Intelligent Mobile Edge Computing: A Deep Learning Based Approach

Abhirup Khanna(✉), Anushree Sah, and Tanupriya Choudhury

University of Petroleum and Energy Studies, Dehradun, India
{akhanna,asah,tanupriya}@ddn.upes.ac.in

Abstract. In recent times researchers across the globe have shown keen interest towards advancements in the domain of edge computing. Mobile Edge Computing (MEC) is a new age computing paradigm wherein cloud services are made accessible at network edges via the use of mobile base stations. It is a promising technology that helps in overcoming the limitations of mobile cloud computing. MEC facilitates seamless integration of various application services, thereby proving cloud resources at the edge of the network, within the vicinity of the end-user's locality. It can effortlessly be integrated with the upcoming 5G architecture, hence supporting the execution of resource-rich applications that require low network latency. In order to enhance the levels of intelligence at mobile base stations, deep learning algorithms can be implemented over network edges for rendering optimized communication and workload balancing. The paper discusses a conceptual architecture for creating a mobile edge computing environment involving the applicability of deep learning algorithms. The paper discusses the fundamentals of MEC along with specific applications of reinforcement and continuous learning in an edge environment. We list the benefits of MEC along with a discussion on how its amalgamation with deep learning models can prove beneficial in case of a computation offloading scenario.

Keywords: Edge computing · Mobile Computation Offloading · Deep learning · Mobile computing

1 Introduction

Predominance of mobile terminals, like smartphones or tablet, computers and so on has a lot of effect on mobile and wireless networks, which in a way creates problems for the mobile system all over the world. Cellular networks experiences high power consumption, less storage capacity, high latency and lower bandwidth. Apart from this, it is believed that the exponential growth of new technologies, i.e., the Internet of Things (IoT), will lead to another convergence of the cellular and wireless networks. Mobile Cloud Computing (MCC), which is the amalgamation of the cloud computing in the mobile environments, provides significant opportunities for mobile devices, which enables them to store, compute, and power through the use of centralized cloud resources. However, while opening a multitude of mobile devices, MCC faces significant

M. Singh et al. (Eds.): ICACDS 2020, CCIS 1244, pp. 107–116, 2020.
https://doi.org/10.1007/978-981-15-6634-9_11

problems [5], like high latency, low coverage, security vulnerabilities, and data transmission latency, which in a way can be cumbersome, especially for the next generation mobile networks (for example, 5G).

It has been seen that MCC is less applied to Quality of service and real life applications. As per a report from Cisco Visual networking index the use of mobile devices is increasing drastically and after 2020 it will further gain more and more usage [2]. The increase of mobile usage is because of increase mobile users and development and availability of mobile applications. In the era of Information technology, EDGE computing has become the leading mobile cellular computing (MEC) in cellular networks.

The term mobile edge computing was first introduced in 2013 when Nokia Siemens Networks and IBM developed the MEC platform, which allows applications to work directly. This platform only speeds up the local area, which does not support the application migration, compatibility. MEC is also recognized as the leading new technology for 5G networks by the European 5G PPP (Public-Private Partnership 5G Infrastructure). The Natural Language Processing helps the Edge Computing to be more robust and increase the rate of data processing so that the output can be generated at much faster rate. It's a type of program which delivers low latency. According to Karim Arabi, edge computing is a broadly network whatever is happening outside the cloud and in the applications where real time processing of data is required.

The world is seeing a constant rise in the numeral of smart cities. Edge computing looks like a viable solution for making the smart city environments as it facilitates the extension of the cloud resources to the network edge [20]. Henceforth, enhancing the service awareness, scalability and low latency. Characteristics high context awareness, single hop connectivity and geo distribution are the reasons for rapid adoption of mobile edge computing. The amalgamation of edge computing along with deep learning mechanisms can be used to create real life health applications. A similar application being, HealthFog, a heart disease analysis application [6]. It delivers healthcare as a service using IoT devices over a fog network.

Organization of rest of the paper is done as follows: Sect. 2 explains some of the contemporaries of edge computing. Section 3 discusses the applicability of deep learning algorithms in an edge environment. It talks through some of the most popular verticals wherein extensive research is happening at a global scale. Section 4 which suggests some of the benefits of mobile edge computing. Finally we have, Sect. 5 presenting the experimental analysis of MEC and deep learning models.

2 Related Technologies

Cloudlet is a small data center that usually happens as a wireless transition from public places such as mobile devices for the convenience of the hospital, mall, office building, etc. It is a convenient approach wherein many blocks of multi-core computers create a cloud that connects remotely to a cloud server [1]. Cloudlet is presented as a promising solution for remote extended area networks (WAN) and latency in cellular Energy consumption due to the use of mobile communication for data transmission. The primary goal of the cloud server cloud cover is to carry the cloud. The end user's technology which provides resource support more time-sensitive applications. Further studies confirmed that using resource wealthy machines close to cellular customers, called cloudlets, provides offerings typically observed inside the cloud, furnished improvements in execution time while some of the responsibilities are offloaded to the edge node. On the alternative hand, offloading each task may also bring about a slowdown due to switch instances between device and nodes, so depending at the workload a foremost configuration may be defined [18].

Fog computing [17] which is also known as edge computing, which supports universally connected equipment. The word fog computing was made CISCO system that brings cloud services to enterprise point of view Network like MEC [6]. In fog computing, processing is mostly done at the end of the LAN on the IoT gateway or fog node. The advantage of allowing only processing equipment in fog computing is the benefit is to collect data from various sensors and act in accordance to it [19]. However, due to their dependence on the wireless connection, there are some limitations in ambiguous calculations, which must be active in order to perform the complicated actions. The estimate of fog and MECs are widely used interchangeably except in some cases where they differ [21, 22].

Blockchain and the use of smart contracts are the best known examples for creating distributed applications. The expression "smart contract" was first authored in mid-1990s by researcher and cryptographer Szabo, who characterized a smart contract as a lot of guarantees, indicated in computerized structure, including rules inside which the different groups perform on these guarantees. In his popular model, Szabo analogized smart agreements to vending machines: machines take in coins, and by means of a simple algorithmic system (e.g., finite automata), dispenses change and item as indicated by the showed cost. Smart contracts go past the candy machine by proposing to install contracts in a wide range of properties by advanced methods. As a rule, smart agreements or contracts can be characterized as the digital rules that digitally encourage, confirm, and uphold the agreements made between at least two gatherings on blockchain. As smart contracts are normally used and verified by blockchain, they have some novel attributes. In the first place, the program code of a smart contract will be recorded and confirmed on blockchain, subsequently making the agreement alter safe. Secondly, the execution of a smart contract is implemented among different, trustless individual hubs without incorporated control, and coordination of third part specialists [20, 23].

3 Deep Learning Applications in Edge

Fog or edge computing is a kind of program which delivers low latency. The following figure illustrates an edge environment. In both IoT and Mobile Computation Offloading tasks and data were communicated to the remote cloud for performing computational operations [4]. Whereas, through edge computing code offloading and data transfer is performed at the network edge level, thus helping in quick response time and enhanced Quality of Service. Further in this section we talk about some of the fields wherein the deep learning algorithms [10] can be implemented in an edge environmental setup. Similar to these areas, a use case for mobile code offloading is discussed in Sect. 5 (Fig. 1).

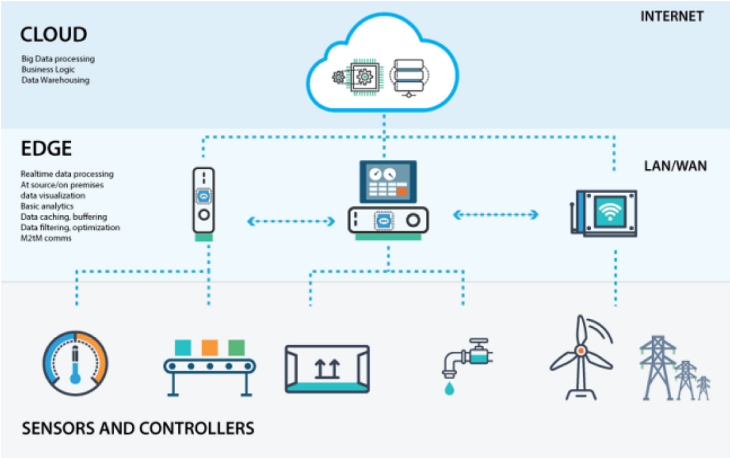

Fig. 1. Conceptual architecture for edge computing [16]

3.1 Internet of Things (IoT)

In today's world of IoT we need lots of storage and massive amount of data to data centers, due to which bandwidth is affected. Besides all these improvements in to-day's world, data centers still cannot guarantee the transfer rate and response times which is very important. Edge software services lessen the volumes of information that must be moved, the resultant site visitors, and the space that records need to move. It provides decrease latency and decreases transmission prices.

3.2 Micro Data Centers and Cloudlets

Computation offloading for actual-time programs, consisting of facial reputation algorithms, confirmed widespread enhancements in reaction times as established in early studies. Further studies confirmed that using resource wealthy machines close to cellular customers, called cloudlets, and providing offerings typically observed inside the cloud, furnishes improvements in the execution time while some of the responsibilities are offloaded to the edge node. On the alternative hand, offloading each task may also bring about a slowdown due to switch instances between device and nodes, so depending at the workload and foremost configuration may be defined.

3.3 Virtual Reality

Another use of the architecture is cloud gaming, in which some components of a sport may want to run within the cloud, even as the rendered video is transferred to lightweight customers together with mobile, VR glasses, and many others. Such sort of streaming is also referred to as pixel streaming. Conventional cloud video games might also suffer from excessive latency and inadequate bandwidth, due to the fact the amount of records transferred is huge because of the excessive resolutions required through some offerings.

In such cases Edge computing plays a vital role. An-other use of edge computing architecture is cloud gaming, in which some components of a sport may want to run within the cloud [8], even as the rendered video is transferred to lightweight customers together with mobile, VR glasses, and many others. Such sort of streaming is also referred to as pixel streaming. Conventional cloud video games might also suffer from excessive latency and inadequate bandwidth, due to the fact the amount of records transferred is big because of the excessive resolutions required through some offerings.

3.4 NLP

Natural Language Processing helps the Edge Computing to be more robust and increase the rate of data processing so that the output can be generated at much faster rate. Natural Language Processing, is an intelligent technology that helps to create an interaction between the machine and the humans. NLP helps to fill the gap between the machine and human by processing various codes, computation linguistics and also computer science to manipulate the human language and help machine to be more precise with the output. NLP helps to deliver the cognitive solutions by improving the output result. NLP helps to improve the latency by processing the data of the idle time a person spends over the service without sending any data packets. This helps to improve the route of the data transfer. NLP helps in real-time analytics to make the CPU workload more intensive. The secret is in knowing your field of interest and market, and by analyzing your activity and making a suitable design to support your system. The key here is develop an 'intimacy' with the customer through new technologies.

3.5 Computer Vision

An example of applicability of deep learning in an edge environment can be, a facial recognition system that verifies the face of the entering and exiting facilities has been used by Department of Defense. This technique requires machine learning and natural language processing, neural networks and other statistical computing for the process. This technology in particular has storage requirement to make the setting feasible. For this, edge computing provides crucial storage required for the processing of data.

In relation to the five research areas discussed previously, the following figure showcases a comparison between works having implemented deep learning models in an edge setup (Table 1).

Table 1. Selected works on union of edge computing and deep learning

Work	Deep learning model	Application	Performance parameters
DeepIoT [9]	VGGNet	Image recognition	Latency, memory
DeepMon [11]	Yolo [12]	Object detection	Latency
VideoEdge [13]	AlexNet	Image classification	Accuracy
DeepThings [14]	Yolo [12]	Object detection	Accuracy, latency
MCDNN [15]	AlexNet	Image classification	Energy, memory

4 Advantages of Mobile Edge Computing

MEC focuses on essential metrics like the concept delay. Moreover, high bandwidth, which is achieved by limiting data movement. Then for MEC servers, and then for a centralized server with a long delay cost. Apart from this, electricity consumption is also a significant problem. Computational work related to external resource systems lead to increased battery life of user devices.

- Speed: Edge computing helps to respond to data almost instantaneously by eliminating the lags. The internet bandwidth is reduced by providing the data processing computing near to the source. Since the resources are remotely available, the efficiency is increased and cost is reduced. The data is not required to be put up on cloud, and the security is ensured for the sensitive data. The most vital advantage of edge computing is their ability to recover the network's performance by combating latency.
- Security: The edge computing architecture's distributed nature makes it easy to execute the security protocol, which can separate the hacked parts without turning the whole network off. For example, a facial recognition system that verifies the face of the entering and exiting facilities has been used by Department of Defense. This technique requires machine learning and natural language processing, neural networks and other statistical computing for the process. This technology in particular has storage requirement to make the setting feasible [7].
- Scalability: Border computing provides a much less expensive way for scalability. The utility of edge computing devices with the ability to process also reduces the cost of development, because each added new device does not apply the critical requirements on the network core bandwidth. At this point, edge computing comes in handy as reduced form factor is critical for IoT implementation and edge computing provides crucial storage required for the processing of data.
- Reliability: With the security advantages of edge computing, it is not surprising that it also provide high reliability. Since IoT edge computing devices and edge data centers are located close to end users, the probability of problems with the net-work in remote locations affecting local customers is less likely. IoT edge computing devices will work fine even if there is some fault with the nearby data centers if the edge computing architecture is followed. When edge computing architecture is used then it is seen that complete service will never disable completely. The key here is develop an 'intimacy' with the customer through new technologies [3].

5 Implementation and Simulation

In this section we would be discussing an application for mobile computation offloading wherein resource intensive tasks of an application are offloaded in order to ensure reduced battery consumption and response time. The sample application comprises of three different categories of tasks that are I/O intensive, CPU intensive and Data intensive in nature. According to the proposed hypothesis "Mobile Computation Offloading performed in an Edge environment making use of a deep learning algorithm is most appropriate in terms of ensuring response time and mobile battery consumption".

The following are the three set of experiments that have been performed for validating the proposed hypothesis.

1. Task Offloading in an MCC Environment
2. Task Offloading in an Edge Environment
3. Task Offloading in an Edge Environment assisted by Deep Learning

Convolution Neural Network (CNN) is the deep learning model used in this experiment. The CNN is trained on the basis of task execution patterns. It identifies relationships between tasks in terms of a task being independent or dependent on other task executions.

Table 2. Task offloading

Experiment number	Number of tasks offloaded	Battery consumption (%)	Latency (ms)
1	9	12.64	32
2	5	12.22	21
3	5	11.87	16

The above-mentioned Table 2 has been further illustrated using the following graphs in form of Fig. 2 and Fig. 3

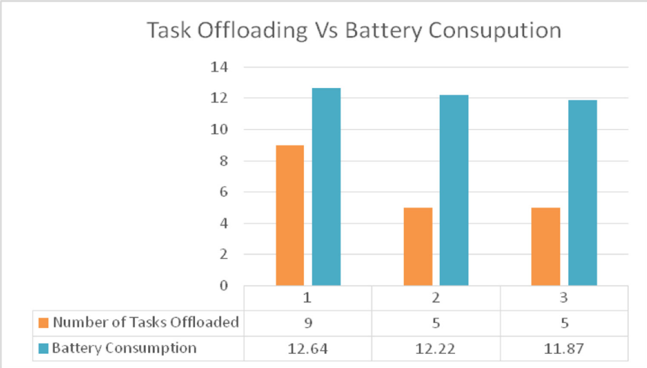

Fig. 2. Effect of task offloading on battery consumption

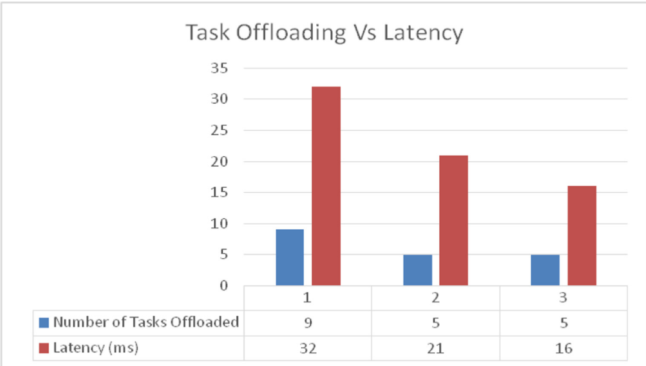

Fig. 3. Effect of task offloading on latency

6 Conclusion

Mobile edge computing (MEC) is one of the new age computing paradigm and is seeing rapid adoption among researchers across the globe. Characteristics such as enhanced scalability and performance are the reasons for it being a successful successor to Mobile Cloud Computing. Significant research is being conducted towards ensuring amalgamation of deep learning models and the edge environment. In recent times, areas like natural language processing and computer vision have seen works involving the use of deep learning algorithms and edge computing. Furthermore, the amalgamation can lead to enhanced performance of mobile computation offloading with respect to performance parameters of mobile battery consumption and latency.

References

1. Xu, X., Chen, Y., Yuan, Y., Huang, T., Zhang, X., Qi, L.: Blockchain-based cloudlet management for multimedia workflow in mobile cloud computing. Multimed. Tools Appl. **79**(15), 9819–9844 (2019). https://doi.org/10.1007/s11042-019-07900-x
2. http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-indexvni/mobile-white-paper-c11-520862.html
3. Wang, S., et al.: Edge server placement in mobile edge computing. J. Parallel Distrib. Comput. **127**, 160–168 (2019)
4. Khanna, A. et al.: Adaptive mobile computation offloading for data stream applications. In: 2017 ICACCA (Fall), pp. 1–6. IEEE, September 2017
5. Khanna, A. et al.: Mobile cloud computing architecture for computation offloading. In: 2016 NGCT, pp. 639–643. IEEE, October 2016
6. Tuli, S., Basumatary, N., et al.: HealthFog: an ensemble deep learning based smart healthcare system for automatic diagnosis of heart diseases in integrated IoT and fog computing environments. Future Gener. Comput. Syst. **104**, 187–200 (2020)
7. Amanullah, M., et al.: Deep learning and big data technologies for IoT security. Comput. Commun. **151**, 495–517 (2020)
8. Tomar, R., Khanna, A., Bansal, A., Fore, V.: An architectural view towards autonomic cloud computing. In: Satapathy, S.C., Bhateja, V., Raju, K.Srujan, Janakiramaiah, B. (eds.) Data Engineering and Intelligent Computing. AISC, vol. 542, pp. 573–582. Springer, Singapore (2018). https://doi.org/10.1007/978-981-10-3223-3_55
9. Yao, S. et al.: DeepIot: compressing deep neural network structures for sensing systems with a compressor-critic framework. In: Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems, p. 4. ACM, November 2017
10. Bahri, Y., Kadmon, J., Pennington, J., Schoenholz, S.S., Sohl-Dickstein, J., Ganguli, S.: Statistical mechanics of deep learning. Annu. Rev. Condens. Matter Phys. **11**, 501–528 (2020)
11. Huynh, L.N. et al.: DeepMon: mobile GPU-based deep learning framework for continuous vision applications. In: Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services, pp. 82–95. ACM, June 2017
12. Redmon, J., Farhadi, A.: YOLO9000: better, faster, stronger. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7263–7271 (2017)
13. Hung, C.C. et al.: VideoEdge: processing camera streams using hierarchical clusters. In: 2018 IEEE/ACM Symposium on Edge Computing (SEC), pp. 115–131. IEEE, October 2018
14. Zhao, Z., et al.: DeepThings: distributed adaptive deep learning inference on resource-constrained IoT edge clusters. IEEE Trans. Comput. Aided Des. Integr. Circuits Syst. **37**(11), 2348–2359 (2018)
15. Han, S. et al.: MCDNN: an approximation-based execution framework for deep stream processing under resource constraints. In: Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services, pp. 123–136. ACM, June 2016
16. IIoT Edge Computing vs. Cloud Computing. https://openautomationsoftware.com/blog/iiot-edge-computing-vs-cloud-computing. Accessed Dec 2018
17. Yang, Y., Luo, X., Chu, X., Zhou, M.T.: Fog computing architecture and technologies. In: Fog-Enabled Intelligent IoT Systems, pp. 39–60. Springer, Cham (2020)
18. Fore, V., Khanna, A., Tomar, R., Mishra, A., Intelligent supply chain management system. In: 2016 International Conference on Advances in Computing and Communication Engineering (ICACCE), pp. 296–302. IEEE, November 2016
19. Calheiros, R.N.: Fog and edge computing: challenges and emerging trends (invited talk). In: 2nd Workshop on Fog Computing and the IoT (Fog-IoT 2020). Schloss Dagstuhl-Leibniz-Zentrum für Informatik (2020)

20. Khan, L.U., Yaqoob, I. et al.: Edge computing enabled smart cities: a comprehensive survey. IEEE Internet Things J., 1 (2020)
21. Kumar, P., Choudhury, T. et al.: Fog computing: common security issues and proposed countermeasures. In: System Modeling Advancement in Research Trends (SMART) (2016)
22. Garg, V., Choudhury, T., et al.: Advance survey of mobile ad-hoc network. IJCST **2**(4), 552–555 (2011)
23. Choudhary, T., Choudhury, V., et al.: An approach to improve task scheduling in a decentralized cloud computing environment. Int. J. Comput. Technol. Appl. **3**(1), 312–316 (2012)