Document downloaded from:

http://hdl.handle.net/10251/189357

This paper must be cited as:

Mukherjee, M.; Guo, M.; Lloret, J.; Zhang, Q. (2020). Leveraging Intelligent Computation Offloading with Fog/Edge Computing for Tactile Internet: Advantages and Limitations. IEEE Network. 34(5):322-329. https://doi.org/10.1109/MNET.001.2000004

The final publication is available at

https://doi.org/10.1109/MNET.001.2000004

Copyright Institute of Electrical and Electronics Engineers

Additional Information

Leveraging Intelligent Computation Offloading with Fog/Edge Computing for Tactile Internet: Advantages and Limitations

Mithun Mukherjee, *Senior Member, IEEE*, Mian Guo, *Member, IEEE,* Jaime Lloret, *Senior Member, IEEE*, and Qi Zhang, *Member, IEEE*

Abstract: With recent advancement in wireless communication and networks, we are at the doorstep of the Tactile Internet. Tactile Internet aims to enable the skills delivery and thereafter democratize the specialized skills for many emerging applications (e.g., remote medical, industrial machinery, remote robotics, autonomous driving). In this article, we start with the motivation of applying intelligent edge computing for computation offloading in Tactile Internet. Afterward, we outline the main research challenges to leverage edge intelligence at the master, network, and controlled domain of the Tactile Internet. The key research challenges in Tactile Internet lie in its stringent requirements such as ultra-low latency, ultra-high reliability, and almost zero service outage. We also discuss major entities in intelligent edge computing and their role in the Tactile Internet. Finally, several potential research challenges in edge intelligence for Tactile Internet are highlighted.

I. Introduction

The seamless and ubiquitous connectivity among the wireless devices and the hardware advancement in haptic devices are taking a step forward to the real-time steering and control of the remote physical or virtual objects. While the 5G technologies have started to be rolling out around the world, the Tactile Internet [1], coined by Fettweis in 2014, is taking to another paradigm shift towards real-time steering and control of the remote physical or virtual objects. By enabling the immersive bi-directional human-machine (and/or human-physical) or human-virtual world interaction, Tactile Internet is envisioned to democratize the specialized skills (often termed as Internet of Skill [2]) in numerous fields ranging from healthcare, remote robotics, remote training, education, to autonomous vehicle. Several working groups, such as IEEE P1918.1 Tactile Internet working group [3], have already started to standardize architectural aspects with the key elements, the interfaces, the functional capabilities of the Tactile Internet.

II. How is the Tactile Internet Different from Traditional Content-centric Internet?

A Tactile Internet generally consists of three main domains: master domain, network domain, and controlled domain. An individual human (or machine) or a group of machines can act as controller in the master domain. In case of human as the controller, a haptic device, often called as human-to-system interface, converts the human input to the haptic signal and also provides the haptic feedback from the controlled domain to the human. The control signals are transmitted over the network domain to the controlled domain. As illustrated in Figure 1, the network domain generally consists of routers, switches, base station, network controller, and even edge/cloud servers. The tele-operated robots, actuators in the controlled domain perform the actions based on the received command/control signals from the master domain. Ultimately, the sensors in the controlled domain gather the state dynamics and provide feedback to the controller in the master domain.

M. Mukherjee is with the School of Artificial Intelligence, Nanjing University of Information Science and Technology, Nanjing 210044, China (e-mail: m.mukherjee@ieee.org).
M. Guo is with the School of Electronic and Information Engineering, Guangdong University of Petrochemical Technology, Maoming 525000, China (e-mail: mian.guo123@gmail.com).
J. Lloret is with the Instituto de Investigación para la Gestión Integrada de Zonas Costeras (IGIC), Universitat Politècnica de València, 46022 Valencia, Spain (e-mail: jlloret@dcom.upv.es).
Q. Zhang is with the DIGIT, Department of Engineering, Aarhus University, 8000 Aarhus, Denmark (e-mail: qz@eng.au.dk).

A. A Broader Concept than uRLLC

A global control loop is established in Tactile Internet with a bilateral communication of the haptic signal, which is one of the fundamental differences with the traditional content-centric Internet. Although in some applications, such as remote cyber-physical systems, an exchange of real-time control/steering and sensing/actuation is carried out, this bilateral communication is constantly present in Tactile Internet. There is a growing misconception that the ultra-reliable and low-latency communication (uRLLC) [4] and Tactile Internet are similar [5]. Indeed, the uRLLC can be considered as the underlying communication infrastructure which plays an important role in Tactile Internet, Tactile Internet is broader than uRLLC. Note that uRLLC mostly focuses on the lower layers in the protocol stack, where the performance metric in terms of bit error rate are used to evaluate the reliability of uRLLC. However, Tactile Internet is more interested to evaluate the delay and reliability from upper layer (e.g., application layer) perspective. For example, the reliability can be evaluated by the accuracy of control task, i.e., quality of tasks (QoT) [15]. Basically, QoT is directly related to the delivery of the specialized skill that is envisioned in Internet of Skill [2]. Moreover, quality-of-experience (QoE) in Tactile Internet measures the difference between the experience with physical interaction and the remote operation. QoE largely depends on the latency and reliability in the underlying uRLLC. This performance metric arises mainly due to the human-in-the-loop concept in Tactile Internet. Note that uRLLC alone does not guarantee to fulfill QoE and QoT in Tactile Internet.

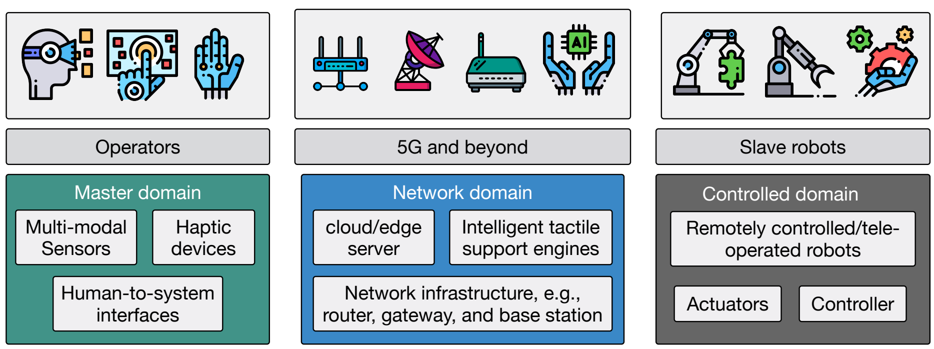

Figure 1: The main domains in Tactile Internet.

B. Functional capabilities:

Latency: The carrier-grade end-to-end latency (i.e., close to 2ms) is generally required in motion control for the industrial applications [6], however, mostly realized through wired connection. Note that the Tactile Internet imposes an ultra-low end-to-end delay between master and controlled domains with an order of 1 ms (see Figure 2). Again, due to the mission-critical nature of the Tactile Internet the hard deadline is required in Tactile Internet. Latency and reliability go hand in hand. Non-mission-critical applications (e.g., remote gaming) do not have high reliability requirement as mission critical ones, therefore, it is possible to relax slightly the hard deadline. That is to say, for stability and transparency reason, the haptic signal and the control signal require hard-deadline-aware packet transmission. The delayed packet often becomes useless for the current time frame for the feedback control system.

Reliability: In content-centric traditional Internet, a few packet loss is acceptable in some particular service provisioning (e.g., the acceptable packet loss $\leqslant$ 0.5% - 5% in video data transmission). However, the envisioned reliability in Tactile Internet is 10^-5. To explain, one out of 10^5 data packets is tolerated to be delivered beyond its predefined delay deadline, otherwise the Quality-of-Service (QoS) will be violated in the Tactile Internet. Basically, this degradation of reliability directly impacts the QoE and QoT in Tactile Internet.

Tactile service availability: Tactile service provider must ensure that there would be no interruption or out of service coverage (literature suggest that tactile service must ensure at most one millisecond outage in a day [3]) over the entire period of service provisioning. It is true that it becomes very hard to satisfy always-on tactile service coupled with ultra-high reliability and ultra-low end-to-end latency [3] [7]. To this end, tactile support engines are envisioned to guarantee tactile service availability using model-mediated teleoperation method.

Bandwidth: To construct the model of the remote environment to calculate the estimated haptic feedback for the controller in the master domain, apart from the haptic and kinesthetic data, high definition audio and video are also equally important. Thus, bandwidth allocation in Tactile Internet should consider the heterogenous packet loss rate, data rate, latency, sampling frequency requirements of these kinds of QoS requirements in Tactile Internet applications. For example, 3D camera may require 137 Mbps-1.6 Gbps with a latency of around 150 ms, whereas a typical haptic feedback such as vibration demands less than 5 ms latency with 400 Kbps [4].

Scalability: As traffic pattern of multiplexed multimodal signal (i.e., haptic, audio, and visual) is very different from that of IoT applications, the overbooking factor for Tactile Internet applications is expected to be lower. In addition, to fulfill the stringent QoS requirements of Tactile Internet, it might be necessary to reserve dedicated computation and communication resources during TI services. Therefore, scalability is not a trivial issue.

To support both mission-critical and non-mission-critical nature of the Tactile Internet applications, Tactile Internet working group came forward to categorize the applications into ultra-grade and normal-grade QoS requirements [3]. A more detail can be found in the recent release of the Tactile Internet reference architecture. Basically, the scalability (i.e., the number of connected haptic/tactile devices) is slightly relaxed with more emphasis on the latency, reliability, and availability in carrier-grade compared to the normal-grade. When the number of connected haptic/tactile devices are increased, the other requirements (such as latency, reliability, and availability of tactile service) must be slashed in the functional capability, termed as normal-grade applications of the Tactile Internet.

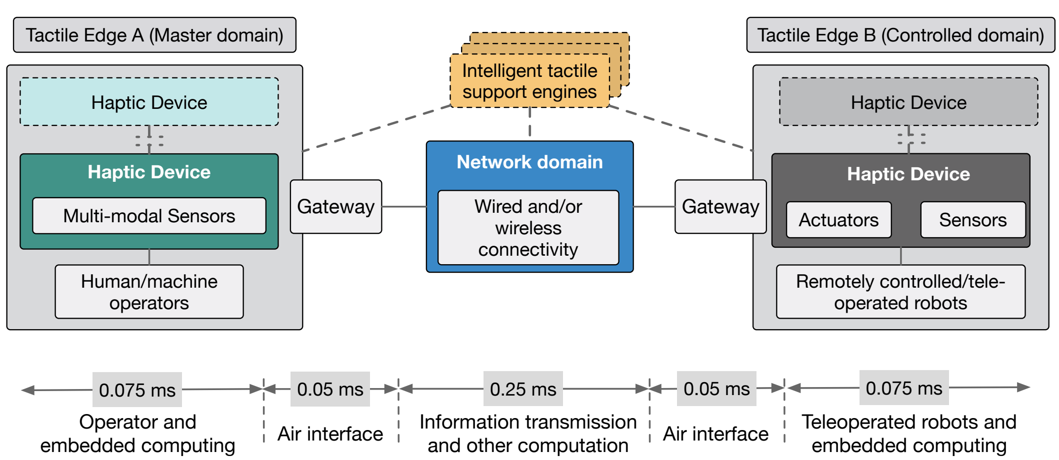

Figure 2: The envisioned end-to-end latency in Tactile Internet

III. COMPUTATION OFFLOADING IN TACTILE INTERNET: MOTIVATION

Fundamental limitation and uncertainty of communication technologies. Consider a fully optical fiber link between the master domain and the controlled domain. Now, even ignoring the delays in data collection and data processing, the maximum distance between master domain and controlled domain can reach only 150 Km within 1ms round-trip time. Thus, we must leverage disruptive technologies and/or framework to fulfil the technical requirements (such as, carrier-grade latency, reliability, and availability with ultra-low latency tele-operation) of Tactile Internet.

A. Huge computation to stabilize global control loop, to construct model of the remote environment at the master domain model, and to command interpretation in controlled domain

Similar to the principles of the tele-operation technique, it is quite evident that the controlled domain does not have prior knowledge on control signal received from the master domain. At the same time, the master domain also needs to be familiar with the behavior of the remotely controlled environment. Basically, the master domain constructs a model of the remote environment based on the haptic and tactile feedback received from the controlled domain. Therefore, to establish a stabilized global control-loop, a huge computation is required due to higher level of data abstraction and large number of force-feedback interactions. Moreover, multiple devices including haptic/tactile and network devices forming an edge network can collect multi-sensory data from to build the prediction control model and refine/update the model afterward.

B. Resource and power constraint tactile/haptic devices in master and controlled domains

Where to process the computation-intensive tasks, how fast the edge servers near to the master domain constructs the model of the remote environment in a precise manner, and how quickly the master domain provides the action based on the updates from the controlled domain are the challenging issues due to following reasons: a) always remote cloud connectivity results significant amount of transmission delay and b) computing devices in the master domain have limited computing and storage resources. Note the above limitations also exist for the tactile/haptic devices in the controlled domain to interpret the command signal from the master domain.

IV. STATE-OF-THE-ART OFFLOADING IN FOG-EDGE COMPUTING

The CPU processing frequency, on-device battery capacity, communication, and storage resources at the network edge are constantly increasing. This motivates to utilize computing, storage, and communication resources of the edge devices to complete the independent tasks by offloading to the suitable edge devices. Recently, a non-linear task offloading strategy has been proposed in [8] considering CPU cycle of the real-time tasks while pooling and sharing the computational resources among the wireless devices. The global knowledge of the computational resources shared by the all participating devices would likely give the best performance for the offloading. However, these resources are often distributed over the network, where a decentralized algorithm is highly desirable that achieves the performance close to the policy based on the global knowledge on the network state. To this end, a game theoretic approach was also considered in [9] to find the optimal task offloading policy to minimize the computation time based on the variational inequality theory. The game model relies on the average system performance with low signaling overhead. The metric that reflects the collaboration among the devices was also studied in [9]. Recently, a deadline-aware fog node collaboration was studied in [10] to meet the diverse deadline requirement of the offloaded tasks. These works laid a strong foundation on the study in dynamic nature of the fog-edge environment to form the network and to balance the workload among the participatory devices, however, the relation with the Tactile Internet that has a stringent requirement is totally missing. Nevertheless, the online fog-edge formation and resource allocation among the network devices would definitely play an important role in the realization of the Tactile Internet.

A. Moving towards Intelligent Edge computing: A relook

We are witnessing the emergence of the Artificial Intelligence (AI), particularly, machine learning techniques to optimize resource allocation and improved service provisioning in wireless communications. Recently, Q-learning, reinforcement learning, and deep reinforcement learning (DRL) are being investigated in the edge computing systems [11]. Unlike other machine learning category, reinforcement learning is being investigated as a potential candidate due to its intrinsic nature of learning from its own previous action-states, thus it becomes less data hungry as in deep learning that requires a high-dimensional raw data for training. The main challenge in applying reinforcement learning (RL) in edge computing is that how to train the agent that gives the optimal policy for a given state from the environment.

Before we are going to discuss the suitability and motivation to apply edge intelligence at the Tactile Internet, we present how we can leverage delay-aware RL at the edge computing scenario. Consider an edge computing system with three types of end-users (e.g., (30,10,5) users, respectively), which tolerate delays of (4,6,10) ms, respectively, and an edge server. The end-users generate task computation requests with probabilities ρ_1, ρ_2, and ρ_3, respectively. To explain, the task request probability is defined as the probability of an end-user generates a computation task in a slot, e.g., $\rho_1 = 0.5$, indicates that, in every simulated slot, a type-1 mobile user generates a computation task with probability of 0.5. The requests can either be processed locally in the mobile device, or computation offloaded to the edge server. We define *delay-guarantee-ratio* as the ratio of the number of tasks completed within their delay deadlines to the number of total arrived tasks. As shown in Figure 3(a), the efficiency of the delay-aware RL scheme is demonstrated by provisioning the highest *delay-guarantee-ratio* under various task request probabilities by comparing with the delay-aware max-weight and greedy schemes. Moreover, Figure 3(b) illustrates that the *delay-guarantee-ratio* under various computation capabilities of local computing nodes with respective to the edge server in delay-aware RL scheme outperforms delay-aware *maxweight* and greedy schemes.

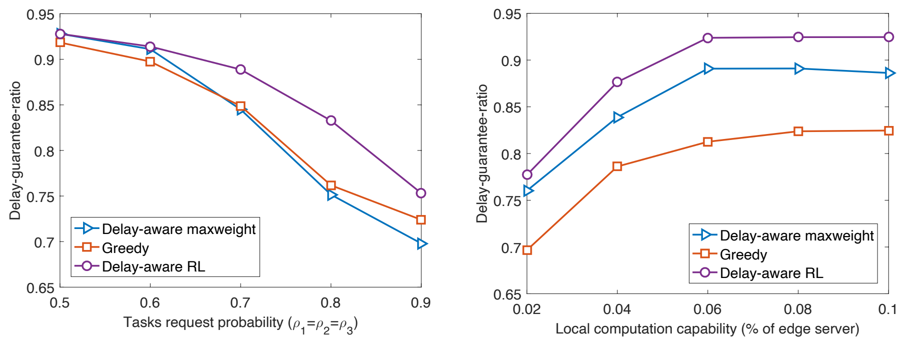

Figure 3: Performance of the *delay-guarantee-ratio* (a) under various task request probabilities and (b) under various computation capabilities.

V. Edge Intelligence in Tactile Internet: Motivation and Objectives

In the following, we discuss the motivation for leveraging the edge intelligence, particularly in the Tactile Internet.

A. Intelligent edge computing can predict the control command or haptic feedback. Based on the above discussion, we can conclude that apart from other research challenges (such as, hardware for haptic devices), the transmission delay will impose the fundamental challenges for the realization of the Tactile Internet. One of the possible solutions is the artificial intelligence-enabled tactile support engines (as shown in Figure 4) that can reside at the master/controlled domain, or at the network domain near to master/controlled domain. With the help of tactile support engines at the network edge, intelligent edge server is able to construct the full or a part of the remote environment and send the haptic feedback to the controller in the master domain. As a result, the controller can achieve stable and transparent operation. At the same time, the tactile support engine can be leveraged to enhance the local model that predicts the control command from the master domain. This predicted control signal can be transferred to the teleoperator in the controlled domain in case the control signal is delayed or lost. Moreover, the tactile support engines can be used for network resources (computing, communications, and caching) allocation. Therefore, the tactile support engine has a huge potential for the realization of Tactile Internet.

B. Intelligent edge computing can satisfy the distinct QoS requirements of the Tactile Internet. The QoS requirement in Tactile Internet is more stringent than that for the IoT. Subsequently, based on the reliability, latency, and number of connected tactile/haptic devices, the two types of the QoS requirements (although these two types have overlapping requirements), normal- and carrier-grade can be achieved by intelligently allocating and execute the tasks in tactile edges. Basically, leveraging the distributed, closer to data sources, and intelligence nature, the intelligent edge computing can enhance the co-operations among the multiple tactile edges at both ends of the Tactile Internet

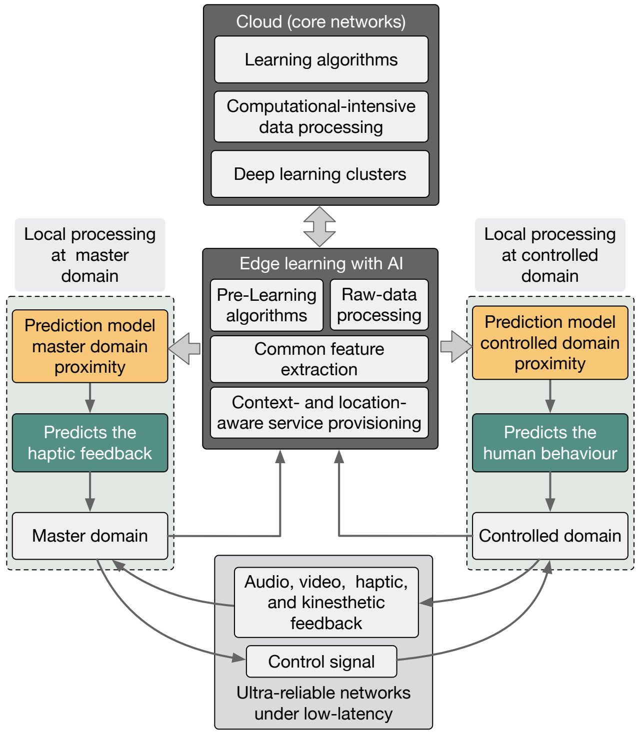

Figure 4: Edge intelligence in Tactile support engines

C. Edge analytics can be used to predict the channel state in the network domain. Communication resource allocation is also an integral part in the network domain. It is always a challenge to accurately predict the time-variant channel state. Edge intelligence can leverage the powerful computation resource and the massive historical channel states to enhance channel state estimation, thereby optimizing the radio resource allocation and improving wireless communication reliability.

D. Intelligent edge computing increments/enhances the learning capability. An important factor is that Tactile Internet services need input not only from one device but from multiple devices. Thus, the collaboration at the network edge can significantly reduce the amount of data to be send to the global controller. Besides, due to the massive amount of data collected by the heterogeneous end-devices including wearable devices and tactile/haptic device, residing at the tactile edge or in tactile support engines can train the neural network model by downloading the parameters from the global controller. After the training, the locally trained parameters can be sent back to the global controller, resulting an improvement in the learning capability.

VI. Main Research challenges to apply Intelligence in Edge-Fog Computing for Tactile Internet

To build the intelligence at the Tactile edge, we outline the following research challenges to meet the ultra-high reliability (note that QoE and QoT are also performance metric in Tactile Internet) and latency at the millisecond scale in the Tactile Internet.

A. Feature Extraction from High Dynamic Environment

To meet the requirements of ultra-low latency and ultra-high reliability, a joint task offloading and resource allocation strategy needs to quickly response to the task computation request. However, the uncertainty of the communication network and the limited computation capability (compared to the cloud server) in the edge servers significantly affect the offloading decision-making for the response. At the same time, remote physical environment at the controlled domain is often high dynamic (although spatial and time correlated). As a result, it is not always preferable to use theoretical model and offline solver. Therefore, extracting the features from the high dynamic remote physical environment, computing and communication environment is desirable in an intelligent offloading and resource allocation decision-making algorithm. The extracted features, called as *state*, and then learning the *state-action* values by model training, an intelligent algorithm can quickly respond to the task computation requests in Tactile Internet. Deep convolutional network is a promising way to extract features and get *state-action* values in intelligent edge environment. However, what environment parameters should be input to the convolutional network and how is the architecture of the learning model, should be further studied.

B. Minimizing the response time

One of the major objectives for processing real-time data is to reduce the response time of the applications. Over the past decade, most of the IoT data are offloaded to the remote resource-rich cloud servers which is not suitable for time-critical applications due to long transmission latency. The emerging fog/edge environment can overcome the bottleneck by processing the data nearer the end-devices. Thus, an intelligent fog/edge environment needs to be designed for making an optimal data offloading to the suitable edge servers close to data sources minimizing the response time thereby meeting multiple constraints in Tactile Internet.

C. Data for model training

Most of the intelligent models including traditional machine learning (supervised and unsupervised learning), reinforcement learning, etc. require either a huge amount of historical data ‘or frequent interaction with the environment for training. Such types of models are suitable for analyzing the data at resource-rich cloud servers; however, due to the limited computation capability and storage, it may not be suitable to train such models at the fog/edge level. Nowadays, Federated Learning (FL) [12] [13], and distributed deep reinforcement learning are introduced to train the overall network using multiple distributed edge servers. Thus, one of the most important research challenges is to train the model at resource-constrained edge server with a limited set of real-time data.

D. Placement of the learning agent

Placement of an agent in a suitable edge server for taking an optimal offloading and scheduling decision is another important research challenge in edge-fog computing. In general, an agent takes an optimal decision based on multiple QoS parameters including resource usage, latency, bandwidth, network congestion, energy consumption, etc. while meeting multiple constraints. This requires an edge server which has enough computing resource capacity for training or analysis the objectives and stores some important result for further evaluation. Moreover, an agent needs to acquire and analyze the real-time data efficiently and feeds them for further training, which requires an optimal edge server. As a result, placing an agent to a suitable edge server for data analyzing and model training is another important research aspect in the fog/edge environment.

E. Sample collection for network parameters

In the real world, it is quite difficult to collect the samples of the network parameters including available bandwidth, network congestion, traffic rate, etc. due to the communication uncertainty between the edge devices. In addition, due to the limited computation power and battery endurance, application scheduling varies from device to device which makes it difficult to receive the synchronized network data at each epoch. As a result, an intelligent edge-fog computing paradigm needs to be addressed for controlling the network parameters and train the network for efficient data transmission with minimum traffic congestion.

VII. Intelligent Edge computing in Tactile Internet: An intuitive Idea

Although the earlier work in intelligent edge computing laid a strong foundation on the application or usage of machine learning in edge computing system, when we aim to apply these techniques in the Tactile Internet, several issues arise. It is worthwhile to note that the assumption underlying model often differs from the real-time environment, resulting sub-optimal solutions. To satisfy the latency and reliability in terms of QoE and QoT, and tactile service availability, the Tactile Internet desires running the intelligent algorithms quickly and accurately in the network edge, however, the limited resource in the network edge (compared to the resource-rich cloud servers) limits the intelligent algorithms, such as deep learning. In addition, considering end-device's privacy, the raw data must not leave its origin. How to obtain an accurate model by training limited amount of data is another issue that should be addressed.

Applicability of Federated Learning in Edge Intelligence

Recently, FL [12] [13], a type of distributed machine learning, is being studied for addressing several real-world situations. Unlike the tradition machine learning approach where the end-devices send massive collected data to remote cloud, to train the deep neural network, in FL the end-devices individually train the network using the local data and send the training gradient weights to the remote cloud instead of data. The remote cloud builds a global model by aggregate the training gradient weights from a large number of end-devices and then push the global model with better inference accuracy back to the end-devices.

One interesting feature of the FL is that the data always stay at the device itself, never leaves from the device, resulting low data leakage (refers to the user privacy). Several categories of FL are classified how the data samples and the features are collected from the multiple distributed devices. For example, horizontal FL shares the same features but on different samples. On the contrary, vertical FL is more suitable to the scenario where the multiple devices contain same sample space; however, the features are different. Recently, federated transfer learning is being studied for the scenario when not only the two data sets differ in their sample space, but also the feature spaces are different. The main idea is the common feature is learned based on the common sample space, then the learned common representation is used to predict the sample points for the one-sided feature space. This type of FL has a huge potential to be applied in several applications [14], even in Tactile Internet where a massive amount of the collected data by the end-devices can be useful using this transfer learning technique.

In fact, the transmission cost is reduced in FL where only the training gradient weights are transferred to the remote cloud compared to the case when the aggregated data collected from the devices are sent for learning process in conventional method. However, keeping the delay deadline in mind, the frequent updating the model parameters from the end-device to the global model that resides at the cloud become a bottleneck. To solve it, several approaches (e.g., reducing the number of parameters to be updated, compression using subsampling) already exist in the literature. Federated average learning has been discussed with reduced communication costs due to uploading the model parameters updates. This could be a possible direction to apply FL while minimizing the updating frequency and reducing the transmission delay for the Tactile Internet applications, particularly, mission-critical applications. A federated learning-based model in master domain is illustrated in Figure 5.

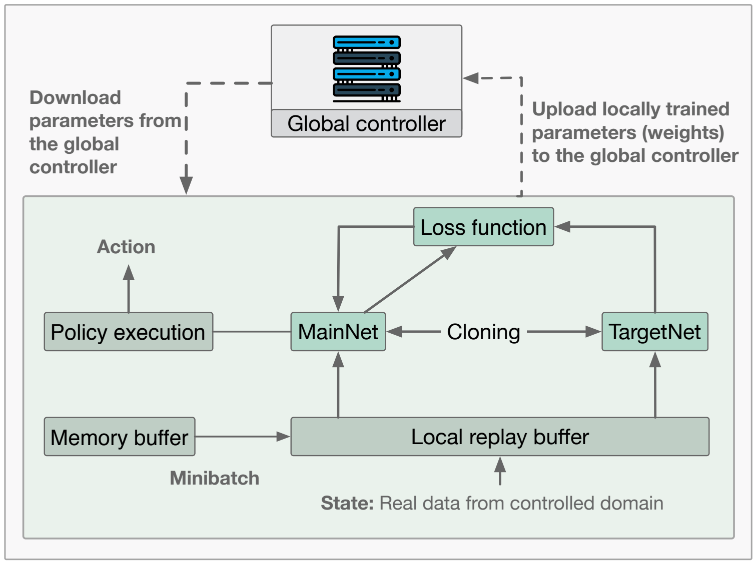

Figure 5: Intelligent edge computing in master domain.

VIII. Future Research Directions and takeaway messages

A. Distributed learning using Model Parallelism or Data Parallelism

Most of the intelligent neural network models including deep learning, Q-learning, etc. are both computation and memory intensive. As a result, it is not feasible to complete the training or acquire the result of the larger and complex model by a single edge server. Therefore, the complex models are trained in distributed edge servers including model parallelism or data parallelism or their combinations. However, model training at distributed edge servers should further consider the data privacy, configuration of the edge devices, wireless communication environment, etc. Nowadays, the FL can copy the whole intelligent neural network model of different edge servers using the concept of data parallelism. As a result, training the whole network model by partitioning the model into multiple independent segments and assigning the segments on the resource-constrained distributed edge servers may be a more feasible and practical solution.

B. Transfer Learning-aware Training at edge servers

To train and deploy a complex intelligent model on edge servers is a challenging task due to the limited computing power and battery endurance. As a result, *Transfer Learning* (TL)-aware [12] training on the resource-constrained edge servers is useful in order to reduce the amount of training data and speed up the training process. By using TL, the edge servers use the cross-model transfer technique for training across different sensing modalities and acceleration of the overall training process. Moreover, the TL-aware training process helps to train the lightweight intelligent models (at edge servers) using the information of well-trained intelligent models (at cloud servers) to improve the accuracy of the network. Thus, TL-aware training is an important research aspect in edge computing with different types of perceptual data.

C. Advancement of FL at edge servers

FL-aware training at the edge servers for improving the accuracy and speed up of the intelligent network model is another open research challenge. Existing FL-aware techniques can mostly process a hundred devices in parallel and train the data in a synchronizing order. However, the limited computation power and battery endurance of the edge servers and various scheduling strategies make it difficult to synchronize the data for training at the end of each iteration. Moreover, at a certain timestamp, maybe some of the edge servers are not available due to task overloading or some uncertainty in the network may cause infrequent task training. The above-mentioned bottlenecks can be optimized by introducing asynchronous-aware FL at the edge servers. As a result, adjusting FL at different edge devices and train the network environment still a research challenge.

D. Reduction of data for computation at edge server

The end-devices from the same area of a globe may request for the recognition results for similar objects which increases the redundancy at the edge servers. By caching the frequently requested and important results in the edge servers can reduce the processing overhead of the devices and reduce the processing delay and energy consumption. As a result, process and analyze the similar applications at edge servers without dabbling intelligent methodologies such as DRL, RL, etc. is another research aspect in this domain.

E. Incentive and secure data offloading at edge servers

The real-time end-devices prefer to offload the data to the nearby edge servers with an aim to reduce the processing delay and energy usage. However, an optimal offloading policy still faces several issues while offloading the data such as a) an incentive mechanism for simulating the real-time data at edge servers with an intelligent model and b) a proper security mechanism for avoiding the risk from anonymous edge servers.

IX. CONCLUSION

In this article, we have outlined the potential advantages of the intelligent offloading using fog-edge computing to meet the carrier-grade QoS requirement in Tactile Internet. In this context, we have discussed how edge intelligence can be leveraged to predict haptic feedback from the controlled domain to reconstruct kinematics and dynamics model and interpretation of the command signal from the master domain in advance. Moreover, the edge intelligence can be further utilized for computational offloading with the help of tactile support engines in Tactile Internet. Finally, the major challenges of the edge intelligence in the context of Tactile Internet applications are identified and we have highlighted the potential research directions in edge intelligence for Tactile Internet. We conclude the article on the note that intelligent edge offloading would be one of the frontiers in future development for Tactile Internet Technologies to meet the QoS and QoT in Tactile Internet. A worthwhile and promising pursuit would be to investigate the cross-technology in applying edge intelligence in Tactile Internet.

ACKNOWLEDGEMENT

This work was supported in part by the National Natural Science Foundation of China under Grant 61901128.

REFERENCES

[1] G. P. Fettweis, "The Tactile Internet: Applications and challenges," IEEE Vehi. Technol. Mag., vol. 9, no. 1, pp. 64-70, 2014.

[2] M. Dohler, T. Mahmoodi, M. A. Lema, M. Condoluci, F. Sardis, K. Antonakoglou and H. Aghvami, "Internet of skills, where robotics meets AI, 5G and the Tactile Internet," in Prof. European Conference on Networks and Communications (EuCNC), 2017.

[3] O. Holland, E. G. Steinbach, R. V. Prasad, Q. Liu, Z. Dawy, A. Aijaz, N. Pappas, K. Chandra, V. S. Rao, S. Oteafy, M. A. Eid, M. A. Luden, A. Bhardwaj, X. Liu, J. Sachs and J. Araujo, "The IEEE 1918.1 "Tactile Internet" standards working group and its standards," Proceedings of the IEEE, vol. 107, no. 2, pp. 256-279, 2019.

[4] M.Eid, J.Cha, and A.E. Saddik,"Admux: An adaptive multiplexer for haptic - audio -visual data communication", IEEE Trans. on Instrumentation and Measurement, vol. 60, no. 1, pp. 2131, Jan 2011.

[5] K.S.Kim, D.K.Kim, C.Chae, S.Choi, Y.Ko, J.Kim, Y.Lim, M. Yang, S. Kim, B. Lim, K. Lee and K. L. Ryu, "Ultrareliable and low-latency communication techniques for Tactile Internet services," Proceedings of the IEEE, vol. 107, no. 2, p. 376– 393, 2019.

[6] A. Aijaz, M. Dohler, A. H. Aghvami, V. Friderikos and M. Frodigh, "Realizing the Tactile Internet: Haptic Communications over Next Generation 5G Cellular Networks," IEEE Wireless Commun, vol. 24, no. 2, pp. 82-89, April 2017.

[7] M. Simsek, A. Aijaz, M. Dohler, J. Sachs and G. Fettweis, "5G-enabled Tactile Internet," IEEE J. Select. Areas Commun., vol. 34, no. 6, pp. 460-473, 2016.

[8] H. Xing, L. Liu, J. Xu and A. Nallanathan, "Joint Task Assignment and Resource Allocation for D2D-Enabled Mobile-Edge Computing," IEEE Transactions on Communications, vol. 67, no. 6, pp. 4193-4207, 2019.

[9] S. Jošilo and G. Dán, "Decentralized Algorithm for Randomized Task Allocation in Fog Computing Systems," IEEE/ACM Transactions on Networking, vol. 27, no. 1, pp. 85-97, 2019.

[10] M Mukherjee, M Guo, J Lloret, R Iqbal, Q Zhang, Deadline-aware Fair Scheduling for Offloaded Tasks in Fog Computing with Inter-fog Dependency, IEEE Communications Letters. 2019

[11] B. Cao, L. Zhang, Y. Li, D. Feng and W. Cao, "Intelligent Offloading in Multi-Access Edge Computing: A State-of-the-Art Review and Framework," IEEE Communications Magazine, vol. 57, no. 3, pp. 56-62, 2019.

[12] Q. Yang, Y. Liu, T. Chen and Y. Tong, "Federated Machine Learning: Concept and Applications," ACM Trans. Intell. Syst. Technol, vol. 10, no. 2, pp. 1-19, 2019.

[13] K. B. e. al, "TOWARDS FEDERATED LEARNING AT SCALE: SYSTEM DESIGN," in Proceedings of the 2 nd SysML Conference, Palo Alto, CA, USA, 2019.

[14] X. Wang, Y. Han, C. Wang, Q. Zhao, X. Chen and M. Chen, "In-Edge AI: Intelligentizing Mobile Edge Computing, Caching and Communication by Federated Learning," IEEE Network, vol. 33, no. 5, pp. 156-165, 2019.

[15] J. Sachs et al., "Adaptive 5G Low-Latency Communication for Tactile Internet Services," in Proceedings of the IEEE, vol. 107, no. 2, pp. 325-349, Feb. 2019

Mithun Mukherjee received the B.E. degree in electronics and communication engineering from the University Institute of Technology, Burdwan University, Bardhaman, India, in 2007, the M.E. degree in information and communication engineering from the Indian Institute of Science and Technology, Shibpur, India, in 2009, and the Ph.D. degree in electrical engineering from the Indian Institute of Technology Patna, Patna, India, in 2015. Currently, he is a Professor with the College of Artificial Intelligence, Nanjing University of Information Science and Technology, Nanjing, China. His research interests include wireless communications, fog computing, and intelligent edge computing.

Mian Guo received her Ph.D. in communication and information systems from South China University of Technology, China in 2012. She was a visiting professor with the University of Ottawa, Ottawa, Canada in 2016, and a visiting professor with Beihang University, China, in 2017. She is currently a Lecturer at Guangdong University of Petrochemical Technology, China. Her research interests include resource allocation, QoS provisioning in computer and communication networks, edge computing, and deep reinforcement learning.

Jaime Lloret received his M.Sc. in physics in 1997, his M.Sc. in electronic engineering in 2003, and his Ph.D. in telecommunication engineering in 2006. He is an associate professor at Universitat Politecnica de Valencia, chair of the Research Institute IGIC, and the Head of the Active and Collaborative Techniques and Use of Technologic Resources in the Education (EITACURTE) Innovation Group. He was the Vice-Chair for the Europe/Africa Region of Cognitive Networks Technical Committee of the IEEE Communications Society for the term 2010–2012 and the Vice-Chair of the Internet Technical Committee of the IEEE Communications Society and the Internet Society for the term 2011--2013. He is also the Chair of the Working Group of the Standard IEEE 1907.1. Since 2016, he has been the Spanish Researcher with highest h-index in the Telecommunications journal list according to the Clarivate Analytics Ranking. He is an ACM Senior and IARIA Fellow.

QI ZHANG received the M.Sc. and Ph.D. degrees in telecommunications from Technical University of Denmark (DTU), Denmark, in 2005 and 2008, respectively. She is an Associate Professor with the Department of Engineering, Aarhus University, Aarhus, Denmark. Besides her academic experiences, she has various industrial experiences. Her research interests include Internet of Things, mobile edge computing, Tactile Internet, compressive sensing, sensor data compression and storage. She is serving as an Editor for EURASIP Journal on Wireless Communications and Networking. She was a Co-Chair of the Co-operative and Cognitive Mobile Networks (CoCoNet) Workshop in the ICC conference 2010-2015 and was a TPC Co-Chair of BodyNets 2015.

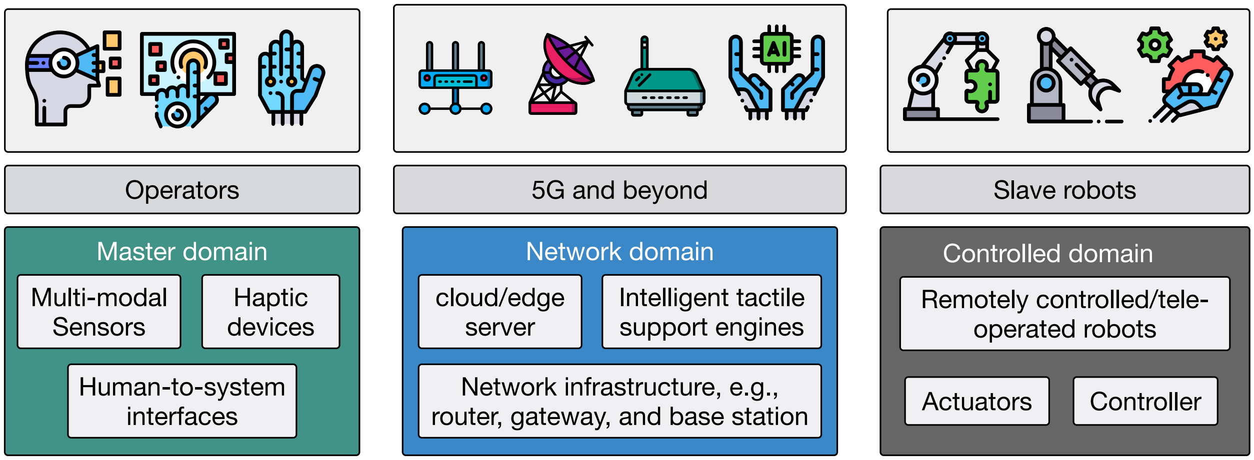

Figure 1: The main domains in Tactile Internet.

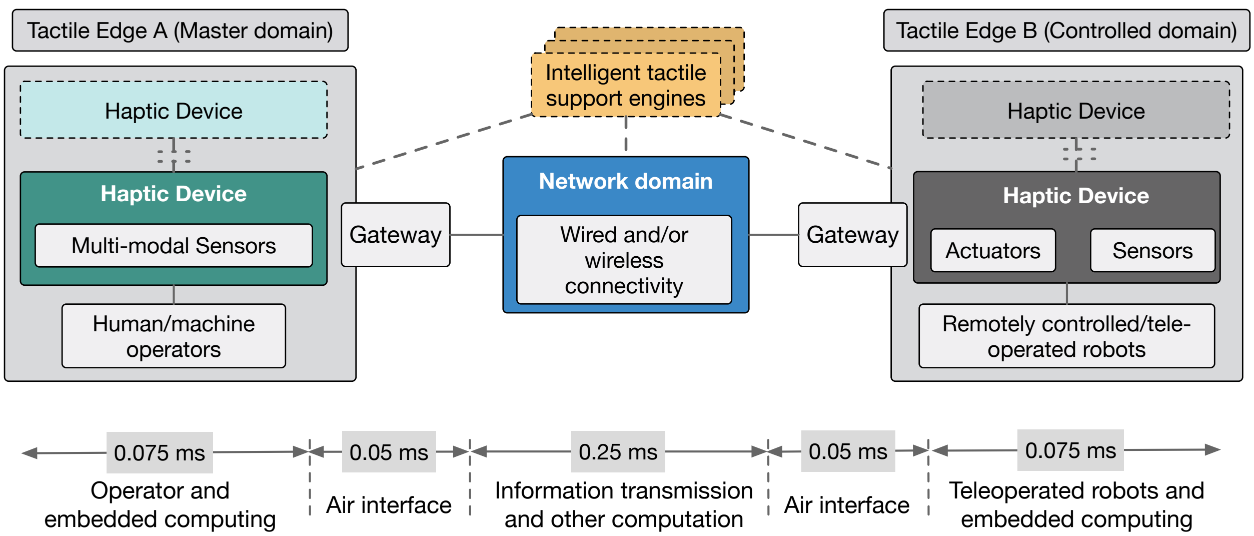

Figure 2: The envisioned end-to-end latency in Tactile Internet

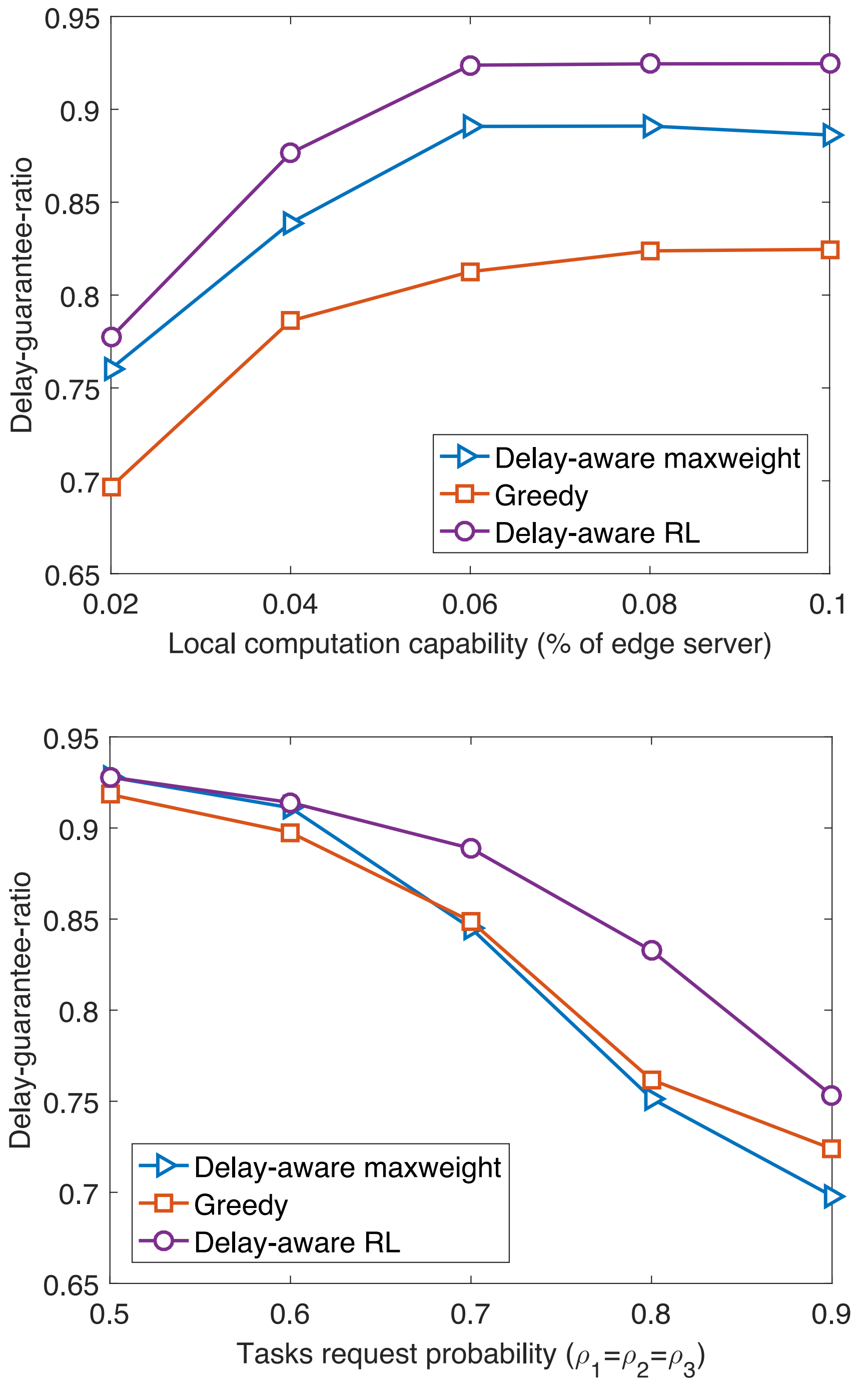

Figure 3: Performance of the *delay-guarantee-ratio* (a) under various task request probabilities and (b) under various computation capabilities.

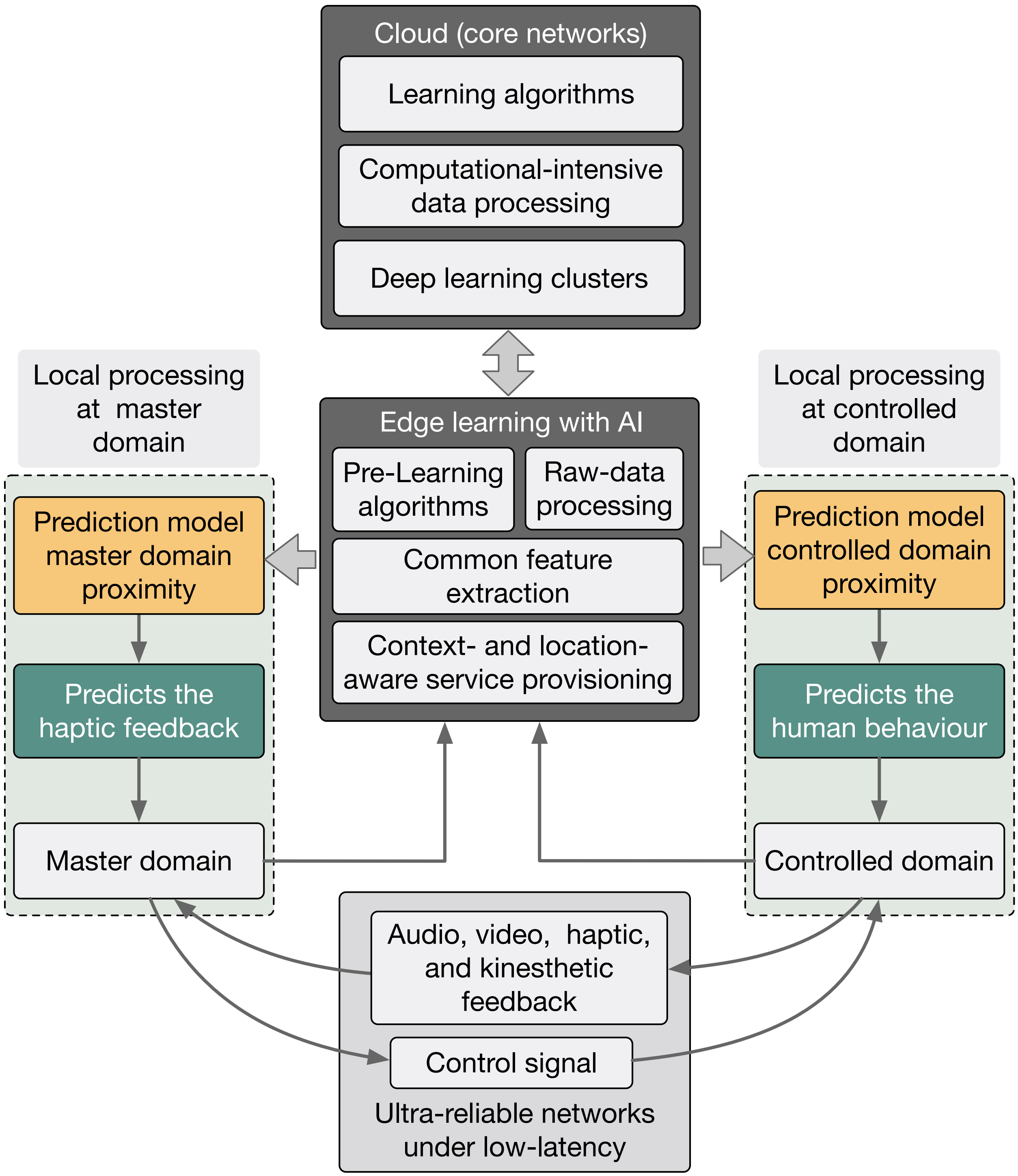

Figure 4: Edge intelligence in Tactile support engines

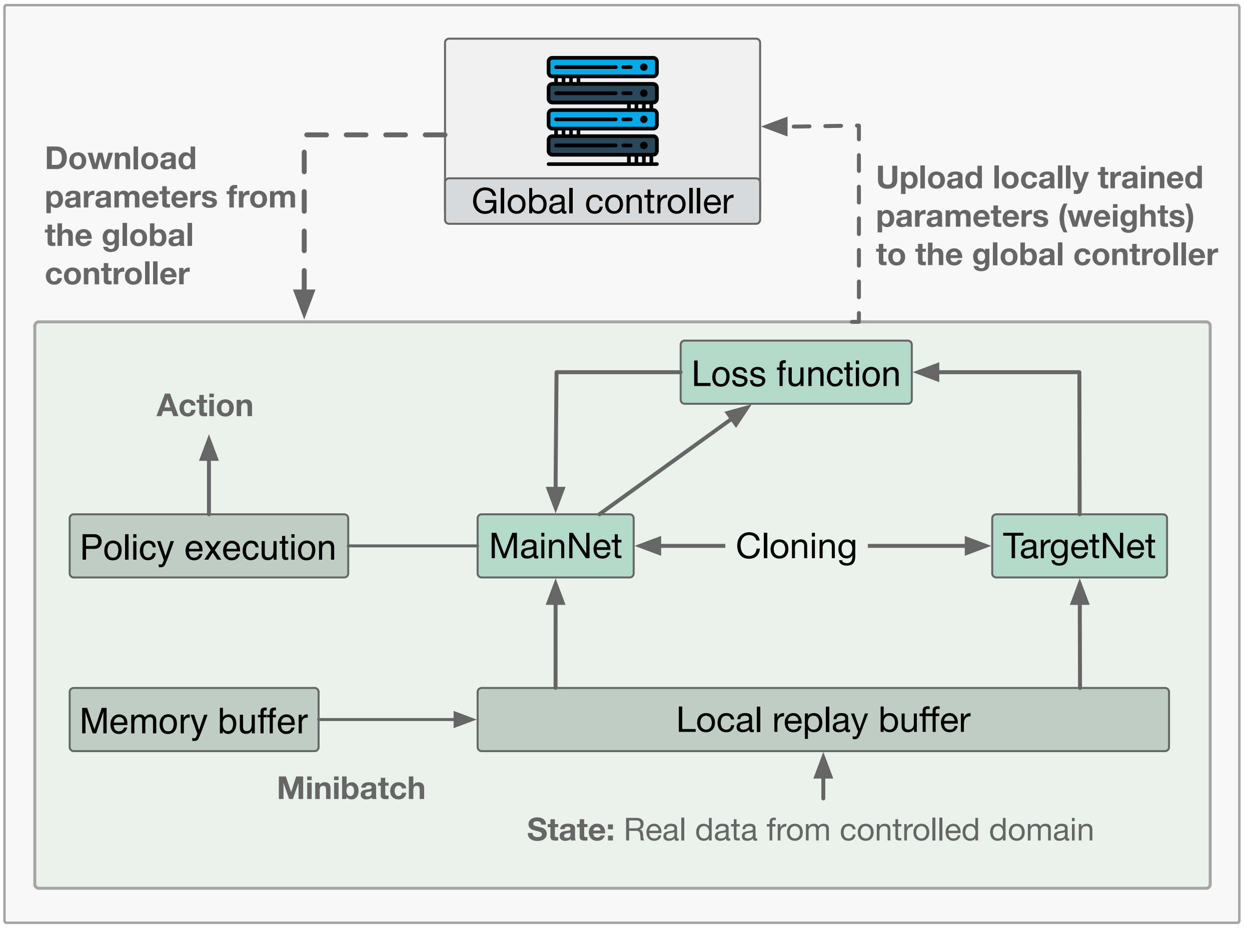

Figure 5: Intelligent edge computing in master domain.