Green 5G Small Cell Network Optimization and Simulation

by

ZANG, Mingyuan

Supervisor: Dr. SI, Pengbo

Submitted to the Fan Gongxiu Honors College, BJUT

in Partial Fulfillment of Requirements for

the Degree of Bachelor of Engineering

in Electronic Information Engineering

Beijing University of Technology

Beijing, China

June, 2018

独 创 性 声 明

本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得北京工业大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

签名：____________ 日期：__________

关于论文使用授权的说明

本人完全了解北京工业大学有关保留、使用学位论文的规定，即：学校有权保留送交论文的复印件，允许论文被查阅和借阅；学校可以公布论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存论文。

（保密的论文在解密后应遵守此规定）

签名：____________ 导师签名：____________ 日期：__________

Abstract

The development of 5G mobile communication systems is aiming at meeting the boosting needs of mobile Internet traffic in the next 10 years. With the advent of various kinds of applications, future mobile networks will face great challenges of higher capacity, lower power consumption, lower cost, etc. Some key technologies have been put forward in the next generation mobile communication technology (5G) such as millimeter waves and Small Cell to meet the requirements. The development of small cell networks offers higher chances in increasing capacity while reducing power consumption, since the small cells are effective supplements to macro cells with less power consumption in each single small cell. The problem is that deploying dense small cells may not actually save the total power of the system. To solve this problem, the small cells could be switched off during the time when traffic load is not that high, and switched on when large number of users appear in the hot spot area. However, switching off the small cells may lead to lower chance of offloading and harm the throughput. It offers the motivation of this project to find an algorithm to realize this switching on/off process automatically to save more energy without sacrificing much capacity. The project focuses on the office scenario where the traffic load can be different on weekday and weekend. Reinforcement Learning Algorithm has been adopted and further developed. The performances of the algorithm are also evaluated by building a simplified LTE model with one macro cell and several small cells in MATLAB. Besides, the main traffic data of the network comes from the simulation in OPNET which is considered as more realistic. To show the details of power consumption changing, we take the simulation time as a whole day (24 hours). Two algorithms to obtain the best policy, on-policy SARSA Algorithm and off-policy Q-Learning Algorithm are evaluated and the results show that Q-Learning performs better in saving energy. By using the algorithm, the system power consumption can be largely reduced during peak hours and the decrement of throughput is within an acceptable range. The deployment of around 110 small cells in a macro cell shows the best performance.

Key words: Small Cell; Reinforcement Learning; Power -saving

摘　要

发展 5G 移动通信系统旨在满足未来 10 年移动通信网络流量的需求。随着各种应用的出现，未来移动网络将面临更高容量、更低功耗、更低成本等重大挑战。第一代移动通信技术（5G）已经提出了一些关键技术，如毫米波和小蜂窝网络（Small Cell）技术来满足要求。小蜂窝网络的发展为增加网络容量提供了机会，它能够有效补充宏蜂窝的网络容量和网络覆盖。同时因为小蜂窝小区能量消耗较少，还能够降低功耗。然而问题在于部署密集的小蜂窝小区在实际情况下可能并不会节省系统的总功耗。为解决这个问题，可以在流量负载不高的时候关闭小蜂窝小区，并在热点地区出现大量用户时开启小蜂窝小区以节省能耗。但是，关闭小蜂窝小区可能导致从宏蜂窝卸载流量的可能性降低并会牺牲掉部分吞吐量。为了解决这些问题，此项目的目标是找到一种可以自动实现开关功能的算法，在没有牺牲太多网络容量的情况下节省更多的能量。该项目主要研究对象是工作日和周末网络负荷不同的办公室场景。通过采用和进一步发展强化学习算法来实现自动开关小蜂窝小区。算法建立后，在 MATLAB 中建立一个具有一个宏蜂窝小区和若干个小蜂窝小区的简化 LTE 模型来进一步验证算法的具体表现。这其中，网络流量负载数据来自于使用 OPNET 软件进行的仿真，使用该软件所生成的仿真结果更贴近实际网络情况。仿真过程中，为了显示网络功耗变化的细节，仿真时间设定为一整天（24 小时）。为了求解最佳策略，模型中考虑了两种算法：SARSA 算法和 Q 学习算法，最终结果显示 Q 学习算法在节约能耗方面表现更好。通过引入该算法，可以在高峰时段显著降低系统功耗，并且吞吐量的降低值在可接受的范围内。同时结果也表明在宏蜂窝小区中部署大约 110 个小蜂窝小区时，本项目提出的算法在节约能耗方面可以展现出最佳性能。

关键词：小蜂窝小区；强化学习；节能

TABLE OF CONTENTS

Chapter 1 Introduction

With the traffic requirements boosting, architectures of previous mobile systems cannot sustain the load anymore. The evolution to 5G mobile communication technology seems to be necessary. According to [1], the major requirements of 5G technology are large mobile data capacity and low power consumption. Various technologies have emerged to meet these requirements from different perspectives. Small Cell is one of the technologies that has been put forward.

Power consumptions are related to many factors. Among them, base station is the most power-intensive element [2]. Even with low level of capacity and coverage, the base station power consumption cannot be ignored. The development of small cell network would efficiently lower the power consumption. Nonetheless, in ultra-dense cellular networks, the density of small cells has a threshold. “When the density is beyond that threshold, the backhaul network capacity will reduce.” [3] In this case, the tradeoff need to be balanced to achieve as low power consumption as possible without sacrifice the network capacity. According to the work in [4], basic types of base stations were modeled and they analyzed the issue from the perspective of energy aware components in the base station. Their work offers a reference of modeling base stations. While the work in [5] viewed the problem from another side. They optimized the network deployment strategies by adjusting the size of small cells dynamically. Both of their works are inspiring and it may be possible to combine their ideas together to reach a better result.

Due to the dynamic feature of the network, it is costly for manpower to achieve that kind of adjustment. So intelligent algorithm need to be introduced into the dynamic controlling process. Q-learning is the common used one. In the research [6], they introduced a new initialization method into the algorithm. But the research concentrates more on interference management than reducing energy consumption. Paper [7] developed cooperative Q to improve traditional algorithm by sharing local knowledge with other cells periodically. But the algorithm's performance is largely dependent on the environmental and operational parameters.

In a nutshell, some work has been done these years to achieve higher capacity and lower energy consumption of the network as much as possible. The goal of this project focuses on the energy saving method in the office scenario where the traffic load is large and can be different on week day and weekend. OPNET Modeler has been used to get more realistic traffic data in the office scenario. MATLAB has been used in modeling work and evaluate the performance of the algorithm in the office scenario. A comparison between the two algorithms, on-policy SARSA Algorithm and off-policy Q-Learning Algorithm, has been made to illustrate why Q-Learning is used more commonly. Further development on parameters of the Q-Learning algorithm has been done to make the algorithm to meet the capacity requirements and show better performance in saving power. More specific details of the algorithm have been tested to evaluate the algorithm from various aspects.

This thesis introduces project details in 4 parts. Chapter 2 introduces some basic concepts about 5G and Small Cell technology. Chapter 3 demonstrates the methodology used in the project. Chapter 4 presents the system model developed to solve the problems. Chapter 5 lists all the results and corresponding analysis.

Chapter 2 The Brief Introduction to 5G Technology and Small Cell

2.1 Introduction to 5G Technology

2.1.1 Requirements of 5G technology

Every new generation of mobile communication network delivers faster speed and more functionalities to the users. The 1st generation technology gives users the very first cell phones. The 2nd generation technology (e.g. Global System for Mobile Communication) allows the users to send SMS (Short Message Service). The 3rd generation technology (e.g. Wideband Code Division Multiple Access) enables the users to connect to the network and realizes the roaming around the world. The 4th generation (e.g. Long Term Evolution) offers much higher bit rates and supports various information transmission like voice, video and image, etc. However, as more kinds of applications come online (e.g. Internet of Things), 4G technology has almost reached its limit in capacity. Under this circumstance, the 5th generation technology is developed to meet these boosting requirements in large amount of data traffic.

The main targets for 5G technology is to make the mobile communication network with larger coverage, higher capacity in hotspot areas, shorter delay, higher reliability and lower power consumption. [8] The detailed requirements are listed below:

1. Data traffic should increase by a factor of 100 and peak throughput should reach at least 100 Gbps/km^2;
2. The number of devices connected to the network is expected to grow by a factor of 100;
3. Peak rate reaches to at least 10 Gbps;
4. User can obtain a rate of 10Mbps;
5. Short delay and high reliability;
6. Spectrum utilization is 5 to 10 times higher than its in 4G network;
7. Low power consumption.

2.1.2 Key technologies

To meet the targets, three technologies have been discussed as key technologies for 5G:

1. Millimeter Waves

The common devices occupy the radio frequencies spectrum from 3kHz to 6GHz. However, as more applications come out, these frequencies start to get more crowded, which will cause lower bit rates and higher drop rates. The solution is to develop more spectrum resources. The spectrum from 6GHz to 300GHz has been considered to be available for mobile devices. The challenges of this solution are that the millimeter waves will have power loss when they travel through buildings or other obstacles. They also stand a chance to be absorbed by plants and rain.

2. Small Cell

Nowadays, the most common technology is to transmit signals through a long distance from macro base stations. However, as millimeter waves cannot travel well through the obstacles, those mobile devices far away from the base station cannot connect to the network. To solve this problem, Small Cell technology can ensure better service quality to mobile device by building thousands of low-power micro base stations. An example of this kind of small cell equipment is shown in figure 2.1. The deployment of small cells will be helpful especially in the city areas where the traffic demands are large. More details about this technology will be discussed in the following section.

Figure 2.1 Example of Small Cell Equipment (Alcatel-Lucent 9760)

3. Massive MIMO

4G base stations now commonly have about dozens of ports for antennas and they are responsible for all cellular traffic. The development of massive MIMO (Multiple-Input Multiple-Output) base stations will improve that and support around a hundred ports. This would raise the capacity of the network by a factor around twenty. The challenge of this technology is that antennas broadcast signals in every direction at once, which can cause large interference between each other.

2.1.3 Power consumption problem in 5G network

The power consumption of a network largely relies on the traffic load. With the increase in the number of users, the power consumption of the network will increase geometrically. Especially in the upcoming 5G era, traffic load will explode and the power consumption of the communication network infrastructure will become a serious problem. Another reason to make this problem more serious is the low efficiency of the communication network infrastructure. As no effective measures have been taken, the communication network works under low traffic load but high power consumption for most of the time.

To build a green 5G network, the power consumption can be cut down from the aspects of network architecture, network deployment, resource scheduling, link-level technologies, etc. The network architecture can be improved by using a flattened IP network where the network layer is simple. The distributed architecture is implemented, making the wireless resource management more flexible and efficient, and achieving the seamless handover of users in the network. Network deployment is considered as increasing the network capacity through the deployment of low-power nodes like Micro Cell, Pico Cell, Femto Cell. Resource scheduling is needed because the network is designed for the resource demands during the peak hours previously. However, in the real case, the traffic load will be high during the daytime while low during the nighttime, as well as the traffic load can be high in the work area while low in the residential areas during the working hours. When network traffic load is low, the capacity of the base stations will be redundant, resulting in wasted energy. The link-level technology is a kind of improvement on link-level resources to increase the network throughput, shorten the delay and reduce the power consumption. One example is massive MIMO technology mentioned in previous section.

2.2 Small Cell

As mentioned in previous section, Small Cell is a key technology to supplement the coverage of macro base stations. It breaks a cell site into several smaller pieces. Small cells are defined as cellular radio access nodes with low power. They are called “small cells” because they have a shorter range of 10 meters to several hundred meters compared to a macro cell. Small Cell has the products of Femto cell, Pico cell, and micro cell. The deployment of SCs (Small Cells) could increase network capacity and make better use of available spectrum. As a supplement to macro cells, deployment of small cells enables the operators to offer a better data services to the subscribers at a lower cost. Figure 2.2[9] shows how the femtocell can be deployed in the indoor scenario and offload the macro cell.

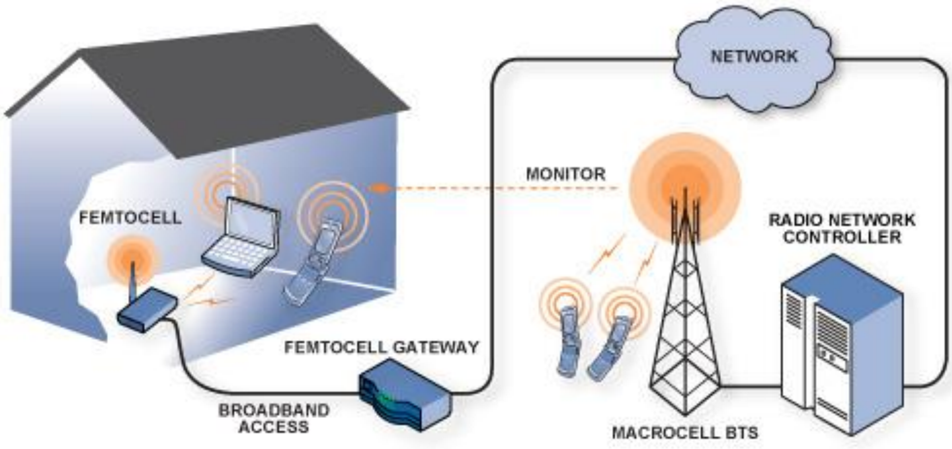

Figure 2.2 Femtocell Deployment Scenario

As shown in figure 2.3, the upper graph shows the scenario of a macro cell. Since there is only one base station in the cell, it is hard for the signals transmitted from base station to reach the mobile device at near the edge of the cell. This will cause the uneven scenario where the power consumption of the cell is large to support the large amount of the devices near the station, whereas those at-the-edge devices still cannot be connected to the network. By deploying the small cells in the network, the micro base stations are put in the hot spot areas to offload the macro base station. Meanwhile, they could also be deployed near the edge of the macro cell to increase the coverage.

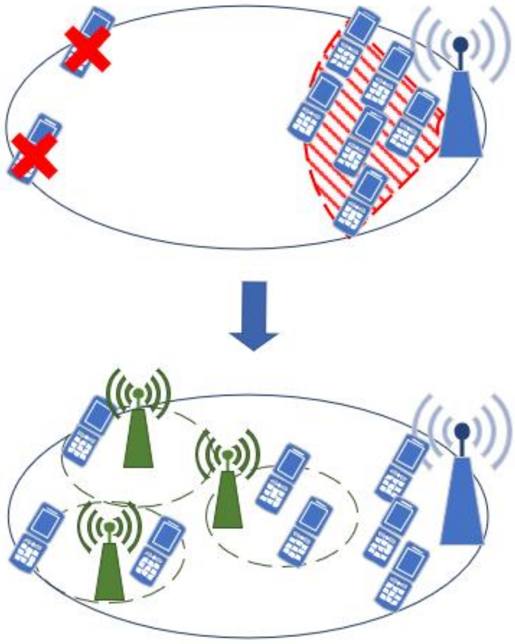

Figure 2.3 Macro Cell vs. Micro Cell

2.2.1 Typical networking scenarios

To supplement of the macro cell, SCs are usually deployed in places where the coverage or capacity of macro cell is limited to support high traffic load of voice and data services from users. They could be deployed indoors and outdoors to supplement the coverage of macro cell and also in hotspot areas to help increase the capacity. Some typical examples are shown in table 2.1[10].

Table 2.1 Typical Networking Scenario of Small Cell

Scenario	Typical Examples
Low Coverage Outdoors	Commercial streets, residential areas, macro cell edge
Low Coverage Indoors	Schools, residential areas, office buildings, shopping malls
Hotspot Areas	Macro cell capacity limited scenarios

2.2.2 Benefits and challenges of Small Cell

1. Benefits

Introducing Small Cell technology into the deployment of network can greatly enhance network capacity, optimize network coverage and provide better user experience. The benefits in details are listed below. [11]

(1) Optimization in three-dimensional coverage: Macro base stations are generally located on the roof and tower height, which have large coverage. However, influenced by the buildings and trees, the coverage is uneven, which means a large number of blind

spots exist. Small cell could form an integrated three-dimensional coverage with the macro station based on its features of small size and easy installation. Small cell can be integrated with the antenna, which would save site costs, shorten the construction cycle and achieve accurate coverage. It can be used for hotspot areas like train stations and residential areas to cover the blind spots.

(2) High-capacity indoor coverage: Mobile broadband network business mainly occurs indoors. In this way, improving indoor coverage would be the key for operators to retain users. When signals from macro base stations penetrate the building into the room, there will be penetration loss. The coverage of indoor voice and data services would be weakened. The introduction of Small cell products can greatly enhance the capacity indoors because of their support for MIMO. Small cell products can be directly deployed in the office or installed in the corridor, through the built-in or external antennas to cover the indoor area.

(3) Effective use of spectrum resources: The introduction of CR (Cognitive Radio) technology in Small Cell can further improve the utilization efficiency of spectrum resources. The CR flexibly reuses free spectrum resources through effective spectrum sensing to meet the data bandwidth requirements. Moreover, the characteristics of high-band spectrum transmission are more suitable for short-distance line-of-sight transmission environments. The spectral multiplexing efficiency is very high, and the interference between transmission links is weak. A large number of high-frequency (6-100 GHz) spectrum resources can be better used by the deployment of Small Cell. [12]

(4) Low power consumption: From the perspective of the UEs (User Equipment), since the signals received from small cell are much stronger than from the macro cell, the phone can maintain service with a relatively low power consumption, which could help extend its battery life time.

(5) Low cost: It could be hard and expensive to find a good location for the macro base station. The deployments of small cell in hotspot or remote areas as well as in emergencies could allow operators to effectively reduce the cost of site, operation and maintenance, which means the reduction in OPEX (Operating Expense).

2. Challenges

Regarding the benefits, there are still some challenges to realize them and achieve perfect performance of Small Cell.

(1) Mobility management: Since the deployment of small cells are pretty flexible, unreasonable configuration of neighboring cells could lead to complex network, which would be detrimental to network planning and mobility management. Too many neighboring cells may result in excessive handovers, affecting the user experience, severely consuming the energy of the UEs, and consuming network resources. The number of neighboring cells and the handover parameters need to be properly configured for different scenarios to reduce unnecessary measurement of neighboring cells and avoid unnecessary switching between macro cell and small cell. [13]

(2) Backhaul network: The backhaul of Small Cell usually relies on the public network, such as DSL (Digital Subscriber Loop). But since the public network would have the load of fixed broadband users at the same time, its bandwidth is usually difficult to guarantee, which could lead to congestion. However, services like OAM (Operation Administration and Maintenance) have high requirements in packet loss and latency and QoS (Quality of Service) of different services must be guaranteed at the transmission level. In this way, the backhaul network should also be taken into consideration during the deployment of small cells.

(3) Load balance: UEs will choose their Serving-eNodeB based on RSRP (Reference Signal Received Power) to get large SINR (Signal to Interference plus Noise Ratio). Since the macro cells' transmission power and coverage area are both larger than small cells', less UEs will choose small cell as Serving-eNodeB. Under such case of uneven service distribution, it is difficult to avoid that some small cells are congested while the other small cells are idle. How to effectively use Small Cell resources and efficiently divert network traffic would be the focus of the research.

(4) Operation and maintenance: The introduction of the Small Cell into the LTE network poses great challenges in terms of operation, maintenance, and optimization, which require highly in network intelligence and simplification. The introduction of technologies like SON (Self-Organized Network) could help minimize the human power in network operation and maintenance.

(5) Interference: In a dense heterogeneous network, a macro base station may bring a certain degree of interference (cross-layer interference) to small cell users due to its high transmission power. Although the transmitting power of small cell eNodeB is relatively low, small cell eNodeB are close to each other and mutual interference exists (interference on the same layer). The more intensive the deployment is, the more serious the interference will be. Eventually, the overall performance of such a heterogeneous network will become less effective.

2.2.3 Energy efficiency in Heterogeneous Networks

Network with the deployment of different types of base stations is defined as Heterogeneous Network (HetNet). HetNet can support large traffic flow for more kinds of applications. In addition to the consideration from the capacity and coverage sides, deploying small cells in the network also potentially consumes less power. [14]

Table 2.2 shows the comparison among the 4 main types of base stations deployed in a HetNet. It is not hard to find that micro cells consume less power in transmission than the macro cells. Correspondingly, the cell radius is smaller. Among them, Femtocell is the kind of small cell base station that technically could be deployed by the users. Other kinds of base stations can only be deployed by the operators. [15] In this project, Femto base station is studied since it can be installed by users and are commonly deployed in indoor scenarios like home or office.

Table 2.2 Comparison of Main Types of Base Stations

Base Station	**Power (dBm)**	**Cell Radius (m)**
Macro	43 - 52	> 1000
Micro	33 - 43	250 - 1000
Pico	24 - 33	100 - 300
Femto	< 23	< 50

Although the transmission power of SC base stations is smaller, the challenge is how to balance the tradeoff between the network capacity and the total power consumption. If the number of SC base stations is large in network, the total power consumption could be larger than only relying on the macro base station. While if that number is small, then it would be hard to meet the coverage and capacity requirements in the urban areas. A small number of micro cells will be neither enough to cover the corners of the area nor meet the throughput requirements from the users since large number of users connected to a same base station will share spectrum resources and result in lower throughput.

As mentioned in introduction section, this project is trying to solve part of the challenges and make some optimization in mobility management by introducing reinforcement learning algorithm to switch on/off micro base stations to save energy without sacrificing much network capacity to meet users' demands.

Chapter 3 Methodology

3.1 Design Overview

A network with one single macro cell is considered specifically in this project to study the small cells in it. To automatically control the switching process of small cells, an algorithm need to be put forward and a model of such a network need to be built to evaluate the performance of the algorithm.

The network traffic data comes from OPNET simulation. An LTE network model has been built in the OPNET software and the network scenario is set to office scenario. Several users are placed in the cell and the users have been divided into two groups, one is a group of stuffs and the other is a group of users who just passed by or visited the area. The simulation has been run for a whole day and get the traffic load of a work day. The following works are then based on these sets of data.

Reinforcement Learning is considered as the algorithm because it is based on Markov Decision Process who can decide the next station based on the current state and the reward gained from the environment. There are several kinds of methods in Reinforcement Learning Algorithm. Since in real case, the network environment cannot be fully known beforehand, Temporal Difference Methods are used to solve out the best policy. Among them, on-policy SARSA Algorithm and off-policy Q-Learning Algorithm are developed and compared. The one with the best performance in power saving is selected.

A simplified network model is then built by using MATLAB to verify the performance of the algorithm. The model includes a macro cell base station and several micro cell base stations with plenty of users scattered in the area. Several micro cell base stations are considered operating within a square macro cell area with the length of 3 km. The macro cell case station is fixed at the edge of this area, whereas micro base stations are randomly positioned at the beginning of the simulation. Users are scattered randomly in the area. To simulate the connection between the users and their serving base stations, k-means clustering algorithm is used to cluster the users into each micro base station's group. Each group has a maximum capacity value, if the number of the users connected to one micro base station exceed this value, the excess users will be connected to the macro base station. The details about the parameters of components in the network will be shown in the chapter 4.

3.2 Problem Description

In real case, the mobile communication network architecture is very complex, but the modeling work doesn't need all components and features in the network. Some of them can be simplified to make the evaluation of the algorithm easier. To enable the network model to evaluate the algorithm easily while still under a relatively realistic scenario, several assumptions are made.

- The modeling work only considers the radio access network.
- It is assumed that the users have the same kinds of applications at time t. Thus, the network traffic load is simplified to the number of users in the network.
- It is assumed that micro base station has 3 levels of transmission power to represent three potential states (switching on/off/ high load).
- The number of users at time t is assumed as a Poisson Distribution with a λ equals to the value of the number of users defined in the OPNET results at time t.
- Not all users move at the next time slot. The number of moved UEs is subject to Poisson Distribution with λ equals to half of the total number of users in the cell at time t.
- When the SINR of the macro cell base station is calculated, the interferences from the users are ignored since they are too subtle.
- It is assumed that switching on a small cell will take 30 seconds, so the simulation is run with 1 minute for each time slot to take the switching delay into consideration.

With the above assumptions, the network model can be simplified while still remained some features in the real case. Based on this model, the algorithm put forward to save energy can be simulated and its performance is evaluated both on power consumption and throughput since the power saving function can influence the network throughput.

3.3 Suggested Solutions

3.3.1 K-means clustering

In the model simulation, k-means clustering algorithm is used for each UE to decide which base station it connects to. The input of this algorithm is a cluster number k and n data objects. The output satisfies the variance minimum standard k clusters. Given an input k, the algorithm divides n data objects into k clusters to make the obtained cluster satisfy the rule that the objects in the same cluster enjoy high similarity while objects in different clusters enjoy low similarity. [16] The main process of the k-means algorithm is illustrated as following:

(1) Choose k objects arbitrarily from n data objects as the initial clustering center;
(2) Calculate the distance between each object and centroid of these objects according to the mean value of each clustering object; and repartition the corresponding object according to the minimum distance;
(3) Recalculate the mean value of each cluster;
(4) Run the loop from (2) to (3) until each cluster doesn't change.

In the modeling process, by using the k-means algorithm, UEs are divided based on their distances to the small cell base station. The number of clusters k equals to the number of small cell base stations in the network. By running the loop until the convergence, all UEs will be clustered to each small cell base stations based on their distance to each small cell base station.

This algorithm is used to simulate the connections of each UE in the network. In the real case, the connection transfer between the UE and base station is done by hard handover. To simplify the model, handover process is simplified by clustering the UEs to decide their serving base station.

3.3.2 Reinforcement Learning Algorithm

In this model, an algorithm is needed to control the switching on/off process of the small cell base station. To meet the requirements, the algorithm need to be the one that can decide the next station based on the current state and the reward gained from the environment. Under these conditions, Reinforcement Learning algorithm is considered.

Reinforcement learning is an algorithm based on the Markov Decision Process (MDP). It can map the situations to actions to get a maximum reward. MDP has an attribute that the next state of the system is only dependent on the current status information and independent on all the states in the past. [17]

A Markov Decision Process can be represented as M = (S, A, P_{sa}, R).

- S represents all the states. $s_i \in S$, s_i represents the statue at the step i;
- A represents all the actions. $a_i \in A$, a_i represents the action at the step i;
- P_{sa} represents the probability of state transition;
- R is the reward function.

With the initial state s_t^0 at time t, the agent will choose an action a_t^0 from A and then transfer to the state a_t^0 with the probability P_{sa}. Then the state vector $s_t = (s_t^0, s_t^1, \dots, s_t^n)$ will be get. Several concepts about the algorithm are shown below.

1. Value Function

Since Reinforcement Learning has the delayed reward, value function is needed to show the long-term influence of the policy π at the current state, which means showing the long-term influence results get from the forward-looking search.

The value function can be written as:

$$V^{\pi}(s) = E_{\pi}[\sum_{i=0}^{\infty} \gamma^i r_i | s_0 = s] \tag{3.1}$$

$V^{\pi}(s)$ means the value function of the policy π at the state s. r_i means the future instant rewards at the step i. γ is the reducing factor and $\gamma \in [0,1]$. It shows the importance of future rewards relative to current rewards. In particular, when $\gamma = 0$, it means only considering the instant reward instead of long-term rewards. When $\gamma = 1$, long-term rewards and instant rewards are considered equivalently important.

The value function can be written as:

$$V^{\pi}(s) = E_{\pi}[r_0 + \gamma r_1 + \gamma^2 r_2 + \gamma^3 r_3 + \cdots | s_0 = s]$$

$$= E_\pi[r_0 + \gamma E[\gamma r_1 + \gamma^2 r_2 + \gamma^3 r_3 + \cdots]|s_0 = s]$$

$$= \mathrm{E}_\pi[r(s'|s,a) + \gamma \mathrm{V}^\pi(s')|s_0 = s] \qquad (3.2)$$

Given the policy π and the initial state s, then under the action $a = \pi(s)$, the state will transfer to the state s' at next step. In this way, the value function can be written as:

$$\mathrm{V}^\pi(\mathrm{s}) = \sum_{i=0} p(s'|s,a)[r(s'|s,a) + \gamma \mathrm{V}^\pi(s')] \qquad (3.3)$$

Define the action value function Q as:

$$Q^\pi(s,a) = E_\pi[\sum_{i=0}^{\infty} \gamma^i r_i | s_0 = s, a_0 = a] \qquad (3.4)$$

With the function (3.3) and given the current state s and current action a, if the function follows the policy π in the future, then the system would transfer to the next state s' at next step. The action value function Q can be overwritten as:

$$\mathrm{Q}^\pi(\mathrm{s},\mathrm{a}) = \sum_{s' \in S} p(s'|s,a)[r(s'|s,a) + \gamma \mathrm{V}^\pi(s')] \qquad (3.5)$$

From the function (3.3) and (3.5), it is not hard to find that the difference between the value function $\mathrm{V}^\pi(\mathrm{s})$ and the action value function $Q^\pi(s,a)$ is that not only the policy π and the initial state s are given, but also the action a is given beforehand.

Having the functions above, the best policy of a Markov Decision Process can be defined as:

$$\pi^* = arg\, max_\pi\, V^\pi(s) \qquad (3.6)$$

It means that given any initial state s, the best policy of a Markov Decision Process is the one that could get the maximum value of the value function.

2. Value Iteration

To solve the function and get the best policy of the Markov Decision Process, Value Iteration is needed. It is an efficient method to shorten the process of calculation, because every round of iteration just sweeps each state once and the best policy will be reached after the convergence. The process of the iteration is that at the $k+1$ round of the iteration, the maximum value of $V^{\pi}(s)$ gained from last round will be passed on to the V_{k+1}. The algorithm only updates the value function every round without storing the policy π.

The whole algorithm can be illustrated as below.

Initialize the vector function as zero

Repeat

$\Delta > 0$

For each $s \in S$:

temp $\leftarrow v(s)$

$v(s) \leftarrow max_a \sum_{s'} p(s'|s,a)[r(s'|s,a) + \gamma v(s')]$

$\Delta \leftarrow \max(\Delta, |\text{temp} - v(s)|)$

Until $\Delta < \theta$ (θ is a small positive number)

Then output the best policy

$\pi(s) = arg\ max_a \sum_{s'} p(s'|s,a)[r(s'|s,a) + \gamma v(s')]$

In this algorithm, estimation is made about the result that when the difference Δ is smaller enough, the result is approximately regarded as the best policy.

3. Temporal-Difference Learning Method

The algorithm shown above, however, need a completely known environment. In this model, the environment cannot be known beforehand since the controlling process is instant. Thus, an algorithm is needed who does not need the environment model and will not be limited by the episode task. It should be able to finish the continuous task. Therefore, based on these conditions, Temporal-Difference learning algorithm is considered. It is an algorithm that are based on the Dynamic Programming methods and Monte Carlo methods. [18]

In the Monte Carlo methods, the value function is updated as

$$V(s_t) \leftarrow V(s_t) + \alpha[R_t - V(s_t)] \tag{3.7}$$

In the Temporal-Difference learning, since the environment is unknown, it is not possible to get the probability of state transition, not to mention to use the function (3.3). In this way, function (3.7) need to be changed. The reward has to be replaced by the instant reward and value function of the next step.

$$V(s_t) \leftarrow V(s_t) + \alpha[r_{t+1} + \gamma V(s_{t+1}) - V(s_t)] \tag{3.8}$$

Then the whole algorithm can be illustrated as below.

Repeat (for each episode)

Initialize the vector function as zero

For each $s \in S$:

a $\leftarrow$ *action taken under the policy* π

$V(s) \leftarrow V(s) + \alpha[r + \gamma V(s') - V(s)]$

s $\leftarrow s'$

Until all $V(s)$ converge

Then output the best policy $\pi(s)$

There are two kinds of algorithms in Temporal-Difference learning algorithm, one is the on-policy SARSA Algorithm, the other is off-policy Q-Learning Algorithm. Both of them are based on the action value function Q instead of the value function. In on-policy algorithm, the policy followed by selecting actions and the one followed by updating action value function are the same. While in off-policy algorithm, the policy followed by updating the action value function is different than selecting the actions.

SARSA Algorithm predicts the action value function Q. It is named as SARSA because each update of the action value function Q depends on $(s_t, a_t, r_{t+1}, s_{t+1}, a_{t+1})$ at time t. The algorithm can be demonstrated as shown below.

Initialize $Q(s, a)$

Repeat (for each episode)

Initialize S

Choose A from S using policy from Q

Repeat (for each episode)

Take action A, get R, S'

Choose A' from S' using policy from Q

$Q(S, A) \leftarrow Q(S, A) + \alpha[R + \gamma Q(S', A') - Q(S, A)]$

$S \leftarrow S'; A \leftarrow A'$;

Until convergence

The difference of using the Q-Learning algorithm is that the update of action value function is independent on the policy of action selection. "Q" means the function that returns the reward of each action taken which shows the "quality" of the action. The function of updating the action value function can be written as follow:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_{t+1} + \gamma max Q(s_{t+1}, a) - Q(s_t, a_t)] \tag{3.9}$$

Then the whole algorithm can be illustrated as below.

Repeat (for each episode)
  Initialize S
  For each $\boldsymbol{s} \in \boldsymbol{S}$:
    a $\leftarrow$ *action taken under the policy* π derived from Q
    $Q(S, A) \leftarrow Q(S, A) + \alpha[R + \gamma max Q(S', a) - Q(S, A)]$
    $S \leftarrow S'$
Until convergence

More details about the performances of SARSA Algorithm and Q-Learning Algorithm will be discussed later in the results section.

3.3.3 OPNET simulation

OPNET Modeler is a software that can establish the existing network model. It is now an industry-leading network development tool that can be used to design and study communications networks, devices, protocols, and applications with much flexibility. Modeler provides developers with an integrated environment for modeling, simulation, and analysis, greatly reducing the workload of programming and data analysis. It is now used by major companies and organizations around the world to promote the development process. [19]

The object-oriented modeling method and graphical editor of the software can reflect the structure of the actual network and network components, so the actual system can be intuitively mapped into the model. In this way, the model and simulation could be easily modified, which is suitable for analyzing and predicting the network performance. Also, OPNET Modeler can be used to model various network parameters like network load, network delay, throughput, error rate etc. In this case, it will be helpful in analyzing the network through the graphic in a visual way.

In this project, to make the data more realistic, simulation of a network has been made by using the software OPNET Modeler. the main process is: create a new project and office scenario, deploy a wireless network in the scenario, add new components, configure the applications and profiles, configure the mobile workstations and define trajectory, configure and run the simulation.

Since small cells can be deployed in 4G network, LTE network architecture has been built by using the OPNET modeler to simulate the daily network traffic as a reference for further modeling simulation. The office scenario is initialized in the OPNET simulation and the network is deployed as shown in figure 3.1.

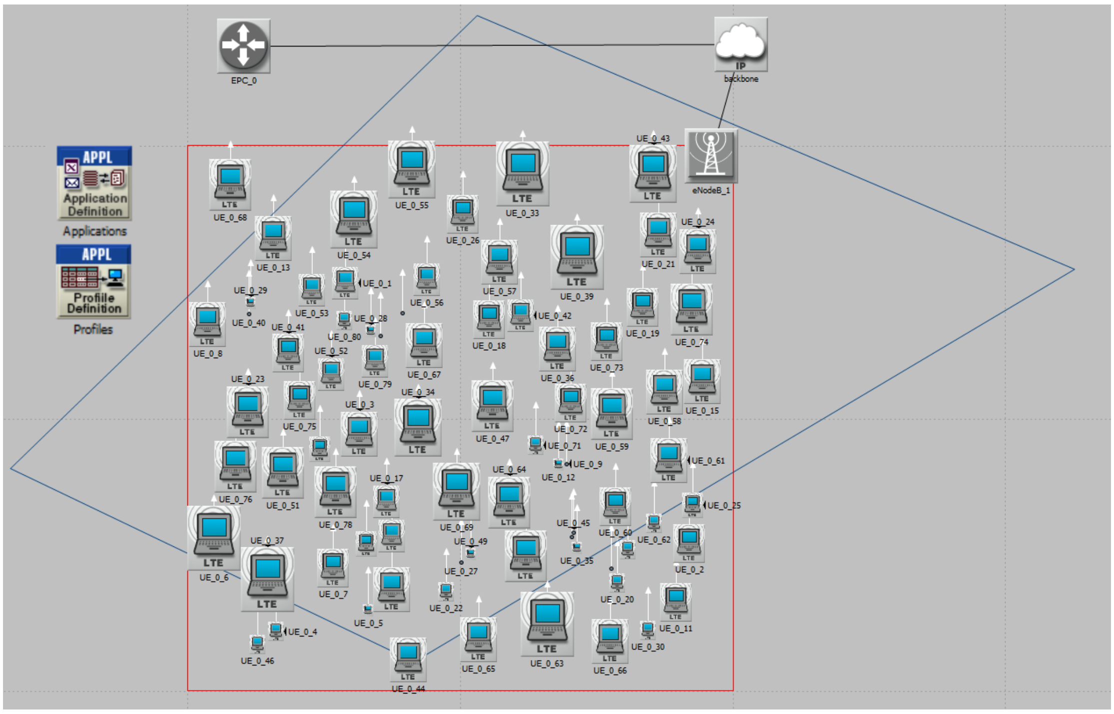

Figure 3.1 LTE Network in Office Scenario

The network structure mainly consists of several components: eNodeB, network backbone, EPC, LTE workstations. The network backbone can tie together different networks. The Application and Profiles components are used for traffic configuration in the network. These components build a macro cell and can offer data and voice service for the users in the cell. In this network, 80 users are configured and move in and out of the network along the time. More details about the network configuration will be discussed later in the section 4.

To illustrate this network architecture in details, figure 3.2 shows the network architecture of an LTE network. [20]

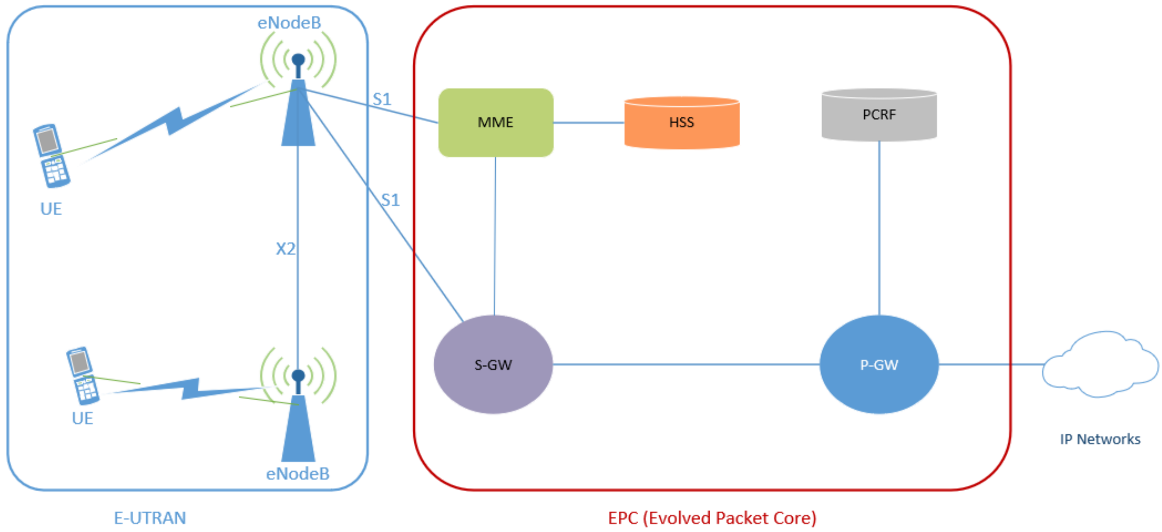

Figure 3.2 LTE Architecture

The LTE architecture mainly consists of two parts: E-UTRAN and EPC. E-UTRAN stands for Evolved UTRAN (UMTS Terrestrial Radio Access Network) which evolves from the 3G network architecture. EPC stands for Evolved Packet Core which is LTE core network.

E-UTRAN has the components of UEs (User Equipment) and eNodeBs (Evolved Node B). UEs are users' mobile devices and eNodeBs are the base stations who also do the RNC job in the previous technology. The eNodeB connects the UEs to the core network and it is responsible for radio resource management.

EPC has the components that build up the core network. MME is the Mobility Management Entity whose work is mobility management as well as authentication and Security. S-GW is the Serving Gateway and P-GW is the Packet Data Network Gateway. S-GW handles all user plane switching and data forwarding. It can get access to the external IP networks via P-GW. PCRF is used for the policy and charging resource function. HSS is Home Subscription Server who is a security database and stores the subscriber data.

In this project, nothing need to be done with the core network, all the controlling process is worked on the E-UTRAN and make some improvements on the radio resources management. The small cells are all deployed in this part and the small cell base stations are all connected to the core network as eNodeB in the figure. Thus, in the simulation network, the core network can be simplified as one single part EPC to do the functions. All the changes need to be simulated are all in the E-UTRAN part.

Chapter 4 System Model

In the real case, the network structure is very complex and the transmission of information can get various parts of the communication network structure involved in. However, in this project, controlling the cells to switch on/off won't involve in all the parts like the backhaul network or the physical layer structure. In this way, a model need to be built to simplify the network structure and simulate the parts needed in evaluating the algorithm efficiency.

4.1 Network Structure Model

4.1.1 Macro cell base station model

In this model, since the network is considered with only one macro cell, only one macro base station is needed to be considered.

Commonly, a cell is modeled as hexagon-shaped, but the modeling work is based on MATLAB programming, so the macro cell is simplified as a 3000*3000 meters square. The macro base station has a fixed location and is defined located at the edge of the cell with the coordinate (0,0). The parameters of the base station are initialized as listed in table 4.1.

Table 4.1 Parameters of Base Station

Parameter	**Value**
Power (dBm)	52
Bandwidth (MHz)	20
Location (m)	(0,0)

4.1.2 Small cell base station model

Micro base stations are modeled to create small cells in a macro cell. Different from the macro base station, the micro base stations are randomly scattered in the macro cell. The power of the micro base station is also smaller than the macro base station. The number of the micro base station can be different in different kinds of areas. In the hot spot area which has a large data traffic load, more small cells need to be deployed to meet the capacity requirements. While in those urban areas where the buildings are dense, some small cells need to be placed to ensure the coverage. In this project, office scenario is researched so the number of small cells should meet the demands of capacity. However, to reduce the simulation time, the number of the small cell base stations placed in the cell is 16 and the type of the small cell base station is defined as Femto base station.

To simulate the switching of the small cell base stations, three levels of power are defined to simulate different working load of the base station. When the power is -20 dBm, the small base station is deemed shut down since the power is near zero. While the power reaches its maximum value 39 dBm, the small cell base station is deemed as working at a high load. The middle level 15 dBm is deemed as a threshold value to keep the small cell base station awake. In the simulation, the power of each small cell base station is initialized by one of the power levels randomly.

The switching problem should also be considered. As shown in figure 4.1, it takes time to switch on a small cell base station until it fully works. It also costs extra power to start up a small cell base station.

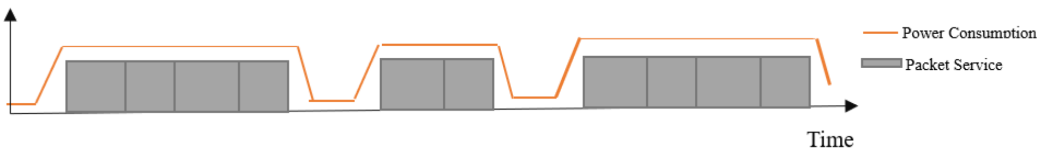

Figure 4.1 Time Delay During the Switching

Simulation parameters of small cell base stations are initialized in the model as listed in table 4.2.

Table 4.2 Parameters of Small Cell Base Station

Parameter	**Value**
Type	Femtocell base station
Number	10
Power (dBm)	-20, 10, 20
Bandwidth (MHz)	20
Location (m)	Randomly scattered in the macro cell
Awaken Power (w)	15
Start-up power (w)	30
Time duration to switch on a small cell (sec)	30

4.1.3 Mobile user model

In the cell, several users need to be modeled. Since the project focuses on the office scenario, the number of UEs (User Equipment) should vary during the day. For example, at midnight, the number of users will be small since people go back home and won't need the data service in the area. While during the working hour (from 8 a.m. to 5 p.m.), the data traffic could be large and varies along with the time. The traffic load is referred from the business curves in figure 4.2. [21]

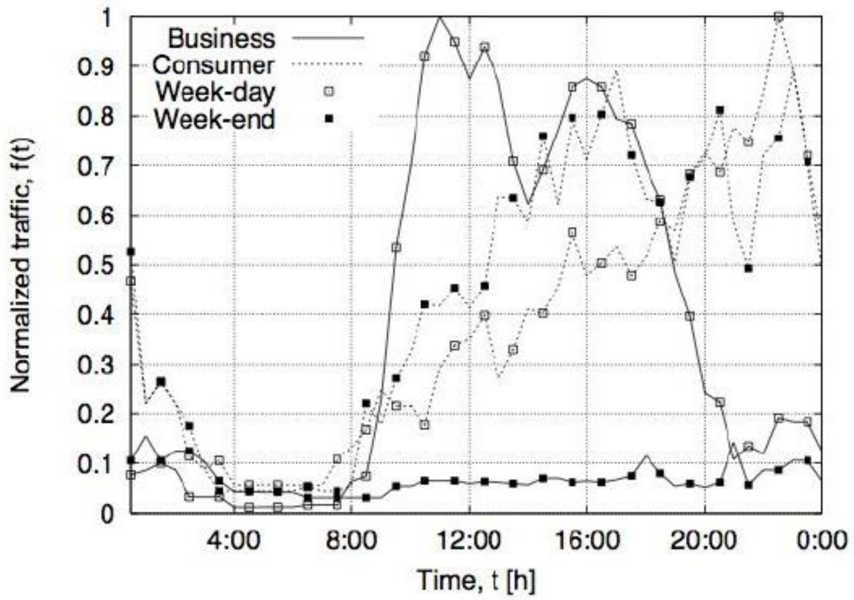

Figure 4.2 Daily Traffic Profile

The traffic load is considered as the number of users in this model. The more users in the cell, the higher traffic of the cell. The number of users is based on the simulation results shown in the next section and it is subject to Poisson Distribution. The locations of the UEs are randomly scattered. To make the simulation close to the real case, some of the UEs move to another place within the range of 500 meters at the next time slot while others stay where they are. The number of moved UEs is subject to Poisson Distribution with λ equals to half of the total number of users in the cell at time t. The distributions of UEs at time 7:00 a.m. and 7:01 a.m. are shown in figure 4.3.

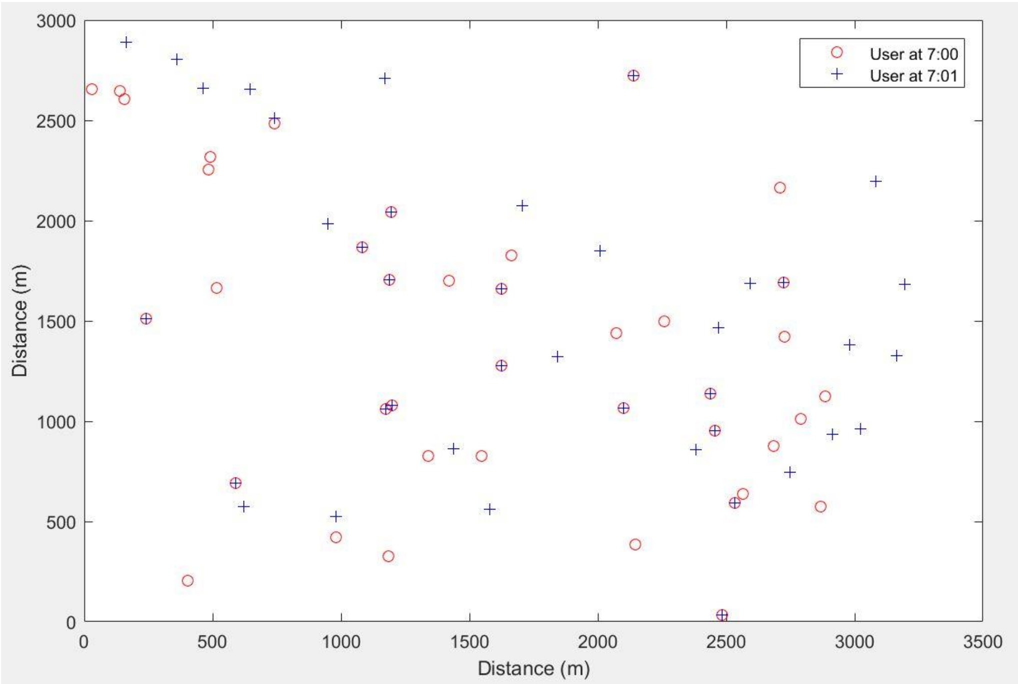

Figure 4.3 Distributions of UEs at 7:00 a.m. and 7:01 a.m.

To decide which base station the UE is connected to, the k-means Clustering algorithm is used. The UEs are clustered based on their distance to each base station and the UE will be connected to the base station it is most close to. The pseudo code of this process is shown below.

Algorithm 1 Pseudo code of k-means Clustering

Input the number and location of both UEs and small cell base stations
Initialize the central points of all base stations
For each UE
    Calculate the distances to all base stations.
    Find the minimum value and record the index of the base station
    For each index
        Count the number of UEs that has that index
        Calculate the location of new center point
    Record the number of UEs in a vector as well as all the corresponding location
Output the vector and the locations

Since a Femtocell base station can only support 4-6 users, after k-means clustering, those excess users have to be connected to the macro cell base station. By far, the initialization of the cell structure has been finished. Figure 4.3 shows the whole scenario of the cell at time 1:40 a.m. In this figure, different color of users shows they are in different small cells. The users in red circle means they are connected to the macro base station. It is not hard to find that some small cells only have one or two users, as marked in green square, these small cells can be switched off to save energy and the users they support can be switched to the macro base station.

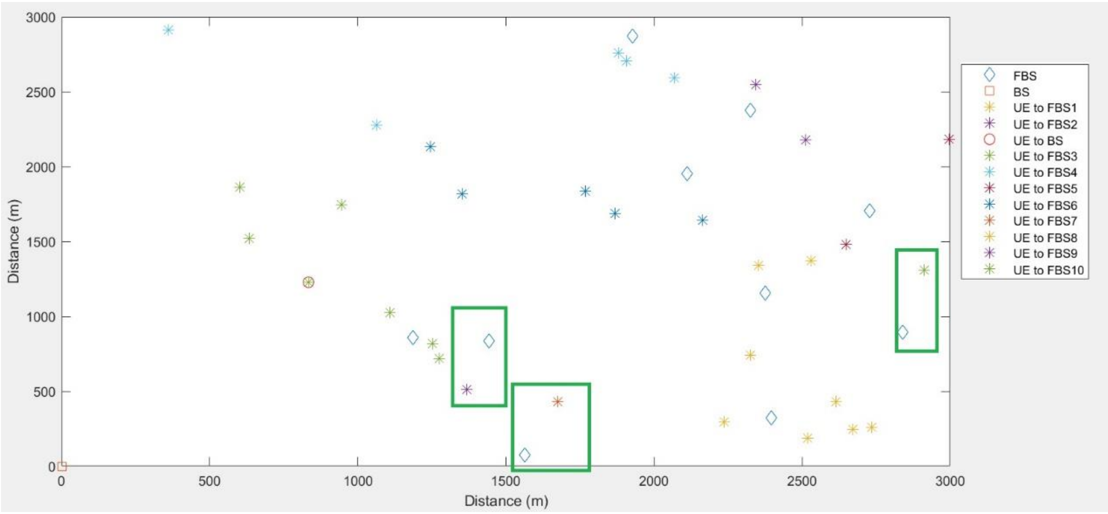

Figure 4.3 Initialization of the Cell at 1:40 a.m.

4.2 Throughput and Power Consumption

The calculation of the power consumption is based on the throughput of the base stations. While the calculation of throughput is based on the SINR of the base stations. The details about the calculation will be illustrated in this part. They will be discussed from the aspect of macro cell base station and the small cell base station separately.

4.2.1 Parameters of macro cell

1. SINR

The calculation of SINR need to take both the transmission power and interference into consideration. The function can be written as (4.1).

$$SINR = \frac{P_{Tx\cdot BS} \cdot G \cdot L}{\sum_i S(i) + \sum_j S(j) + N} \tag{4.1}$$

$$S(i) = P_{Tx\cdot BS} \cdot G_i \cdot L \cdot PL_{IJ}$$

$$S(j) = P_{Tx\cdot FBS} \cdot G_j \cdot L \cdot PL_{IJ}$$

where $P_{Tx\cdot BS}$ and $P_{Tx\cdot FBS}$ are the transmission power of macro base station and Femtocell base station, respectively. L is the penetration loss. G is the shadow fading.

$\sum_j S(j)$ is referred as the total interference to the macro base station. N represents the total noise during the transmission. The interference is considered as the interference between the macro cell base station and small cell base station. In the calculation of the interference, the path loss PL need to calculated due to the power loss during the transmission. The function of the path loss is shown as (4.2).

$$PL = 32.4 + 20 \cdot log_{10}(f) + 20 \cdot log_{10}(d) \quad (4.2)$$

where d is the distance between the Femtocell base station and the macro base station, f is the channel bandwidth frequency.

2. Throughput

The calculation of throughput is based on the result of SINR. By using the channel capacity concept in Shannon Theorem, the throughput of the macro cell can be estimated as function (4.3).

$$C = BW \cdot log_2(1 + SINR) \quad (4.3)$$

Besides the Femtocell base stations, those UEs that exceed the upper bound of the Femtocell station's capacity will be connected to the macro base station and will also be calculated in the throughput of the macro cell.

3. Power Consumption

The relationship between the power consumption and the throughput is regarded as a linear one. The function is written as (4.4).

$$P_{macro} = P_0 + \beta \cdot \rho \quad (4.4)$$

where P_0 is the baseline power consumption of a macro cell and ρ is the normalized throughput and $\rho \in [0,1]$. The factor β is the transmission power of the base station.

4.2.2 Parameters of small cell

1. SINR

Like the SINR calculation in macro base station, the SINR calculation also need to consider the transmission interference and the noise. For the small cell base station, the interference on the transmission between the user and Femto base station is from the macro base station, other Femtocell base stations and other users. [22] The function is shown as (4.5).

$$SINR = \frac{S(p)}{\sum_p S(p) + \sum_q S(q) + \sum_m S(m) + N} \tag{4.5}$$

$$S(p) = P_{Tx \cdot FBS} \cdot G_p \cdot L \cdot PL_{PQ}$$

$$S(q) = P_{Tx \cdot BS} \cdot G_q \cdot L \cdot PL_{PQ}$$

$$S(m) = P_{Tx \cdot UE} \cdot G_m \cdot L \cdot PL_{PM}$$

where $P_{Tx \cdot UE}$ is the transmission power of Femtocell base station per user. The path loss is calculated as function (4.6).

$$PL = 32.4 + 20 \cdot log_{10}(f) + 20 \cdot log_{10}(d) \tag{4.6}$$

The element d is the distance between the Femtocell base station and the targeted object.

Meanwhile, the case when the Femtocell base station is off is also taken into consideration. If the transmission power of the Femtocell base station is a negative value, it is deemed as shutting down and its SINR will be defined as -3 dB automatically.

2. Throughput

The calculation of the throughput in micro cell also uses the channel capacity concept in Shannon Theorem as the calculation for macro cell in function (4.3).

3. Power Consumption

The power consumption of one single small cell is also regarded as a linear function depending on the normalized throughput. The function is shown as (4.7).

$$P_{micro} = P_0 + \beta \cdot \rho \tag{4.7}$$

P_0 is the baseline power consumption of the small cell and will be initialized before the simulation. The factor β is the transmission power of the Femtocell base station.

The total power consumption is the sum of the macro cell power consumption and the small cell power consumption. Besides, the power consumption of starting up a

Femtocell base station also need to be considered. In this way, the switching actions are counted every time to record the number of Femtocell base stations turned on. Therefore, the total power consumption also includes power used to start up a small cell at time t. For the small cell which is switched on, its power consumption can be written as (4.8).

$$P_{micro(off\rightarrow on)} = P_0 + \beta \cdot \rho + P_{start-up} \tag{4.8}$$

4.3 Controlling Process

To increase the power efficiency, small cell base stations need to be turned down when the network traffic load is relative low. The switching on/off decision is made based on the Reinforcement Learning algorithm as discussed in the previous sector.

In the algorithm, each small cell base station is regarded as an agent. The switching on/off decision is a kind of policy. Each agent will lean the policy independently through the real-time interaction with the environment. Every time when the agent "tries" an action, a reward value will be returned from the environment. The Q-value of each agent i is $Q(x_t^i, a_t^i)$ which only depends on the state x_t^i and action a_t^i which is relates to the current time t. As the traffic load varies in the environment, the states of each agent are different. The decision-making process about these states is based on the Markov Decision Process with states vector $x_t = (x_t^1, x_t^2, \dots, x_t^n)$ .The whole process is: at the time t, agent i randomly chooses a state x_t^i from the state set and make the action a_t^i. Then at the time $t+1$, the agent receives the reward r_t^i from the environment and the Q-value $Q(x_t^i, a_t^i)$ is then updated. The agent will choose the state with the highest Q-value to be the next state. The process can be illustrated by the graph as shown in figure 4.4.

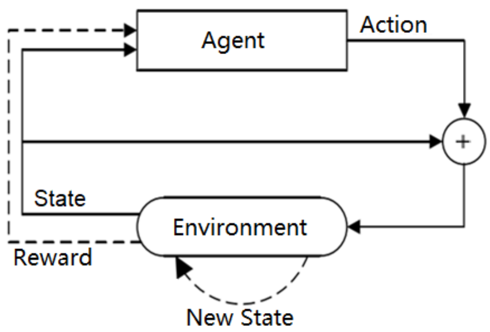

Figure 4.4 Learning Process

The above process can be presented as:

$$Q(S,A) \leftarrow Q(S,A) + \alpha[R + \gamma max Q(S', a) - Q(S,A)] \qquad (4.9)$$

where γ is the reducing factor and $\gamma \in [0,1]$. It shows the importance of future rewards relative to current rewards. α is the learning rate and $\alpha \in [0,1]$. S is the set of states and A is the set of the action. R is the reward.

In the system model, the state $x_t^i = P_{tx_t}^{\ i}$ of the agent i at the time t depends on the transmission power level $P_{tx_t}^{\ i}$ of the small cell base station. The reward value r_t^i is discussed based on the drop rate D_t^i and the state of agent i, which is defined as:

$$r_t^i = \begin{cases} 0, & D_t^i > D_{th} \\ P_0 + P_{start-up}, & D_t^i \le D_{th}\ and\ i\ is\ on \\ P_t^i\ , & D_t^i \le D_{th}\ and\ i\ is\ off \end{cases} \qquad (4.10)$$

When the drop rate is too high, the reward value will be set to 0 to force the small cell i to be switched off. When the small cell is on, the reward value is the power consumption of small cell i switching on at time t to help the offloading from macro base station. When $P_t^i > P_0 + P_{start-up}$, it means the throughput of the small cell gets pretty high which is caused by the small number of users in the cell. The small cell will then be switched off to save energy.

The algorithm is trained through the time of a day and get a set of Q-values. The programming process can be simplified to a pseudo code as followed:

Algorithm 2 Pseudo code of learning process
1: **procedure Training** $(Q(x_t^i, a_t^i))$
2: $Q(x_t^i, a_t^i) \leftarrow 0$
3: **while do** $t \in D$
4: **run Q-learning**
5: $Q(S,A) \leftarrow Q(S,A) + \alpha[R + \gamma max Q(S', a) - Q(S,A)]$
6: $QTable \leftarrow Q(x_t^i, a_t^i)$
7: $S \leftarrow S'$
8: **end while**
9: **end procedure**
where D is the set of minutes in a day

The whole learning process can be illustrated as a flow chart below.

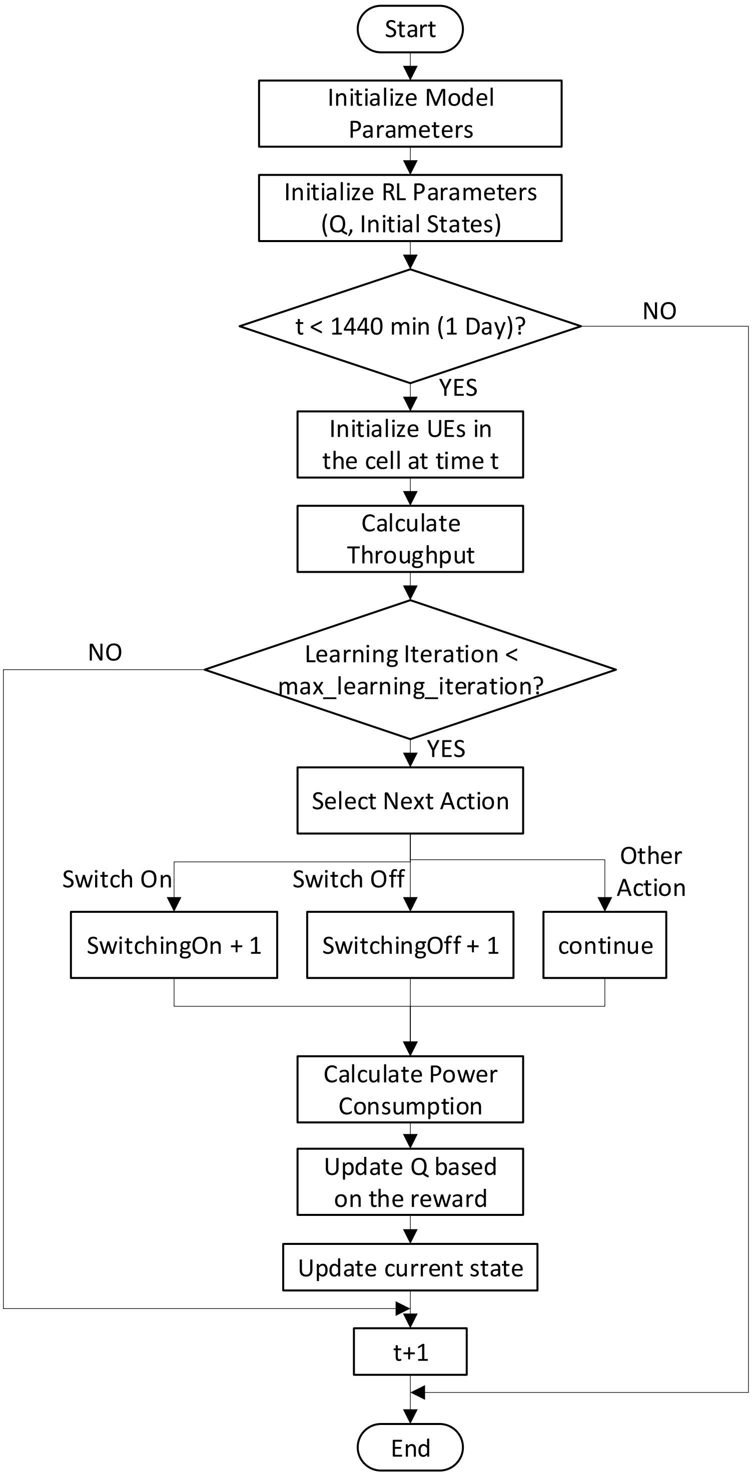

Figure 4.6 Flow Chart of Learning Process

Chapter 5 Simulation Results

5.1 OPNET Simulation Results

Since OPNET is a software that can simulate the communication network, an LTE network is deployed in this software to get a reference about the simulation results of the traffic load in the macro cell. The LTE network is initialized by using the function "Deploy Wireless Network" in the software. The initial setting is as shown in figure 5.1.

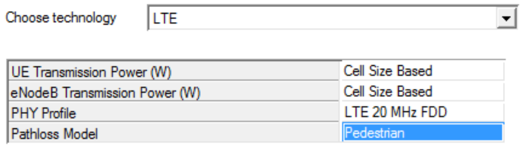

Figure 5.1 Profile Settings of LTE Network

Having the initial network been set, the applications and profile files need to be set then. Since this project consider the office scenario, applications relevant to office working are considered. Two rows are set in the profiles, one is staff and the other is user. They are set in this way because customers or other users will step into this office area and use the network. User will not be able to get access to some of the applications in the office area, so the application settings in these two profiles are different. Figure 5.2 shows the profile settings.

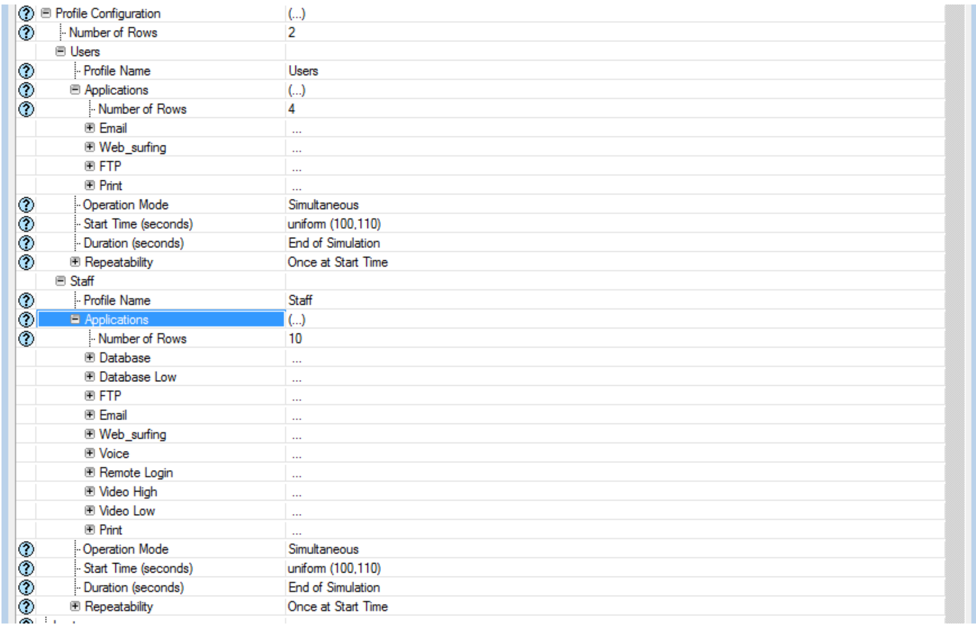

Figure 5.2 Profile Settings

Application details are shown in figure 5.3. The applications include database, web surfing, FTP, print, video, voice and remote login and they are commonly usedi.

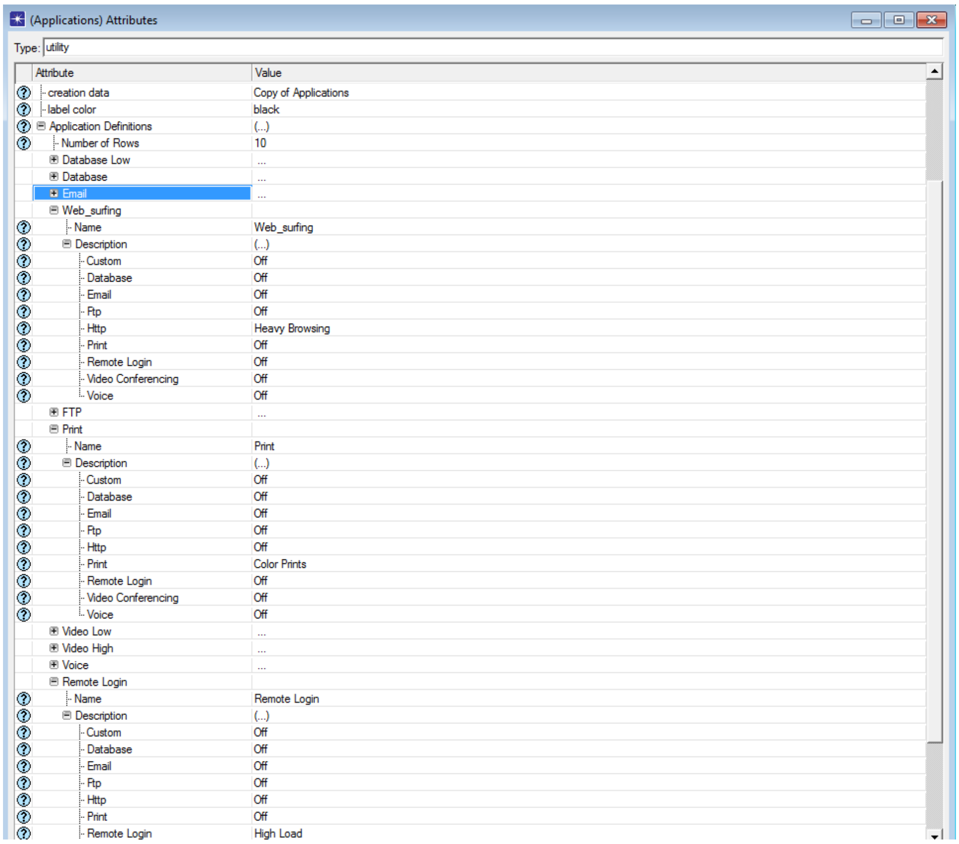

Figure 5.3 Application Settings

After the settings, simulation is run with the parameter setting as 1-day duration and 128 seeds. The result of the traffic load of the cell is shown in figure 5.4.

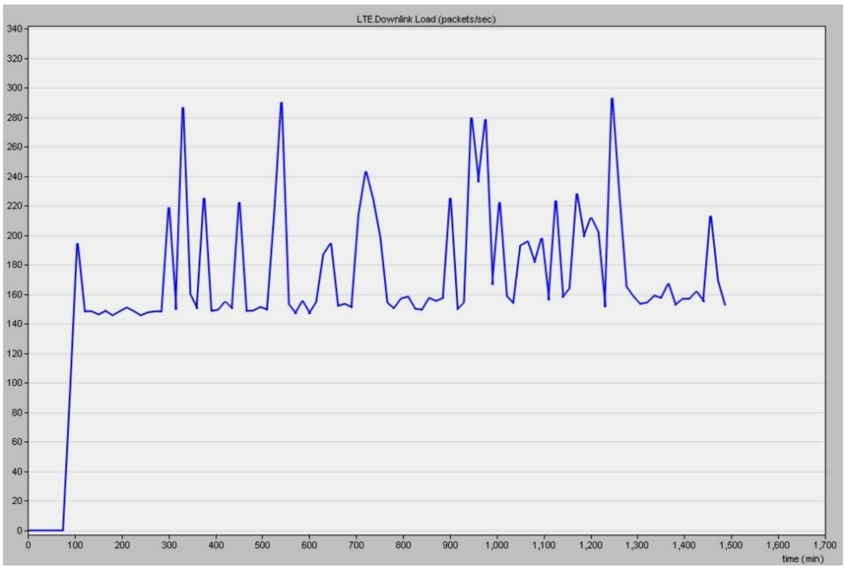

Figure 5.4 Traffic Load Result (packet/sec)

The figure shows the peak hours of a work day in office scenario, which is basically in accordance with the reference graph in figure 4.2. The subsequent simulation model on MATLAB will be based on this result.

5.2 MATLAB Simulation Results

5.2.1 Simulation settings

To validate the model built in previous section, MATLAB simulation has been made. The components in the network include Femtocell base stations (FBS) and macro base station (BS). Table 5.1 shows the simulation parameters of these components.

Table 5.1 Simulation Parameters

Parameter	Value
Number of FBS	10
Distance of Area	3 km × 3 km
Baseline Power of BS (w)	750
Transmission Power of BS (w)	600
Antenna Gain of BS (dBi)	0
Cable Loss of BS (dBi)	0
Noise Figure of BS (dB)	6
Baseline Power of FBS (w)	105.6
Transmission Power of FBS (w)	39
Antenna Gain of FBS (dBi)	0
Cable Loss of FBS (dBi)	0
Noise Figure of FBS (dB)	4
Transmission Power of FBS per UE (w)	0.5
Antenna Gain of FBS (dBi)	0
Cable Loss of FBS (dBi)	0
Noise Figure of FBS (dB)	8

5.2.2 Algorithm performance evaluation

1. On-policy SARSA Algorithm vs. off-policy Q-Learning Algorithm

To determine which algorithm performing better in the controlling process, comparison of the performance in energy saving efficiency has been made. These two algorithms have the same reward function, but their methods to update the Q value are different.

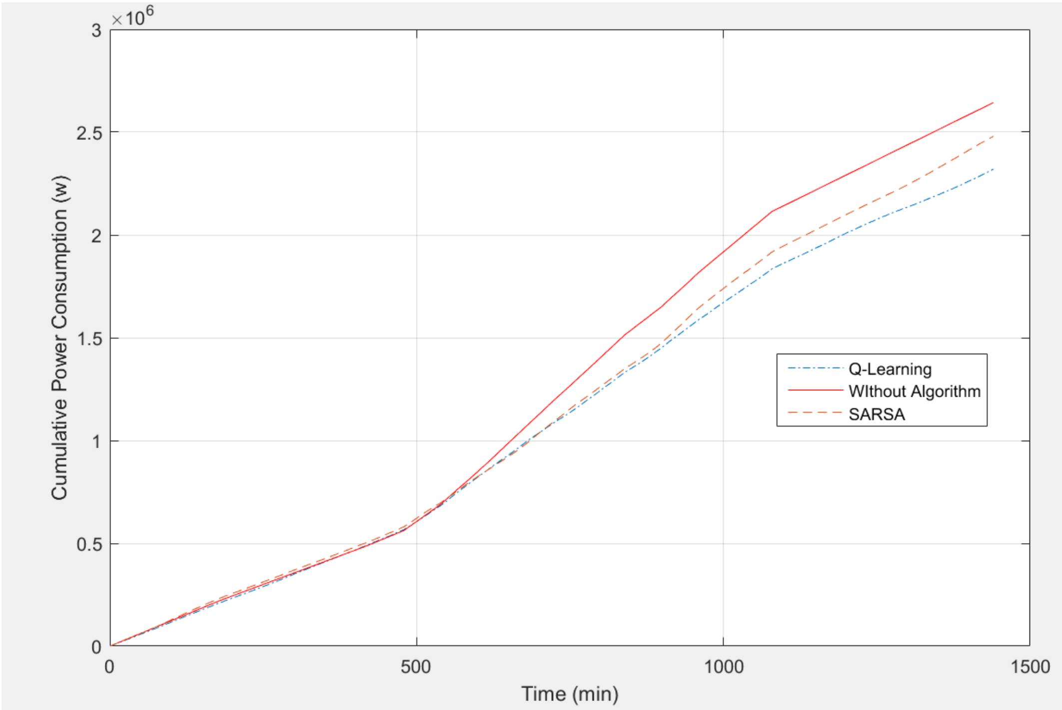

Figure 5.5 SARSA Algorithm vs. Q-Learning Algorithm

In figure 5.5, at initial stage when the traffic load is fair, the performances of both algorithms are equivalent. While during the peak hours when traffic load in the network get higher, Q-Learning shows a better performance compared with the SARSA.

2. Performance evaluation of reduction factor γ

γ ranges from 0 to 1. $\gamma = 0$ means only instant reward matters, while $\gamma = 1$ means the equivalent importance. Since γ shows the importance of future rewards, simulation has been made to determine its value to set. This simulation is run based on the α value of 0.5.

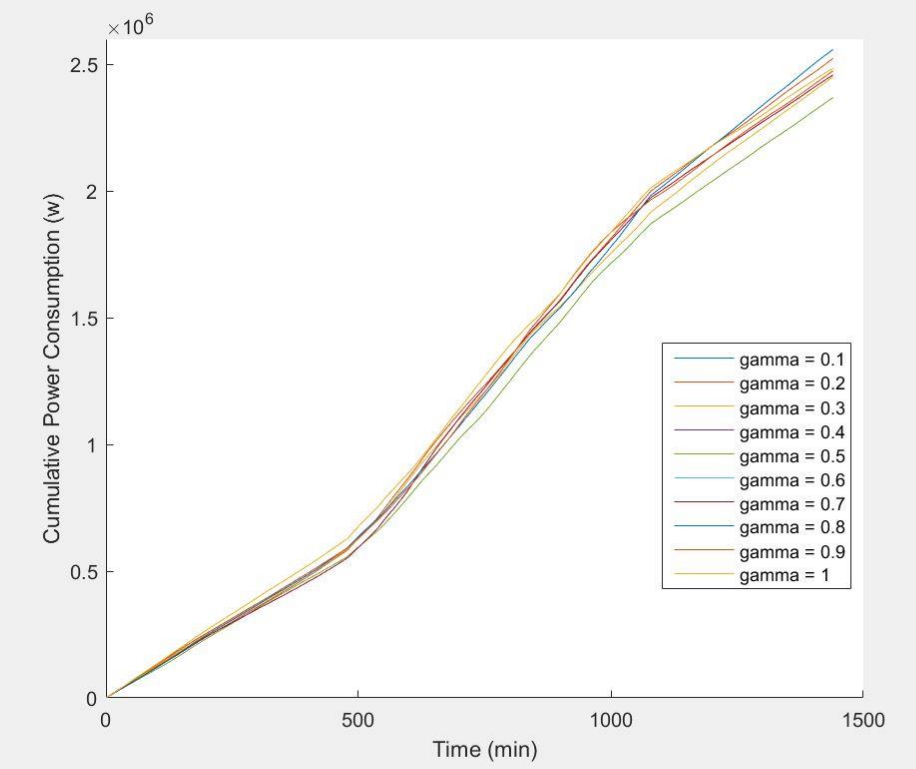

Figure 5.6 Performance Evaluation of γ

From figure 5.6, it's not hard to find that the cumulative power consumption is the lowest when $\gamma = 0.5$. Especially during peak hours in daytime, this value shows better power-saving performance.

3. Performance evaluation of learning rate α

The learning rate α has the range from 0 to 1. As α gets larger, the update of Q value will depend less on past experience. Since α shows the weight about the past experience during the learning process, simulation has been made to determine its value to set. This simulation is run based on the γ value of 0.5.

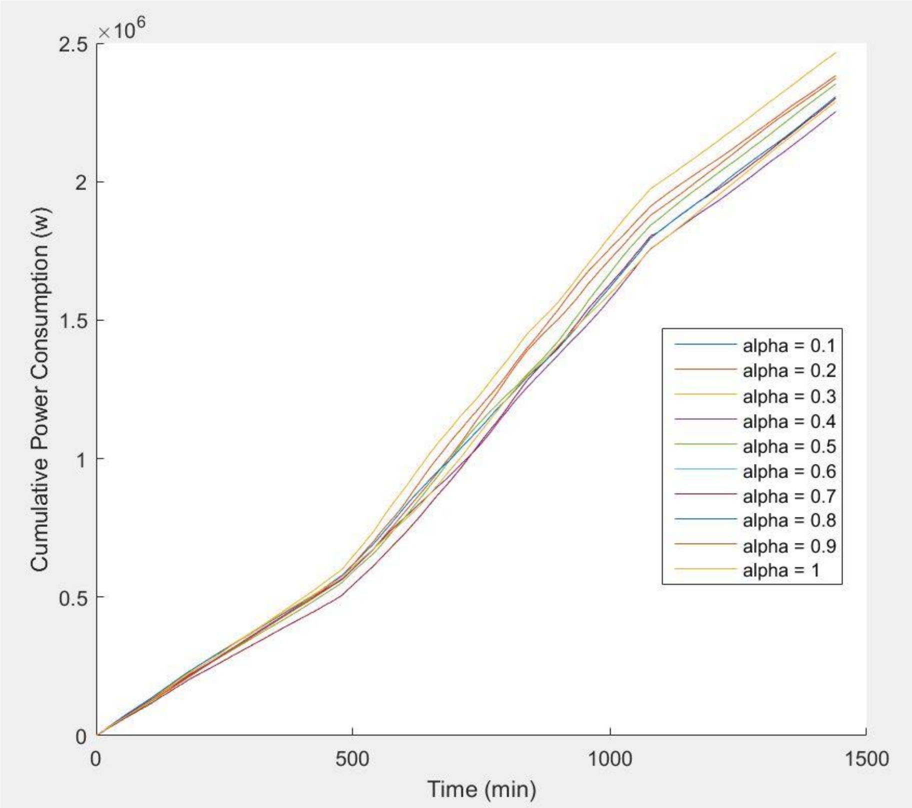

Figure 5.7 Performance Evaluation of α

The figure shows that $\alpha = 0.4$ shows the best performance specially during the hours when the traffic load is not that large.

Based on the results listed above, Q-Learning algorithm is chosen as a better algorithm for the controlling process and the algorithm is simulated with a learning rate $\alpha = 0.4$ and a reduction factor $\gamma = 0.5$.

5.2.3 Network performance evaluation

1. Power Efficiency Performance

The power consumption is plotted with the duration of 24 hours. Each time slot is a minute. The curve is compared with the one that without switching algorithm and have all the small cells turning on to meet the capacity demands of peak hours in the office scenario.

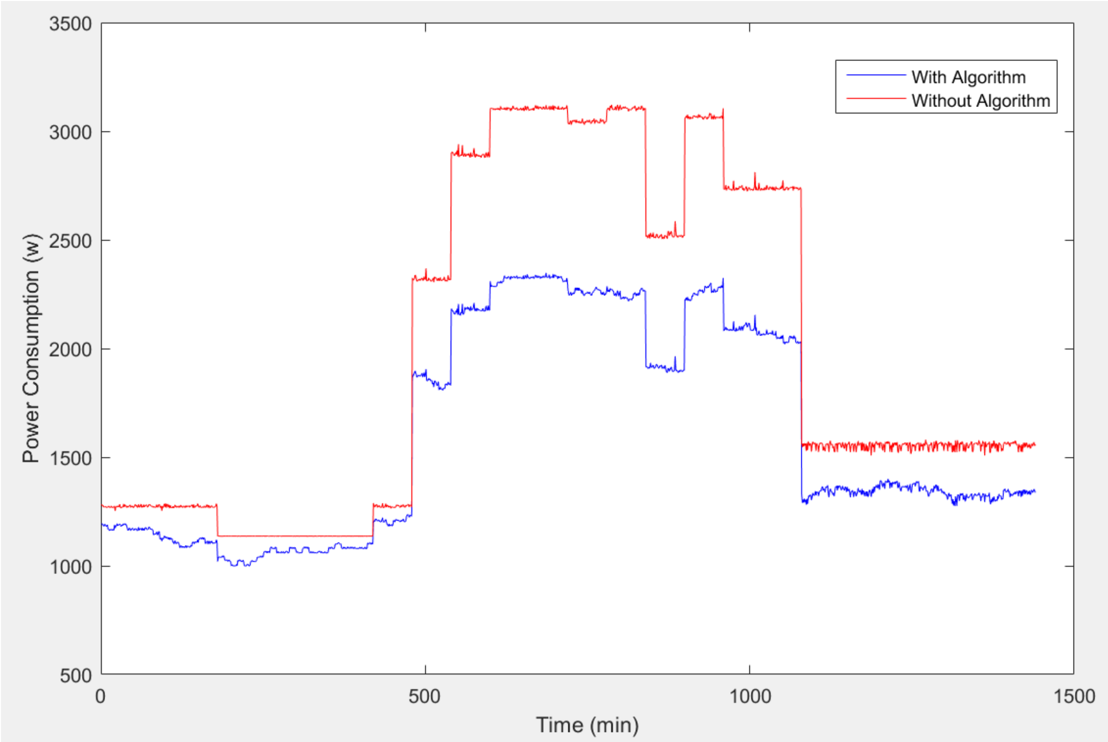

Figure 5.8 Power Consumption with Algorithm vs. without Algorithm

Figure 5.8 shows the details about power consumption in a macro cell before and after using the Q-Learning algorithm. The simulation duration is one day and each iteration is one minute. In general, the power consumption with the algorithm is less than without algorithm. In details, every time when the traffic load in the network increase, which leads to the increment of power consumption, power consumption with the algorithm will increase first and then drop until reach a relatively stable state. This shows the effect of switching. The reward will have delay on action selection. Despite the good performance during the peak hours, the performance of the off-peak hours is not that remarkable. This is because the large proportion of the power consumption comes from macro cell base station, which makes the power saving from Femtocells is relatively subtle to be seen.

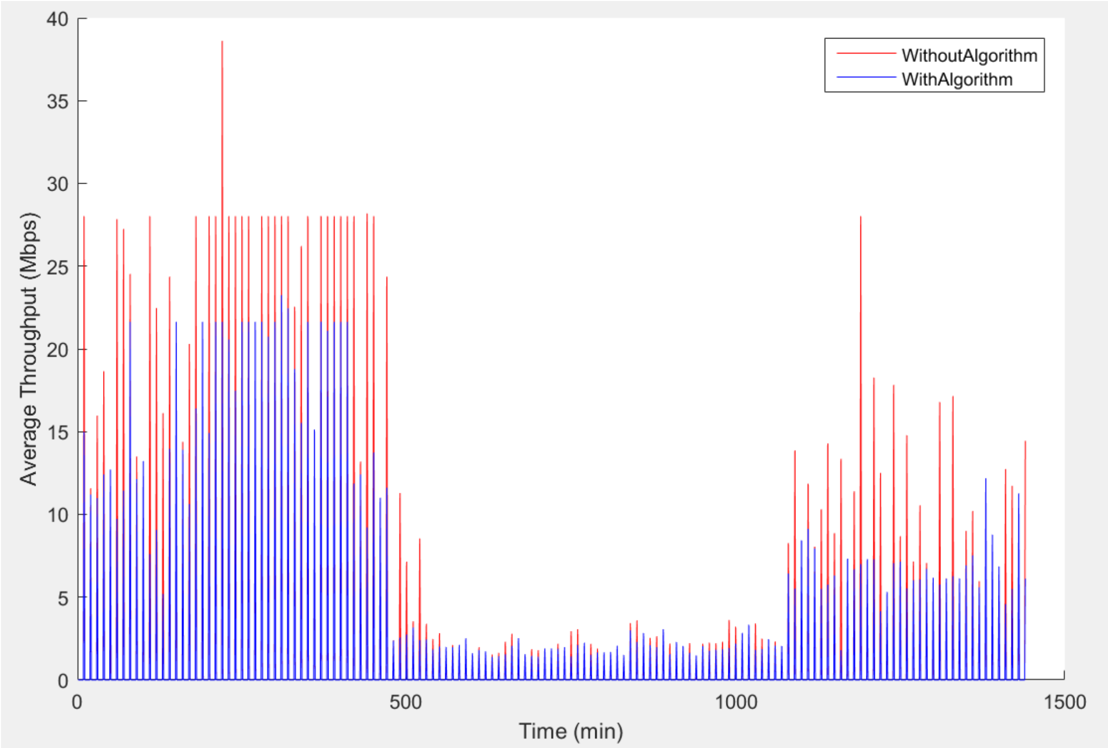

Figure 5.9 Average Throughput in Macro Cell with Algorithm vs. without Algorithm (1 Day)

Figure 5.9 shows the comparison of average throughput in the macro cell between using the algorithm and not using the algorithm. Combining the result in figure 5.8, it is not hard to find that the average throughputs are similar during the peak hours since all small cells are switched on. While during those hours when the traffic loads are relatively low, the average throughputs are lower when the algorithm is used, i.e. some small cells are switching off to save energy. It is because by using the algorithm, some UEs have to be connected to the macro base station and add more traffic loads to BS. The throughput experienced by each user will decrease if many users share the resource. This helps illustrate the tradeoff between the higher throughput and lower power consumption. If better services are offered, more power will be consumed. However, it is not necessary to offer the highest throughput, the value within a sustainable range will be fine to save energy and offer service users needed.

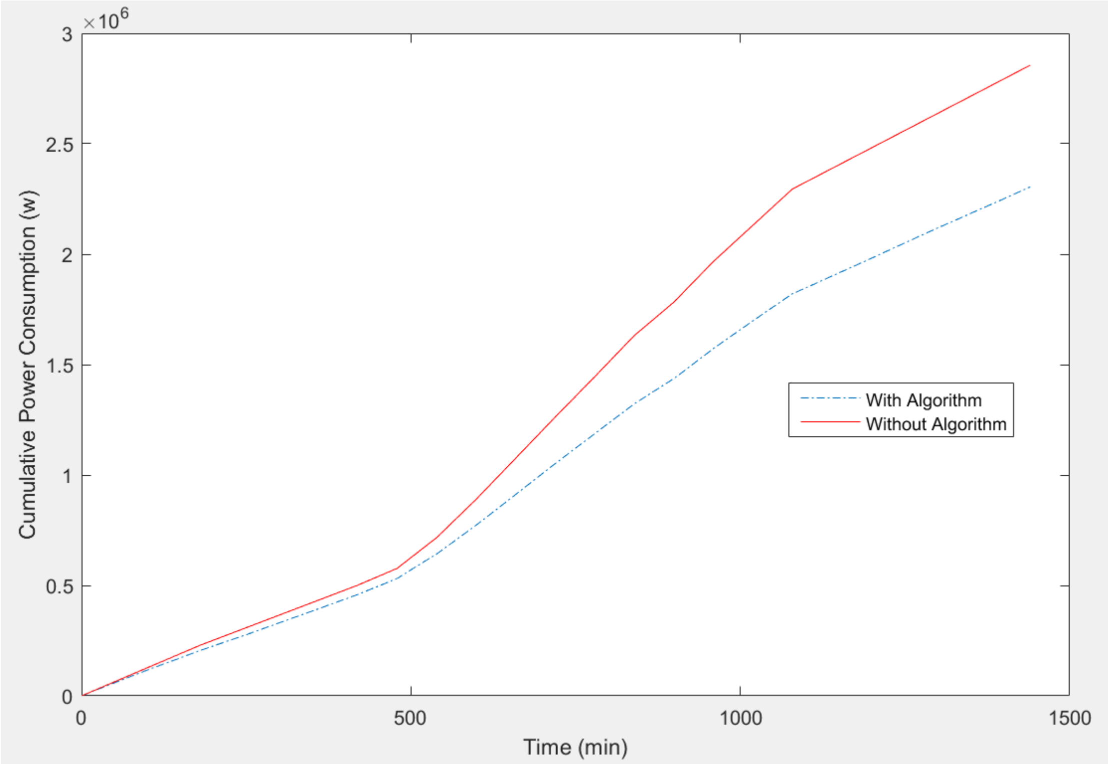

Figure 5.10 Cumulative Power Consumption with Algorithm vs. without Algorithm (1 Day)

Figure 5.10 shows the comparison of cumulative power consumption. This plot can clearly show the performance evaluation about the algorithm. It is not hard to find from the plot that at the initial stage of the simulation, the performance of using the algorithm is not that obvious, since the traffic load is not that high and large proportion of the power consumption is the baseline power consumption of BS. So, the impact of switching on/off the small cells is not that apparently to see. However, as the traffic load increase, the network without the algorithm will largely increase due to the high load. In this case, the algorithm can show its effect in switching the small cells to offload the macro base station and slowing down the increment speed of power consumption.

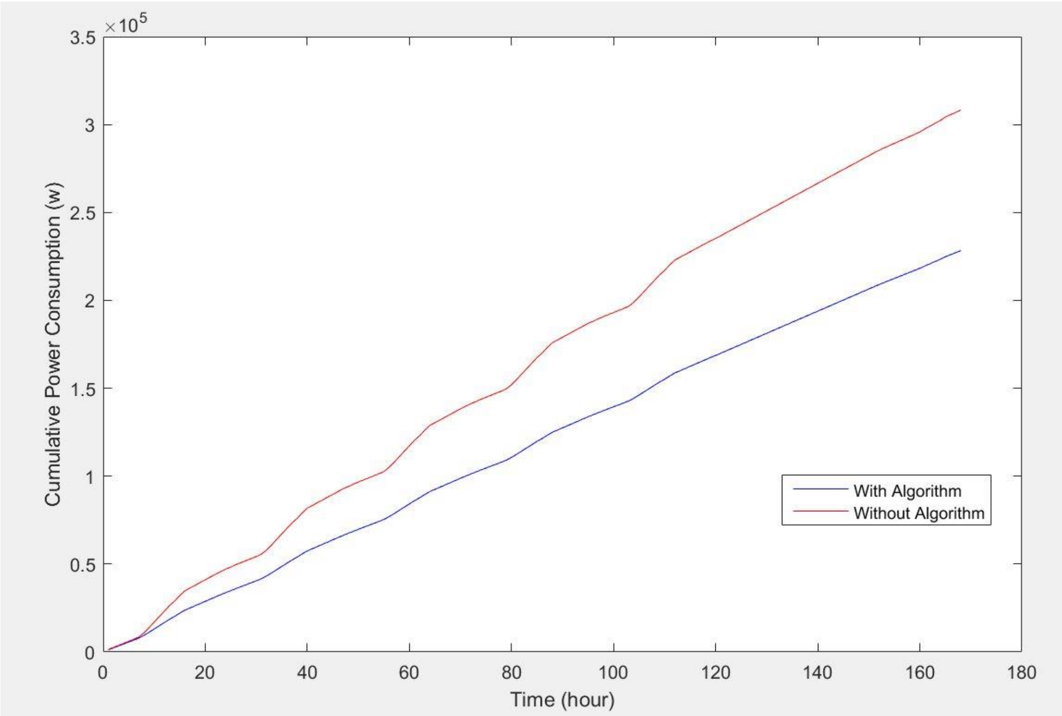

Figure 5.11 Cumulative Power Consumption with Algorithm vs. without Algorithm (1 Week)

In office scenario, the traffic load can be totally different during weekday and the weekend. On weekend, as shown in figure 4.2, the traffic load is pretty low since most people go back home to spend their weekend and few people will choose to go to the office building. In this way, figure 5.11 shows the power consumption of one week to evaluate the performance of the algorithm both from weekday and the weekend. Similarly, the effect of using the algorithm is not obvious during the off-peak hours because of the low traffic load. But during those weekdays when stuffs sit in their office rooms and use all kinds of applications, the algorithm shows a good performance in reducing the power consumption.

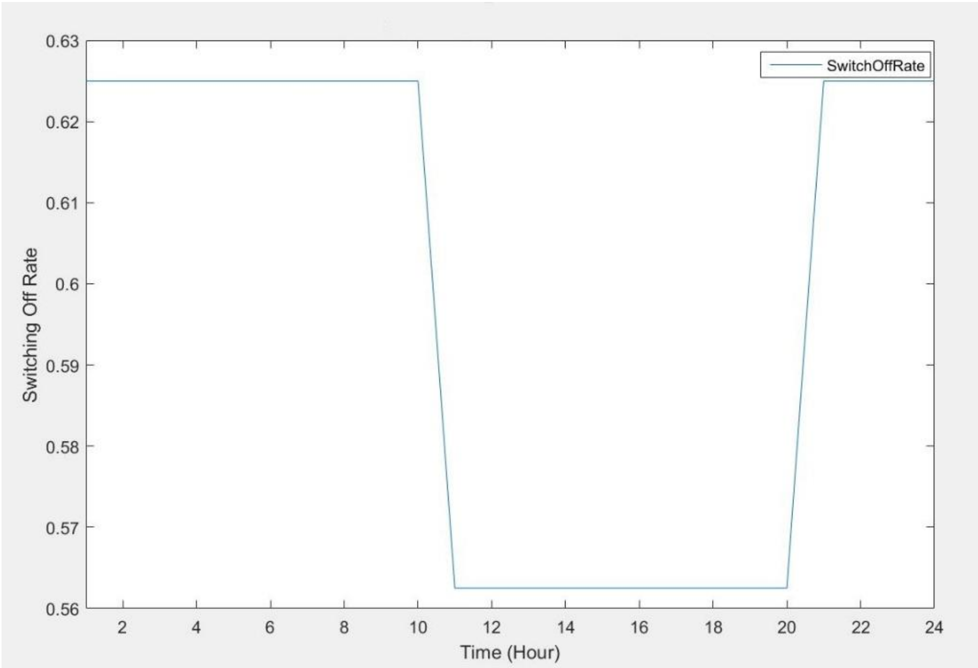

Figure 5.12 Switching Off Rate of SCs During a Day

Figure 5.12 shows the details about the switching Off rate of small cells in the network. To show the result clearly, the simulation is run based on an hour for each iteration. The plot shows the trend that during those hours when traffic load is low, many small cells are switched off to save energy. While in those traffic hours, small cells are switched on to meet the capacity requirements in the cell.

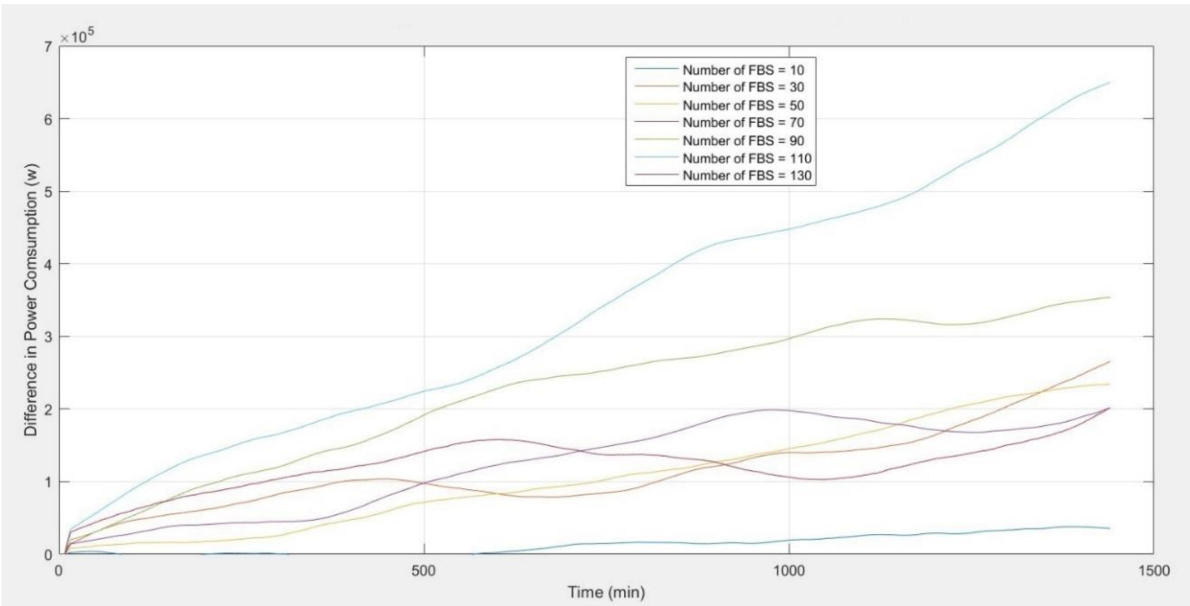

Figure 5.13 Power Saving with Different number of SCs

Figure 5.13 shows the power saving graph of one day with different number of small cells in the network. The Y axis of the figure shows the difference value between system power consumption with the algorithm and without the algorithm. Previous simulations are based on 10 small cells in the macro cell to minimum the simulation time. However, in real case, larger number of small cells can be deployed in the network to increase the capacity or the coverage. In this way, comparison is made among different number of small cells in the macro cell to further evaluate the performance of the algorithm. From figure 5.13, it is easy to find that with 110 small cells deployed in the network, the power saving efficiency is the best. While when the number of small cells is small, like 10, the power saving efficiency is not that remarkable as larger number.

5.3 Results Summary

Results of this project including results in the throughput simulation, algorithm performance and network performance have been listed and analyzed above. From these plots and results, it could be concluded that off-policy Q-Learning Algorithm shows better performance in network power saving than on-policy SARSA Algorithm. Different choices of learning rate and reduction factor of the Q-Learning algorithm can show slightly different effect on power saving performance. By using the Q-Learning algorithm, small cells with less users can be switched off during the off-peak hours while switched on during the peak hours. By achieving this, the network power consumption can be reduced especially during the peak hours. However, delay is shown in the controlling process and power consumption during the off-peak hours cannot be reduced remarkably because the macro cell case station consumes large proportion of the total power consumption. By switching off some small cells, the throughput gets lower than that without the algorithm. It shows the tradeoff between the higher throughput and lower power consumption. Less small cells remaining working in the network means more users share the total spectrum resources, which lead to the decline of the average throughput. This kind of tradeoff can be balanced when the decrement of throughput is within an acceptable range. The results show that the average throughput after using the algorithm is still enough to meet the users' demands. The algorithm can show different performance when there are different numbers of Femtocell base stations deployed in the network. The best performance in power saving will be shown when about 110 Femtocells deployed in the network. These results together show that the Q-Learning algorithm developed in this project can indeed help saving power without sacrificing much network capacity.

Conclusions

This project develops an algorithm to switch on/off small cells to help save network energy and meanwhile meet the capacity requirement. A distributed implementation of switching on/off algorithm in an office scenario network with one macro cell and several small cells has been made. After the algorithm being introduced, small cells can offload the macro cell dynamically to help save the network energy. The algorithm is based on the Markov Decision Process. The two methods to get the best policy, on-policy SARSA Algorithm and off-policy Q-Learning Algorithm, have been taken into consideration. The final results show that the Q-Learning Algorithm shows better performance in saving power. More details about the parameters of Q-Learning Algorithm have been compared to get a better performance. To evaluate the performance of the algorithm, a simplified network model has been built. The data of traffic load comes from the simulation by using OPNET Modeler, whose results are more realistic in the office scenario. The main evaluation is run in MATLAB and the final results show that the algorithm can save power consumption especially during the peak hours when traffic load is large. When the traffic load is relatively low, i.e. large proportion of the power consumption is baseline power consumption of base station, the performance is not that obvious but still can save energy by switching off some small cells.

The solution can still be improved in the future. Although the traffic data are based on the realistic simulation results from OPNET Modeler. Model built in MATLAB is simplified and may different from the result of much complex network in real case. To get more realistic results, the model in MATLAB can be called in OPNET Modeler to evaluate the performance of the algorithm in the office scenario in LTE network. The implementation is relatively complicated and need to change the model files in OPNET Modeler. At current stage, due to the limit in time and capability, the algorithm is evaluated in a simplified model rather than a complex network in real case.

Bibliography

[1] Chen, S. and J. Zhao. "The requirements, challenges, and technologies for 5G of terrestrial mobile telecommunication." *IEEE Communications Magazine.* 52 (5). 2014. p. 36-43.

[2] Bohli, A. and R. Bouallegue. "Energy Efficiency Performance for 5G Cellular Networks." *International Federation for Information Processing.* 2017. 99-111.

[3] Ge, X. et al. "5G Ultra-Dense Cellular Networks." *Paper presented at the Proceedings of IEEE Wireless Communications Workshops*. 2015.

[4] Arnold, O. et al. "Power Consumption Modeling of Different Base Station Types in Heterogeneous Cellular Networks." *Paper presented at the Proceedings of Future Network and Mobile Summit 2010 Conference.* 2010.

[5] Bhaumik, S. et al. "Breathe to stay cool: adjusting cell sizes to reduce energy consumption." *Paper presented at the Proceedings of the first ACM SIGCOMM workshop on Green networking, New Delhi, India*. 2010.

[6] Simsek, M. and A. Czylwik. "Improved decentralized fuzzy q-learning for interference reduction in heterogeneous LTE-networks in OFDM." *Paper presented at the Proceedings of 17th International OFDM Workshop 2012 (InOWo'12).* 2012.

[7] M., E. et al. "Cooperative Q: Energy-efficient channel access based on cooperative reinforcement learning." *Paper presented at the 2015 IEEE International Conference on Communication Workshop (ICCW).* 2015.

[8] You, S. and Z. Qi. "Research and Prospect of the Status Quo of Green Communication Technology in 5G Network." *Mobile Communication.* 40(20). 2016. p. 31-35.

[9] Cameron, T. and P. Forbes. "Analog Front End for 3G Femto Base Stations Brings Wireless Connectivity Home." *Analog Dialogue*. 42(12). 2008. p. 1.

[10] Holma, H., A. Toskala and J. Reunanen. L*TE Small Cell Optimization: 3GPP Evolution to Release 13*. John Wiley & Sons Ltd. 2016. p. 59-74.

[11] Hoydis J and M. Debbah. “Green, Cost-effective, Flexible, Small Cell Networks.” *IEEE Communications Society MMTC*. 5 (3). 2010. p. 23-26.

[12] Willkomm, D. et al. “Primary Users in Cellular Networks: A Large-Scale Measurement Study.” *2008 3rd IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks.* Chicago. IL. 2008. p. 1-11.

[13] Lopez-Perez, D., I. Guvenc and X. Chu. “Mobility Management Challenges in 3GPP Heterogeneous Networks.” *IEEE Wireless Communication Magazine*. 2012(12). p. 70-78.

[14] Richter, F., A. J. Fehske and G. P. Fettweis. “Energy Efficiency Aspects of Base Station Deployment Strategies for Cellular Networks.” *2009 IEEE 70th Vehicular Technology Conference Fall. Anchorage*. AK. 2009. p. 1-5.

[15] Wang Y. and S. Lee. *Small-cell planning in LTE HetNet to improve energy Efficiency*. Int J Commun Syst. 2018.

[16] Hartigan, J. A. and M. A. Wong. “Algorithm AS 136: A K-Means Clustering Algorithm.” *Journal of the Royal Statistical Society. Series C (Applied Statistics)*. 28 (1). 1979. p. 100-108.

[17] Miozzo, M. et al. “Distributed Q-learning for energy harvesting Heterogeneous Networks.” *2015 IEEE International Conference on Communication Workshop (ICCW)*. London. 2015. p. 2006-2011.

[18] Sutton, R. S. and A.G. Barto. *Reinforcement Learning: An Introduction*. Cambridge MA: MIT Press. 1998.

[19] Lu, Z. and H Yang. *Unlocking the Power of OPNET Modeler*. Cambridge: Cambridge University Press. 2012.

[20] Sauter, M. *From GSM to LTE-advanced: an introduction to mobile networks and mobile broadband (Revised 2nd edition)*. John Wiley & Sons Ltd. 2014. p. 237-248.

[21] Marsan, M. A. et al. “Towards zero grid electricity networking: Powering BSs with renewable energy sources.” *2013 IEEE International Conference on Communications Workshops (ICC)*. Budapest. 2013. p. 596-601.

[22] Alshami, M., T. Arslan and J. Thompson. “Femtocell Interference, SINR and the

Probability of Connection at Downlink." *Int. J. Com. Dig. Sys.* 3(2). 2014. p. 79-89.

Acknowledgements

This bachelor thesis was finished during the period when I was studying as an exchange student at Technical University of Denmark (DTU). At the end of the thesis, I would like to express my sincere gratitude to those who have offered me help.

Firstly, I would like to thank for great help from my supervisor, Dr. Si Pengbo. Although there are 6-7 hours of time difference, Dr. Si Pengbo offered me great help about my work and gave me advice when I got stuck on a specific point. Only with his care could I finish this project and get the expected results. I would also like to appreciate for the help from prof. Henrik Lehrmann Christiansen at DTU FOTONIK. Last semester, I took Henrik’s course “Introduction to Mobile Communication”. The course helped me a lot in my project and prof. Henrik Lehrmann Christiansen gave me some suggestions about my modeling work after class. Without his help, it would be hard for me to start up my project. I would also express my gratitude to prof. Michael Stübert Berger. He was my supervisor of a group project on “Broadband Network” course. He gave us much advice on how to handle the scientific work in an appropriate scientific way. Thirdly, I would like to appreciate everyone who gave me a hand during my exchange period. In particularly thanks to the help from teacher Zhou Zexi and my classmate Sun Yiming, for they helped me deal with some problems when I was not able to show up personally.

Finally, gratitude for all supports and trusts from my family, all these supports gave me power to face all the difficulties during this whole period.

Thanks again for all the helps. Only with these helps could I finish this bachelor thesis on time.