Project Specifications

* Abstract (Not Now)
* Introduction and Background (motivation)

With the rapid development of technology, the mobile edge computing (MEC) was introduced to fill the gap of mobile devices when facing processing computationally intensive and time-sensitive applications such as AR/VR and face recognition. The task offloading has been introduced for dealing these resources to the remote cloud and provide powerful computing ability and unlimited storage. However, the long transmission distance between user equipment (UE) and cloud may cause unpredictable delay, which could also cause network congestion. The resource allocation, and task migration have been introduced to provide computation and networking services for UEs more quickly and efficiently.

In the existing work, the Deep Neural Network is used to generate the offload decision, and the linear regression is used for prediction-based offloading in MEC. The limitations of these solutions are: (1) they were not suitable for real-time offloading. (2) and could only be used in the LAN network. Also, the convex optimization, semi-positive definite relaxation (SDR) and game theory methods were used for managing resource allocation. However, those methods might face difficulties in below: (1) could be received uncertain user requests due to privacy and unstable system states. (2) the mobility problem has not been considered. The reason of that is because most existing resource allocation algorithms can only be achieved at a certain moment that is not suitable for the real world. Many methods support mobility from different aspects, such as service migration and fine-grained control. Some existing work have be applied to solve this problem, but still has some issues need to be improved on: (1)An energy-aware mobility management (EMM) scheme: is an user-centric approach. This method considered energy efficiency in order to optimize delay in an ultra-dense network, but it hasn’t considered about the decision of offloading proportion. (2)A dynamic mobility aware partial offloading (DMPO) algorithm: to optimize both the offloading proportion and the communication path of offloading after every UE movement. But the limitation is it only provide one dimension.

To overcome these limitations, in this project, the task prediction based intelligent framework and deep Q-network (DQN) task migration will been applied in the MEC. The goal is to achieve optimal performance in MEC environment by reducing energy consumption and time delay by combining the improvement of computation offload, managing resource allocation, and executing task migration. The framework is integrated the Deep Reinforcement Learning (DRL) and Federated Learning (FL) for optimizing computing and caching. The collected data could be mining deep and cached depend on the popularity of the content, which provide a cognitive system for MEC. GRU algorithm formulates the optimal offloading strategy by using cognitive functions. And DQN helps the agent to make optimum migration decision by establish the map between action and state to gain reward values.

* Research Question and Objectives

**Research Question**: How the resources can be offloading and cache within the mobility of users with an intelligent framework in the MEC.