander, and J. E. Mangs, "Resource allocation for distributed cloud: concepts and research challenges," *IEEE Network*, vol. 25, no. 4, pp. 42–46, July 2011.

[4] https://www.mathworks.com/products/matlab.html
[5] http://lpsolve.sourceforge.net/5.5/
[6] http://cvxr.com/cvx/
[7] http://www-03.ibm.com/software/products/fr/ibmilogcpleoptistud
[8] http://www.gurobi.com/
[9] https://omnetpp.org/
[10] https://simpy.readthedocs.io/en/latest/
[11] https://www.grid5000.fr/

[10] A. Ahmed and E. Ahmed, "A survey on mobile edge computing," in *2016 10th International Conference on Intelligent Systems and Control (ISCO)*, Jan 2016, pp. 1–8.

[11] I. Stojmenovic, "Fog computing: A cloud to the ground support for smart things and machine-to-machine networks," in *2014 Australasian Telecommunication Networks and Applications Conference (ATNAC)*, Nov 2014, pp. 117–122.

[12] M. Chiang and T. Zhang, "Fog and IoT: An overview of research opportunities," *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 854–864, Dec 2016.

[13] S. Yi, C. Li, and Q. Li, "A survey of fog computing: Concepts, applications and issues," in *Proceedings of the 2015 Workshop on Mobile Big Data (Mobidata '15)*. New York, NY, USA: ACM, 2015, pp. 37–42.

[14] R. Mahmud and R. Buyya, "Fog computing: A taxonomy, survey and future directions," 2016. [Online]. Available: http://arxiv.org/abs/1611.05539

[15] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, "Mobile Edge Computing: Survey and Research Outlook," Jan. 2017. [Online]. Available: https://arxiv.org/abs/1701.01090

[16] A. Bhattacharya and P. De, "A survey of adaptation techniques in computation offloading," *Journal of Network and Computer Applications*, vol. 78, pp. 97 – 115, 2017.

[17] F. Rebecchi, M. D. de Amorim, V. Conan, A. Passarella, R. Bruno, and M. Conti, "Data offloading techniques in cellular networks: A survey," *IEEE Communications Surveys Tutorials*, vol. 17, no. 2, pp. 580–603, 2015.

[18] L. Chunlin and L. LaYuan, "Cost and energy aware service provisioning for mobile client in cloud computing environment," *The Journal of Supercomputing*, vol. 71, no. 4, pp. 1196–1223, 2015.

[19] S. Shekhar, A. Chhokra, A. Bhattacharjee, G. Aupy, and A. Gokhale, "INDICES: Exploiting edge resources for performance-aware cloud-hosted services," 2017. [Online]. Available: http://www.isis.vanderbilt.edu/sites/default/files/indices.pdf

[20] K. Toczé and S. Nadjm-Tehrani, "Resource management at the edge: A survey," To be submitted.

[21] M. Satyanarayanan, P. Bahl, R. Caceres, and N. Davies, "The case for VM-based cloudlets in mobile computing," *IEEE Pervasive Computing*, vol. 8, no. 4, pp. 14–23, Oct 2009.

[22] H. Frank, W. F. Fuhrmann, and B. V. Ghita, "Mobile edge computing: Requirements for powerful mobile near real-time applications," in *Proceedings of the Eleventh International Network Conference (INC 2016)*. Plymouth University, 2016, pp. 63–66.

[23] M. Aazam, M. St-Hilaire, C. H. Lung, and I. Lambadaris, "PRE-Fog: IoT trace based probabilistic resource estimation at fog," in *2016 13th IEEE Annual Consumer Communications Networking Conference (CCNC)*, Jan 2016, pp. 12–17.

[24] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: A comprehensive survey," *IEEE Communications Surveys Tutorials*, vol. 18, no. 3, pp. 1617–1655, 2016.

[25] F. Z. Yousaf and T. Taleb, "Fine-grained resource-aware virtual network function management for 5G carrier cloud," *IEEE Network*, vol. 30, no. 2, pp. 110–115, March 2016.

[26] K. Wang, K. Yang, X. Wang, and C. S. Magurawalage, "Cost-effective resource allocation in C-RAN with mobile cloud," in *2016 IEEE International Conference on Communications (ICC)*, May 2016, pp. 1–6.

[27] Y. Liu, M. J. Lee, and Y. Zheng, "Adaptive multi-resource allocation for cloudlet-based mobile cloud computing system," *IEEE Transactions on Mobile Computing*, vol. 15, no. 10, pp. 2398–2410, Oct 2016.

[28] B. Confais, A. Lèbre, and B. Parrein, "An object store service for a fog/edge computing infrastructure based on IPFS and scale-out NAS," in *1st IEEE International Conference on Fog and Edge Computing - ICFEC'2017*, Madrid, Spain, May 2017, <hal-01483702>.

[29] Q. Fan, N. Ansari, and X. Sun, "Energy driven avatar migration in green cloudlet networks," *IEEE Communications Letters*, 2017, preprint, https://doi.org/10.1109/LCOMM.2017.2684812.

[30] E. Vergara, S. Nadjm-Tehrani, and M. Asplund, "Fairness and incentive considerations in energy apportionment policies," *ACM Transactions on Modelling and Performance Evaluation of Computer Systems (TOMPECS) 2(1)*, 2016.

[31] L. Gu, D. Zeng, S. Guo, A. Barnawi, and Y. Xiang, "Cost efficient resource management in fog computing supported medical cyber-physical system," *IEEE Transactions on Emerging Topics in Computing*, vol. 5, no. 1, pp. 108–119, Jan 2017.

[32] W. Tärneberg, A. Mehta, E. Wadbro, J. Tordsson, J. Eker, M. Kihl, and E. Elmroth, "Dynamic application placement in the mobile cloud network," *Future Generation Computer Systems*, vol. 70, pp. 163 – 177, 2017.

[33] J. Plachy, Z. Becvar, and E. C. Strinati, "Dynamic resource allocation exploiting mobility prediction in mobile edge computing," in *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Sept 2016, pp. 1–6.

[34] M. Aazam and E. N. Huh, "Fog computing micro datacenter based dynamic resource estimation and pricing model for IoT," in *2015 IEEE 29th International Conference on Advanced Information Networking and Applications*, March 2015, pp. 687–694.

[35] J. Oueis, "Joint communication and computation resources allocation for cloud-empowered future wireless networks," Ph.D. dissertation, Université Grenoble Alpes, 2016, https://tel.archives-ouvertes.fr/tel-01366449.

[36] P. Borylo, A. Lason, J. Rzasa, A. Szymanski, and A. Jajszczyk, "Energy-aware fog and cloud interplay supported by wide area software defined networking," in *2016 IEEE International Conference on Communications (ICC)*, May 2016, pp. 1–7.

[37] H. R. Arkian, A. Diyanat, and A. Pourkhalili, "MIST: Fog-based data analytics scheme with cost-efficient resource provisioning for iot crowdsensing applications," *Journal of Network and Computer Applications*, vol. 82, pp. 152 – 165, 2017.

[38] L. Tang, S. He, and Q. Li, "Double-sided bidding mechanism for resource sharing in mobile cloud," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 2, pp. 1798–1809, Feb 2017.

[39] C. T. Do, N. H. Tran, C. Pham, M. G. R. Alam, J. H. Son, and C. S. Hong, "A proximal algorithm for joint resource allocation and minimizing carbon footprint in geo-distributed fog computing," in *2015 International Conference on Information Networking (ICOIN)*, Jan 2015, pp. 324–329.

[40] M. Satyanarayanan, P. Simoens, Y. Xiao, P. Pillai, Z. Chen, K. Ha, W. Hu, and B. Amos, "Edge analytics in the internet of things," *IEEE Pervasive Computing*, vol. 14, no. 2, pp. 24–31, Apr 2015.

[41] B. Zhou, A. V. Dastjerdi, R. N. Calheiros, S. N. Srirama, and R. Buyya, "A context sensitive offloading scheme for mobile cloud computing service," in *2015 IEEE 8th International Conference on Cloud Computing*, June 2015, pp. 869–876.

[42] R. Mijumbi, S. Hasija, S. Davy, A. Davy, B. Jennings, and R. Boutaba, "Topology-aware prediction of virtual network function resource requirements," *IEEE Transactions on Network and Service Management*, vol. 14, no. 1, pp. 106–120, March 2017.

[43] D. Zhang, Z. Chang, F. R. Yu, X. Chen, and T. Hämäläinen, "A double auction mechanism for virtual resource allocation in SDN-based cellular network," in *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Sept 2016, pp. 1–6.

[44] NSF Edge Computing Workshop Committee, "NSF workshop report on grand challenges in edge computing," 2016. [Online]. Available: http://iot.eng.wayne.edu/edge/NSF%20Edge%20Workshop%20Report.pdf

[45] R. N. Calheiros, R. Ranjan, A. Beloglazov, C. A. F. De Rose, and R. Buyya, "Cloudsim: A toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms," *Softw. Pract. Exper.*, vol. 41, no. 1, pp. 23–50, Jan. 2011.

[46] H. Gupta, A. V. Dastjerdi, S. K. Gosh, and R. Buyya. (2016, June) iFogSim: A toolkit for modeling and simulation of resource management techniques in internet of things, edge and fog computing environments. [Online]. Available: https://arxiv.org/abs/1606.02007v1

[47] H. Gedawy, S. Tariq, A. Mtibaa, and K. Harras, "Cumulus: A distributed and flexible computing testbed for edge cloud computational offloading," in *2016 Cloudification of the Internet of Things (CIoT)*, Nov 2016, pp. 1–6.

[48] Y. Liu and M. J. Lee, "Security-aware resource allocation for mobile cloud computing systems," in *2015 24th International Conference on Computer Communication and Networks (ICCCN)*, Aug 2015, pp. 1–8.

[49] H. Liang, D. Huang, L. X. Cai, X. Shen, and D. Peng, "Resource allocation for security services in mobile cloud computing," in *2011 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, April 2011, pp. 191–195.