國立交通大學

電信工程研究所

碩士論文

異質無線網路環境下以模擬退火演算法為基礎之群換手決策機制

A Simulated Annealing Based Group Handover Decision Scheme in Heterogeneous Wireless Networks

研究生：蒲恆皜

指導教授：張仲儒 博士

中華民國 壹百零壹年八月

異質無線網路環境下以模擬退火演算法為基礎之群換手決策機制

A Simulated Annealing Based Group Handover Decision Scheme in Heterogeneous Wireless Networks

研究生：蒲恆皜	Student：Heng-Hao Pu
指導教授：張仲儒 博士	Advisor：Dr. Chung-Ju Chang

國立交通大學
電信工程研究所
碩士論文

A Thesis
Submitted to Institute of Communication Engineering
College of Electrical and Computer Engineering
National Chiao Tung University
in Partial Fulfillment of the Requirements
for the Degree of Master of Science
in
Communication Engineering
August 2012
Hsinchu, Taiwan

中華民國壹百零壹年八月

異質無線網路環境下以模擬退火演算法為基礎之群換手決策機制

研究生：蒲恆皜　　　　　指導教授：張仲儒

國立交通大學電信工程研究所

摘要

因為各種不同無線網路技術的發展，創造了異質無線網路的環境。為了提供群體(group)存在於大眾交通工具上的行動終端裝置(mobile terminal)無縫的(seamless)網路接取，賦予行動終端裝置多模(multi-mode)的能力讓行動終端裝置能夠能夠連接到多元異質網路的能力在此環境下是不可或缺的，這些行動終端裝置根據通道狀況回報訊息給基地台進行換手(handover)決定適宜的網路的同時，能夠考慮到使用者的需求以及系統業者的支出便是異質無線網路下群換手決策的重要議題。

在由 GSM/EDGE、WCDMA、LTE 組成的異質無線網路中，為了降低群體通話阻斷比率、減少換手的頻率、平衡網路之間的交通負載，並且希望保證使用者的服務品質需求(QoS requirement)，在本篇論文中，我們提出了一個以模擬退火(simulated annealing)演算法為基礎並且使用成本函數(cost function)模型來決定群體行動裝置適當網路的群換手決策機制。在模擬結果中顯示我們提出的方法可以有效的降低群體通話阻斷比率，同時能平衡網路之間的負載，除此之外，在群體數量中等時，換手的頻率也有一定幅度的下降，並且能符合品質需求。

A Simulated Annealing Based Group Handover Decision Scheme in Heterogeneous Wireless Networks

Student: Heng-Hao Pu Advisor: Chung-Ju Chang
Institute of Communication Engineering
National Chiao Tung University

English Abstract

Abstract

Nowadays, a variety of wireless network technologies is developed. Different wireless networks (e.g. GSM/EDGE, WCDMA, LTE) create a heterogeneous wireless network environment. For those multi-mode mobile terminals(MTs), it is important to provide a seamless service for them. When there is a group of MTs stay closely in a vehicle and decide to handover to another networks, it is an important issue to help these MTs to decide a proper target network which is good for both users and the system provider.

In this thesis, we tend to reduce the group handover blocking ratio, reduce the number of handover, balance the traffic load between networks, and meet the quality of service (QoS) requirements. Therefore, a simulated annealing based group handover decision scheme using cost function is proposed. In the simulation result, the proposed scheme can reduce the group handover blocking ratio efficiently and balance the load. Besides, when the number of group MTs is not quite large, the number of handover also reductive. Also, it can meet the QoS requirements.

Acknowledgement

誌 謝

兩年的碩士生涯很快的就過去了，經過兩年來對論文方向的摸索、定題目、深入研究以及程式模擬，在每個過程中都讓我成長許多，能順利的完成這篇碩士論文，必須感謝非常多人的鼓勵與幫忙。首先要感謝張仲儒教授這兩年來在專業知識和待人處世方面的指導，讓我在學業和生活上都獲益良多。感謝文敬、聖章、耀興、振宇和益興學長們，每在我遇到困難的時候，能夠提供我寶貴的意見。還有在進實驗室時熱情歡迎我們的碩士班學長姐俊魁、苡芾和竣威。當然，還要謝謝和我一起努力的兼源、長青、佩珊和moro，大家一起玩耍、聊天真的很快樂。還有努力帶給我們實驗室歡樂時光的學弟華立、韋成和學妹巧翎，在我需要更多電腦來跑程式時給予我很大的幫助。以及辛苦的助理玉棋及佳婷。有你們的陪伴，研究的路途即使有著苦悶，更多了很多很多的歡笑。

最後，我要感謝我的父母、大姐、二姐還有支持我的朋友們，有了他們對我的支持，我才有動力度過每一個難關，順利完成碩士的學業。

蒲恒皜 謹誌
民國一百零一年八月

Contents

List of Figures

List of Tables

Chapter 1
Introduction

In recent years, the use of smart phones and tablet personal computers has increased noticeably. People change their manners from watching newspaper to watching videos or browsing the web by smart phones when they are taking bus, train, or mass rapid transit (MRT). Therefore, there have a large group of people who seat closely and move together using their smart phones in the same vehicle. In the mean time, there are different access cellular networks for smart phones, such as global system for mobile communications (GSM), the enhanced data rate for GSM evolution (EDGE), wideband code division multiple access (WCDMA) network, long term evolution (LTE) and etc. These networks comprise a heterogeneous network and serve those users in the same area.

In a vehicle, there have many mobile terminals (MTs) that are served by different wireless networks. When the vehicle drives away from a base station (BS) of one of the wireless networks, there have more than one MTs that served by the BS triggering their handover at the same time duo to that their channel conditions are similar, which is caused by that they are close to each other. These MTs with same serving BS and same serving network forms a handover group and a group handover decision (GHOD) is going to make for these MTs.

In traditional handover decision scheme, it only consider that there is only one MT doing handover at a time. Therefore, the metrics such as received signal power, signal-to-interference and noise ratio, and so on that are used to decide the target BS of a MT. The MT may choose the best BS which have largest received signal power to be the target BS. If the single MT handover decision scheme is used in group handover decision scheme, the MTs in the handover group are more likely to handover to the same “best” BS due to that it doesn't consider the impact of other MTs such as traffic loading, traffic type, and so on. This will result in network congestion if the target BS cannot provide enough resources. In other words, the handover blocking ratio of the handover group (group handover blocking ratio) will increase. In addition, it also causes inefficient utilization of system resource due to the traffic loading between each BS is not balanced, which may result in a highly new call blocking rate if there have some MTs want to access the network. Besides, there are many kinds of traffic types nowadays, such as voice, video, HTTP, and FTP. Each traffic type has its own requests. Therefore, the design of GHOD scheme needs to satisfy their requirements. Furthermore, each handover requires network resources to reroute the MT to a new BS. So a high number of handover spend more network resources. Also, it increases more delay in the processing of handover requests and the chance that a call will be denied access to a BS. Therefore, decreasing the number of handover also needs to be considered when designing the GHOD scheme.

The GHOD involves measurements and information collection. Based on which device do the measurement and collect the information, and which device do the group handover decision, the GHOD can be categorize into four types [1].

Type 1: Mobile - controlled GHOD (MCGHOD) [2] [3]

In MCGHOD, each MT must take its own measurement and also make the handover decision by its own. It is the highest decentralize handover decision scheme.

Type 2: Network - assisted GHOD (NAGHOD) [4]

In NAGHOD, network collects information such as network capacity, traffic load of other MTs, and so on to each MT. After that, each MT make the handover decision by its own.

Type 3: Mobile - assisted GHOD (MAGHOD) [3] [5] [6]

The MTs will in charge of doing measurements and collecting information to inform their serving network. After receiving the information, the serving network will do the GHOD.

Type 4: Network - controlled GHOD (NCGHOD)

The measurements and information collection are all doing by network. This type will increase the burden of the network side. Hence, it is seldom to use.

In GHOD, load balancing is important to increase the overall system utilization. In [2], Cai and Liu proposed a MCGHOD scheme using random delay. Each MT makes the group handover decision after a random delay, and decides a target network with minimal load intensity. This scheme let each MT avoid making decision simultaneously in order to improve the load balancing between each network. But it does not consider the QoS guarantee for different service, and the random delay will increase the handover latency.

In [3], S. Lei et al. proposed two MCGHOD scheme. The first scheme is that different traffic types use different random delay interval to decide when to start the GHOD, and selects a target network with minimal load intensity. The second scheme is that different traffic types has different predefined selecting probability of each network. This scheme will not bring the unnecessary delay during handover. However, it cannot get the optimal decision results due to the fact that the predefined selecting probability is subjective, and it does not guarantee that the group handover blocking ratio is less than a predefined value.

In [4], W. Lee and D. H. Cho proposed a NAGHOD scheme using both adjusted delay and adjusted probability that can limit the group handover blocking ratio under a predefined value, which is, each MT makes the GHOD after an adjusted delay, and decide its target network by using adjusted probability. However, this scheme does not consider the QoS guarantee for different services.

In [3], S. Lei et al. also proposed a MAGHOD scheme. it deploy a common radio resource management (CRRM) entity between each network. The CRRM entity collects the handover information from the group of handover MT and multiple networks. The GHOD scheme in CRRM entity decides the target network of all the real time MTs by minimize the average transmission delay, and the target network of all the non-real time MTs that by minimize the average packet loss rate.

Once more, the group handover blocking ratio is an important factor in GHOD. In [5], L. Shan et al. proposed a predictive GHOD with channel borrowing using mobile relay scheme to reduce the group handover blocking ratio. First, it deploy a mobile relay in a bus to reduce the system signaling overhead which is caused by group of MTs performing handover frequently. Second, this scheme will predict the target BS and send the handover channel borrowing requests to the target BS when there is no free channels in the target BS for the coming handover MTs. But the channel cannot be borrowed when there is no connection using more channels than the minimal number of channel in the target BS.

In [6], G. Zhang and F. Liu proposed an auction based GHOD with mobility prediction scheme. It formulate the decision problem as a multi-assignment problem to maximize the system throughput, reduce the number of handover, and keep the load of all candidate networks under a target level. After that, an auction algorithm is adopted to solve this problem. However, the residence time of a vehicle in candidate networks, which used to judge the number of handover is not precise. And this paper does not consider the group handover blocking ratio, which is an important factor in GHOD.

The group handover decision is in charge of deciding proper target network for different MTs. It is recognized as a NP-Hard problem [7], which means a great deal of time is needed to find the optimal solution. In this thesis, we use the simulated annealing algorithm to solve the problem. It is a good technique for minimizing functions of many variables [8]. Moreover, it is more efficient than Genetic algorithm when they achieve the same performance [8]. Therefore, a simulated annealing based group handover decision (SAGHOD) scheme is proposed for heterogeneous wireless networks. The scheme of SAGHOD intends to minimize the group handover blocking ratio, balance the loading between networks, minimize the number of handover, and support QoS requirement such as average packet delay, average packet dropping rate, and minimum transmission rate. Based on the goals in the above, we create a cost function and try to find a solution that can minimize the value of the cost function and also satisfy two constraints.

The rest of the thesis is organized as follows. Chapter 2 describes the system model. Chapter 3 formulates the GHOD problem, introduces the simulated annealing algorithm, and describes the proposed SAGHOD scheme in detail. The simulation results and discussions are shown in Chapter 4, and the conclusions are given in Chapter 5.

Chapter 2
System Model

2.1 Heterogeneous Wireless Network Environment

Fig. 2.1 shows the heterogeneous wireless network environment which containing a GSM/EDGE cellular network, a WCDMA cellular network, and a LTE cellular network, where the coverage of these three networks are overlaid on the same area. For a MT *i* in the heterogeneous network, we define the traffic class of the MT as *h*, where $h \in \{1, 2, 3, 4\}$. The MT with $h = 1, 2, 3, \text{or } 4$ represents the MT with voice traffic, video traffic, HTTP traffic, or FTP traffic, respectively. A cellular network is defined by *j*, where $j \in \{1, 2, 3, 4\}$. $j = 1$ represents GSM system in GSM/EDGE network with inter-side distance (ISD) 1 km. $j = 2$ represents EDGE system in GSM/EDGE network. $j = 3$ and $j = 4$ represent WCDMA network with ISD 0.645 km and LTE network with ISD 0.5 km, respectively. In network *j*, the load intensity of network *j* is denoted by ρ_j, the average voice/video packet delay and the average voice/video packet dropping rate are denoted by $d_{j,h}$ and $p_{j,h}$, $h = 1, 2$, and the average HTTP/FTP transmission rate of network *j* are denoted by $R_{j,h}$, $h = 3, 4$. In the next subsections, we introduces the network environment of the four networks.

2.1.1 GSM/EDGE Network

In GSM/EDGE network, we assume that each BS has N_f downlink frequency channels, and each channel has 200 kHz bandwidth. For each downlink frequency channel, as shown in Fig. 2.2 [9], it use 52 TDMA frames structure to form a multiframe. In a multiframe, two frames are reserved for transmission of the packet time advanced control channel (PTCCH), and the other two frames are idle frames.

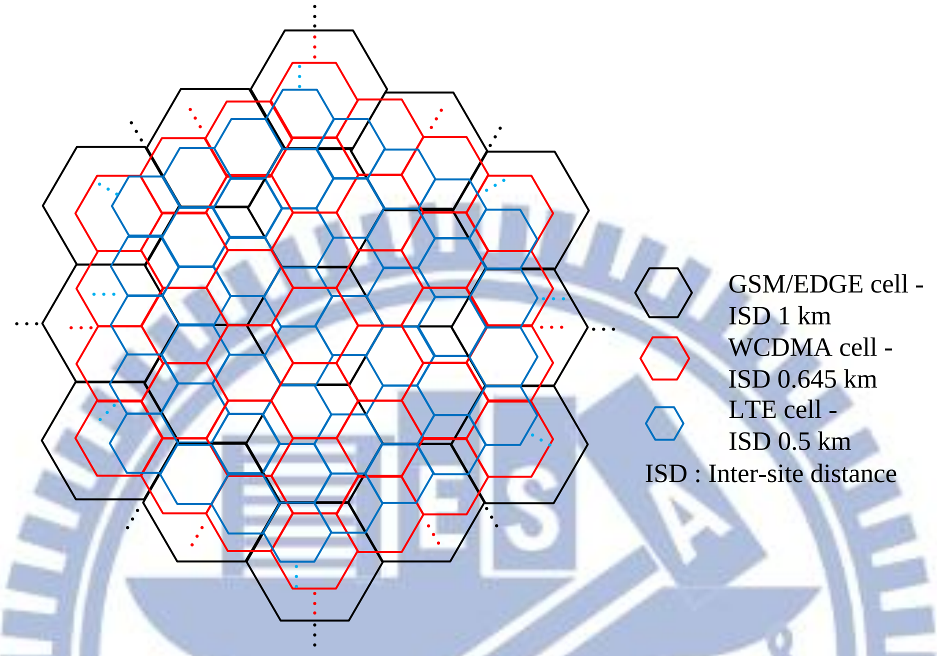


Fig. 2.1 : Heterogeneous Wireless Network Environment

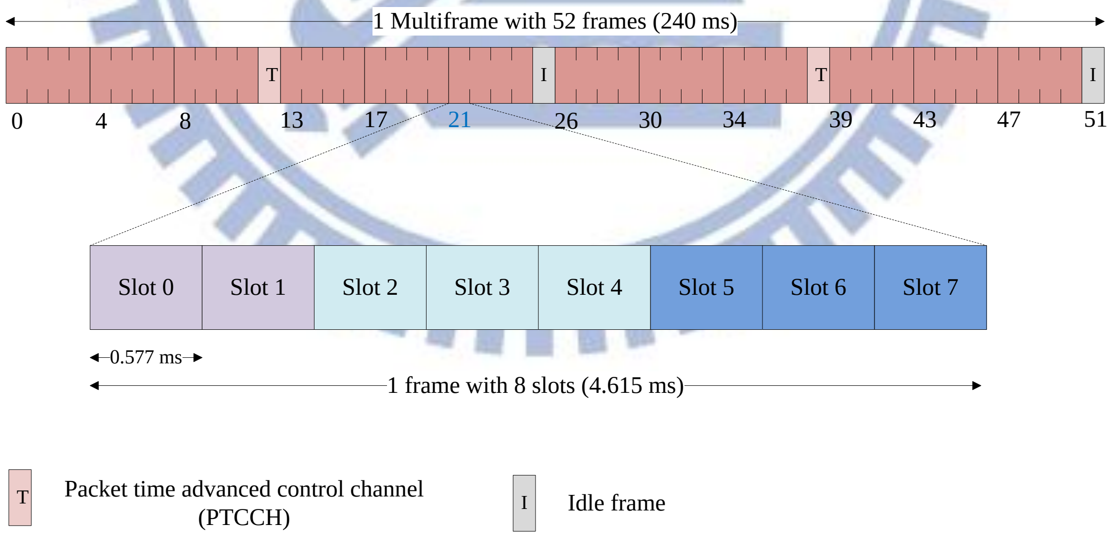


Fig. 2.2 : GSM/EDGE Multiframe Structure

The slots in the remaining 48 frames are allocated to different information. Slot 0 and slot 1 are used to transmit the control message of every MT. Slot 2, slot 3, and slot 4 are used to transmit GSM voice signal, and each slot here can only serve one voice MT. Slot 5, slot 6, and sot 7 are used to transmit EDGE data MT, and each slot here can serve multiple data users. Herein, every data user can use three slots in one frame at most. The time of each frame is 4.615 ms with eight slots, and each slot is 0.577 ms.

After that, we can calculate load intensity by knowing the characteristic of resources. A load intensity of a BS is defined as a ratio between 0 and 1. It represents whether the traffic loading is light or heavy in the BS. Because of GSM/EDGE use circuit switch for voice MTs (slot 2 to 4) and use packet switch for data MTs (slot 5 to 7), it needs two load intensity to represent these two different services. In addition, we only consider the downlink side resource allocation due to the data is transmitted from base station to MT mostly. The load intensity of a GSM BS is defined as

$$\rho_1 = \sum_{i'=1}^{N_G} \Delta\rho_{i',1}, \tag{2.1}$$

where N_G is the total number of voice MTs that are served by the BS of GSM. $\Delta\rho_{i',1}$ is the increased load of a BS of GSM caused by MT i', and is defined as

$$\Delta\rho_{i',1} = \frac{1}{N_t \times N_f}, \tag{2.2}$$

where N_t is the total number of time slots for voice MT in a frame. N_f is the total number of frequency channels.

The EDGE load intensity, denoted by ρ_2, is defined as

$$\rho_2 = \sum_{i'=1}^{N_E} \Delta\rho_{i',2}, \tag{2.3}$$

where N_E is the total number of MTs that are served by the BS of EDGE. $\Delta\rho_{i',2}$ is the increased load of a BS of EDGE caused by MT i', and is defined as

$$\Delta\rho_{i',2} = \frac{c_{i'}}{R_E^*}, \tag{2.4}$$

where $c_{i'}$ is the required mean data rate of MT i', and R_E^* is the maximum transmission rate of EDGE BS.

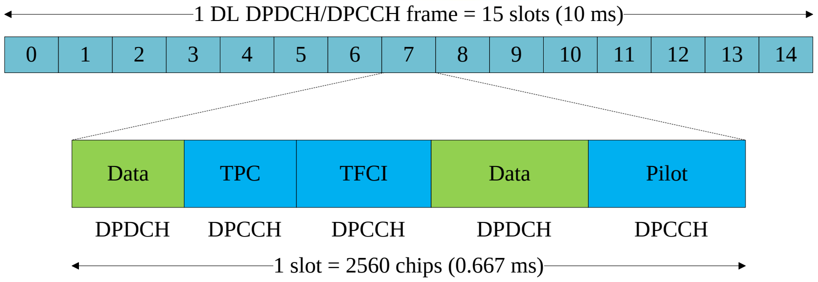
The handover parameter of GSM/EDGE that we use is received signal power in dBm, denoted as P_{Rx}, can be obtained by

$$P_{Rx} = P_{Tx} + F_l(r), \tag{2.5}$$

where P_{Tx} is transmitted signal power in dBm from serving BS. $F_l(r)$ is large scale fading in dB. The handover triggered condition is $P_{Rx} \le \eta_E$, where η_E is the handover threshold in dBm.

2.1.2 WCDMA Network

In WCDMA network, the bandwidth of DL frequency channel of WCDMA BS is 5 MHz. Furthermore, it use different orthogonal codes to differentiate different physical channel in the same frequency and same time duration. MT's data is sent by dedicated physical data channel (DPDCH) and the control information is sent by dedicated physical control channel (DPCCH). The frame structure of downlink DPDCH/DPCCH is shown in Fig. 2.3.



TPC Transmit Power Control
TFCI Transport Format Combination Indicator

Fig. 2.3 : Downlink DPDCH/DPCCH Data/Control Multiplexing

In one power control period, there is 10 ms length frame which is split into 15 slots and there are 2560 chips in each slot. It means that the chip rate will be 3.84 Mbps. It is Assumed that we can use pilot control signal in DPCCH to broadcast a BS load intensity in each frame. The BS load intensity of WCDMA is defined as [10]

$$\rho_3 = \sum_{i'=1}^{N_W} \Delta\rho_{i',3}, \tag{2.6}$$

where N_W is the total number of MTs that are served by the BS of WCDMA in a frame. $\Delta\rho_{i',3}$ is the increased load of a BS of WCDMA caused by MT i', which is defined as

$$\Delta\rho_{i',3} = v_{i'} \cdot \frac{(E_b / N_0)_{i'}}{W / c_{i'}} \cdot [(1-\alpha_{i'}) + \beta_{i'}] \tag{2.7}$$

The $v_{i'}$ is the activity factor of MT i'. The $(E_b / N_0)_{i'}$ is the required signal energy per bit divided by noise spectral density of MT i'. The W is the chip rate. $c_{i'}$ is the required mean data rate of MT i'. The $\alpha_{i'}$ is the channel orthogonality of MT i'. The $\beta_{i'}$ is the ratio of other cell BS power to own cell BS power, received by MT i'.

The handover parameter of WCDMA is the received signal code power (RSCP) in CPICH. The own RSCP is defined as P_{RSC} in dBm, P_{RSC} can be obtained by

$$P_{RSC} = P_{TSC} + F_l(r), \tag{2.8}$$

where the P_{TSC} is the transmitted signal code power from serving BS in dBm. The $F_l(r)$ is the large scale fading in dB. While the handover triggered condition is $P_{RSC} \le \eta_W$, where η_W is handover threshold in dB.

2.1.3 LTE Network

In LTE network, it is assumed that the total downlink bandwidth is 3 MHz. In 3 MHz, there are 15 subchannels and the bandwidth of each subchannel is 180 kHz. One subchannel is divided into 12 subcarriers, and each subcarrier is with 15 kHz. In time domain, one frame is 10 ms and is consist of 10 subframes. And one subframe is formed by two OFDM control symbols and 12 OFDM symbols. In addition, we defined one resource unit (RU) as one subframe over one subchannel as shown in Fig. 2.4. [11]

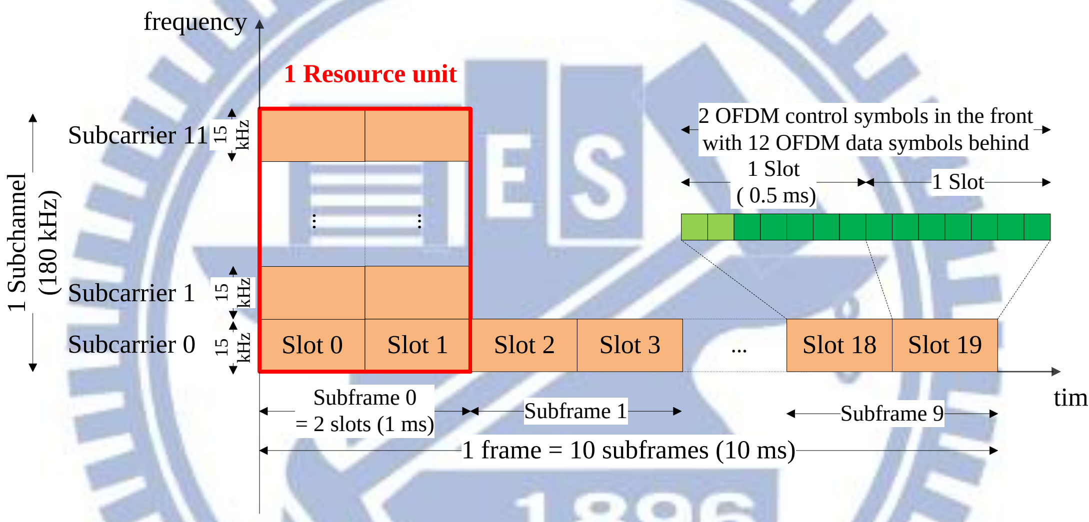


Fig. 2.4 : Resources in LTE System

There are 150 RUs in one frame over 3MHz. In LTE resource allocation, RU is a basic unit of resource, which means that one RU can only serve one user, or many RUs serve the same user. By ignoring the coding rate and the deployment of antenna, we can calculate the number of bits in a resource unit by multiplying 12 data symbols in a subframe, 12 subcarriers in a subchannel, and *m* bits per symbols of this resource unit. That is, 144*m* bits. The BS load intensity of LTE is defined as

$$\rho_4 = \sum_{i'=1}^{N_L} \Delta\rho_{i',4}, \qquad (2.9)$$

where the N_L is the total number of MTs that are served by the BS of LTE in a frame. $\Delta\rho_{i',4}$ is the increased load of a BS of LTE caused by MT i', and is defined as

$$\Delta\rho_{i',4} = \frac{c_{i'}}{R_L^*}, \tag{2.10}$$

where $c_{i'}$ is the required mean data rate of MT i', and R_L^* is the maximum transmission rate of LTE BS.

The handover parameter of LTE that we use is received signal power in dBm, denoted as P'_{Rx}, can be obtained by

$$P'_{Rx} = P'_{Tx} + F_l(r), \tag{2.11}$$

where P'_{Tx} is transmitted signal power in dBm from serving BS. $F_l(r)$ is the large scale fading which measured in dB. While the handover triggered condition is $P'_{Rx} \le \eta_L$, where η_L is a handover threshold in dBm.

2.2 Channel Model

The wireless fading channel is composed of large-scale fading and small-scale fading. The large scale fading includes path loss and shadowing effect, which represents the average signal strength loss due to the distance between transmitter-receiver (T-R). The small-scale fading reflects the rapid fluctuation of the radio signal over a short period of time. The large-scale fading is modeled as [12] [13]

$$F_l(r) = \overline{PL}(r) + X_\sigma, \tag{2.12}$$

where $\overline{PL}(r)$ is the average path loss for an arbitrary T-R separation, which is defined by

$$\overline{PL}(r)=\overline{PL}(r_0)+10n\log_{10}\left(\frac{r}{r_0}\right) \tag{2.13}$$

where $\overline{PL}(r_0)$ is measured path loss at the close-in reference distance r_0 km. n is path loss exponent. The value of n depends on the specific propagation environment. That is, n = 3 ~ 5 for shadowed urban cellular radio environment. The r_0 is the close-in reference distance, which is determined from measurement close to the transmitter. The d is T-R separation distance. The X_σ is zero-mean Gaussian distributed random variable with standard deviation σ_X. This is caused by the fact that the surrounding environmental clutter may be vastly different at two different locations having the same T-R separation.

In this thesis, the small-scale fading is not considered. It is because that the small-scale fading will cause ping-pong handover [14]. The ping-pong handover is not our main consideration.

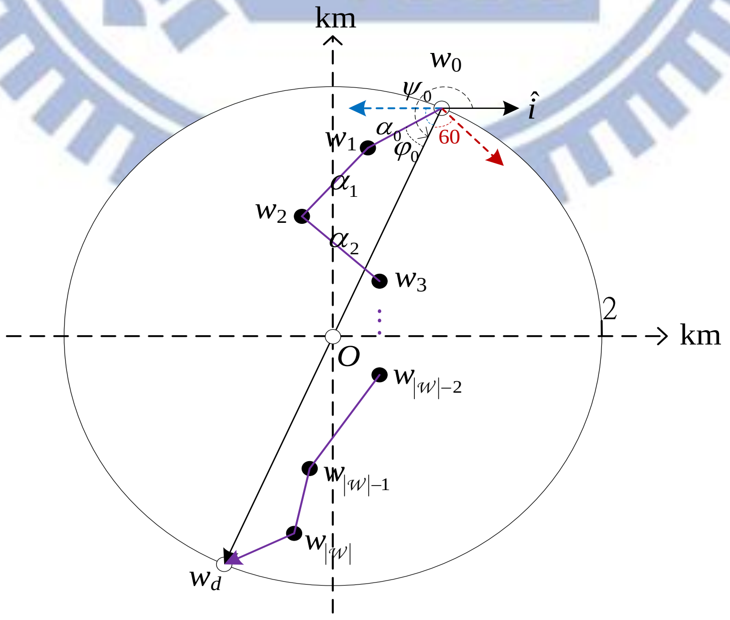


Fig. 2.5 : Mobility Model

2.3 Mobility Model

In our mobility model, we modify the random waypoint model [15] to model the mobility of a vehicle. The proposed modified random waypoint model is created by the following steps:

Step 1: Generate an initial waypoint of a bus w_0 and a destination waypoint of the bus w_d.

The coordinates of w_0 is (x_0, y_0) which is randomly selected on the boundary of the circle with center $O(0,0)$ and radius 2 km as shown in Fig. 2.5. The coordinates of w_d is $(-x_0, -y_0)$ which is symmetric about the center with w_0.

Step 2: Generate a waypoint set between the initial waypoint and the destination waypoint: $\mathcal{W} = \{w_m \mid \|\overrightarrow{w_m O}\| \le 2, m = 1, 2, \dots, W\}$

The mth waypoint w_m with coordinates (x_m, y_m) is generated from the (m-1)th waypoint w_{m-1} with coordinates (x_{m-1}, y_{m-1}) by

$$x_m = \begin{cases} x_{m-1} + \alpha_{m-1}\cos(\psi_{m-1} + \varphi_{m-1}), & \text{if } \|\overrightarrow{w_{m-1}w_d}\| > 0.3, \\ -x_0, & \text{if } \|\overrightarrow{w_{m-1}w_d}\| \le 0.3, \end{cases} \tag{2.14}$$

$$y_m = \begin{cases} y_{m-1} + \alpha_{m-1}\sin(\psi_{m-1} + \varphi_{m-1}), & \text{if } \|\overrightarrow{w_{m-1}w_d}\| > 0.3, \\ -y_0, & \text{if } \|\overrightarrow{w_{m-1}w_d}\| \le 0.3, \end{cases} \tag{2.15}$$

where $m = 1, 2, \dots, W$, W is the number of element of $\mathcal{W}$, the symbol $\|\overrightarrow{w_{m-1}w_d}\|$ denotes the length of a vector which initial point is w_{m-1} and the terminal point is w_d. The parameter α_{m-1} is the distance from the waypoint w_{m-1} to the waypoint w_m in unit km, which is defined as

$$\alpha_{m-1} = random[0.2, 0.3], \tag{2.16}$$

where z = random[x, y] represents that z is a random variable uniform distributed from 0.2 to 0.3. It illustrates that the vehicle will move a straight line distance between 0.2 km and 0.3 km, then start to make a turn. The parameter ψ_{m-1} is the angle between the vector $\overrightarrow{w_{m-1}w_d}$ and the unit vector of *x*-axis $\hat{i}$, which is calculated by

$$\psi_{m-1} = \begin{cases} \cos^{-1}\left(\dfrac{<\overrightarrow{w_{m-1}w_d}, \hat{i}>}{\left\|\overrightarrow{w_{m-1}w_d}\right\|}\right), & \text{if } (-y_0) - y_{m-1} \geq 0, \\ 360 - \cos^{-1}\left(\dfrac{<\overrightarrow{w_{m-1}w_d}, \hat{i}>}{\left\|\overrightarrow{w_{m-1}w_d}\right\|}\right), & \text{if } (-y_0) - y_{m-1} < 0, \end{cases} \tag{2.17}$$

where $<\overrightarrow{w_{m-1}w_d}, \hat{i}>$ is the inner product of the two vectors $\overrightarrow{w_{m-1}w_d}$ and $\hat{i}$. the parameter φ_{m-1} is an angle deviation from the vector $\overrightarrow{w_{m-1}w_d}$ to $\overrightarrow{w_{m-1}w_m}$ in unit degree, which is defined as

$$\varphi_{m-1} = random[-60, 60]. \tag{2.18}$$

The equation (2.18) means that the angle of a turn is limited in 60 degree. The constraint $\left\|\overrightarrow{w_m O}\right\| \leq 2$ indicates that the generated waypoint will be bound by the circle with radius 2 km. If the generated waypoint w_m is out of the boundary, it will generate again until satisfies the constraint.

Step 3: The bus moves along a straight line between each waypoint with a constant velocity *v*. After reaching each waypoint w_m, the bus will wait in a waiting time duration $U(w_m)$ before continuing moving to the next waypoint.

The bus will keep moving from the initial waypoint w_0, passing through W waypoints, then arriving at destination waypoint w_d.

In the proposed modified random waypoint mobility model, it is assumed that networks can get the information of the route map of each bus due to that the route map of each bus is usually fixed. According to this assumption, we can calculate the dwelling time of a bus in network *j* precisely.

It is assumed that a bus towards a waypoint w_k, and the group handover decision in the bus is happening at point g with coordinates (g_x, g_y). Firstly, we defined the waypoint set in the cell range of network *j* as

$$\mathcal{W}_j = \{w_m \mid \|\overrightarrow{n_j w_m}\| < \gamma_j, m = k, k+1, \ldots, W\}, \tag{2.19}$$

where n_j is the BS location of network *j* with coordinates (x'_j, y'_j) and γ_j is the BS cell radius of network *j*, $j = 1, 2, 3, 4$. Then we define a total point set in the cell range of network *j* as

$$\Omega_j = \{g\} \cup \mathcal{W}_j \equiv \{u_1, u_2, \ldots, u_{\omega_j}\}, \tag{2.20}$$

where the coordinates of u_κ is $(\tilde{x}_\kappa, \tilde{y}_\kappa)$, $\kappa = 1, 2, \ldots, \omega_j$, $u_1 = g$. The dwelling time of a bus in network *j* can be calculated by

$$T_j = \begin{cases} \dfrac{P_j}{v}, & \omega_j = 1, \\ \dfrac{1}{v}\left(\displaystyle\sum_{\kappa=1}^{\omega_j - 1} \|\overrightarrow{u_\kappa u_{\kappa+1}}\| + P_j\right) + \displaystyle\sum_{\kappa=2}^{\omega_j} U(u_\kappa), & \omega_j > 1, \end{cases} \tag{2.21}$$

where P_j is the path length in the cell range of network *j* from waypoint u_{ω_j} to the next waypoint $w_{k+|\mathcal{W}_j|}$ which is out of the cell range of network j, and is calculated by

$$P_j = \begin{cases} 0, & \text{if } u_{\omega_j} = w_d, \\ \|\overrightarrow{u_{\omega_j} n_j}\| \cos\theta_j + \sqrt{\gamma_j^2 + (\|\overrightarrow{u_{\omega_j} n_j}\| \sin\theta_j)^2}, & \text{if } u_{\omega_j} \neq w_d, \end{cases} \tag{2.22}$$

where θ_j is the angle between the vector $\overrightarrow{u_{\omega_j} n_j}$ and the vector $\overrightarrow{u_{\omega_j} w_{k+|\mathcal{W}_j|}}$ and is calculated by

$$\theta_j = \cos^{-1}\left(\frac{<\overrightarrow{u_{\omega_j} n_j}, \overrightarrow{u_{\omega_j} w_{k+|\mathcal{W}_j|}}>}{\left\|\overrightarrow{u_{\omega_j} n_j}\right\| \cdot \left\|\overrightarrow{u_{\omega_j} w_{k+|\mathcal{W}_j|}}\right\|}\right). \tag{2.23}$$

Fig. 2.5 shows three kinds of path situations when a bus passing through the cell range of network *j*.

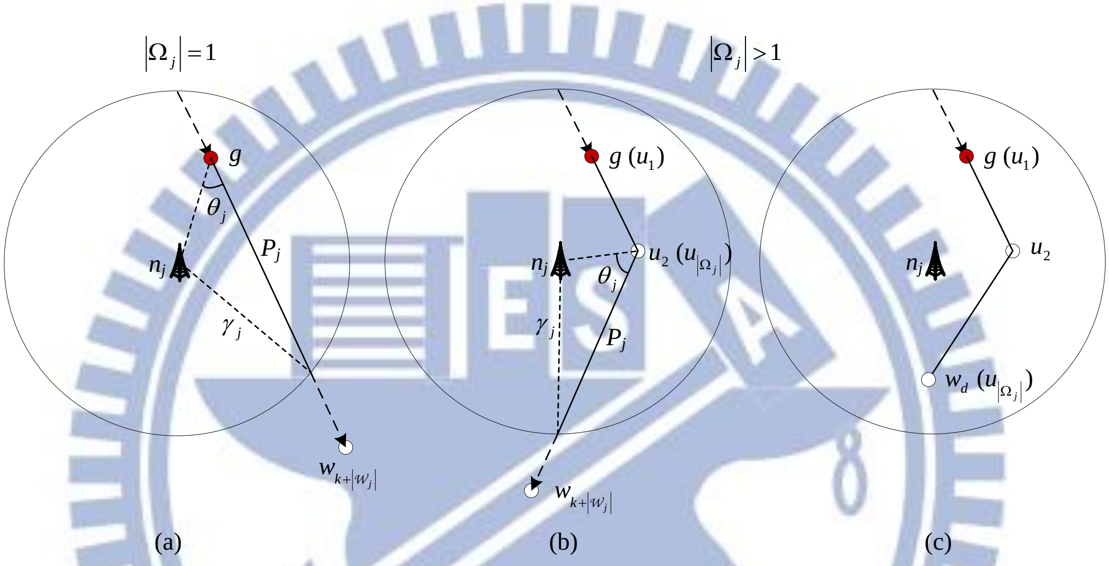


Fig. 2.6 Three types of Paths of a Vehicle in Network *j*

2.4 Traffic Class

We consider four service classes: conversational class (*h*=1), video streaming class (*h*=2), interactive class (*h*=3) and background class (*h*=4). The conversational class represents real-time (RT) multi-media applications such as voice over IP. The RT streaming class includes streaming type of applications, like video on demand. The interactive class is a non-real-time (NRT) service which composed of applications such as Web browsing (HTTP). The background class is the service using best effort transmission, such as file transfer protocol (FTP).

The QoS requirements of RT service are bit error rate, maximum delay tolerance, and maximum allowable dropping rate. The QoS requirements of NRT and BE services are bit error rate and minimum average transmission rate. For MT i with traffic class h, we donote $d_{i,h}, PDR_{i,h}$, and $R_{i,h}$ as the requirements of the transmission delay, the packet dropping rate, and the average transmission rate, respectively. If the transmission delay of RT service is beyond the maximum delay tolerance (d_h^*), the packet will be dropped. On the other side, the packets of NRT services will be queued in the buffer and it will be dropped if the buffer has no vacancy.

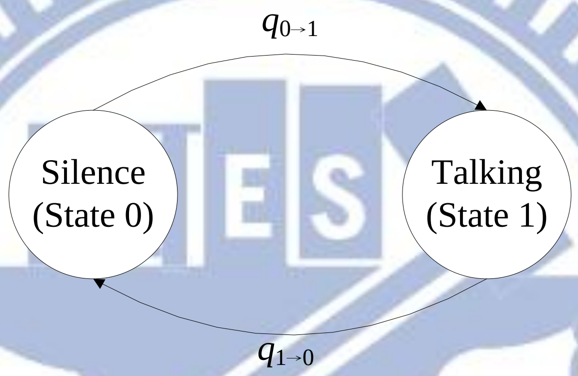


Fig. 2.7: Voice traffic model

The conversational class service is modeled as two-state ON-OFF (Silence-Talking) model [16] shown in Fig. 2.5. During ON period, voice packets are generated with D_v packet/sec. During OFF period, there is no packet generated. In addition, the periods of ON and OFF follow the exponential distributions with mean $\frac{1}{\alpha}$ and $\frac{1}{\beta}$. This model has a transition rate with value $q_{0\to1}$ in the ON state and a transition rate with value $q_{1\to0}$ in the OFF state.

Fig. 2.7 depicts the packet trace of one video streaming session model, which is composed of a sequence of video frames generated regularly with a constant interval T_f [17]. Each video frame consists of a fixed number of slices N_s, where each slice corresponds to a single packet. The size of packet is denoted by P_s, and the inter-arrival time between each packet is T_p.

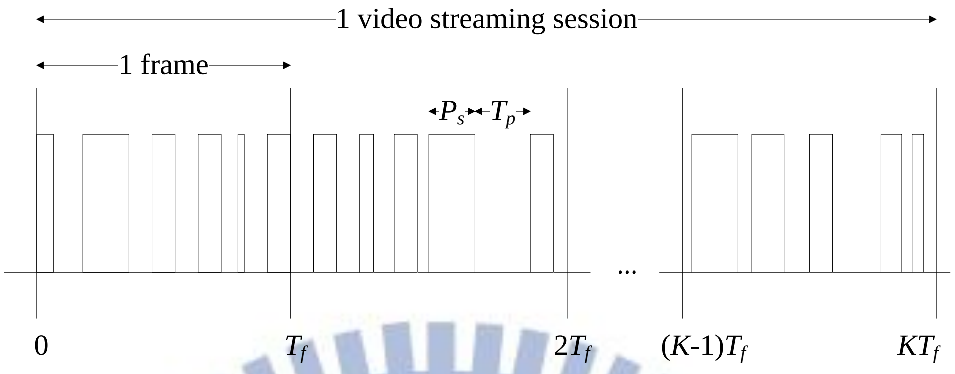


Fig. 2.8: Video traffic model

The interactive class service can be modeled as a sequence of packet calls (pages) as shown in Fig. 2.8. Each packet call consists of a sequence of packet arrivals, which is composed of a main object and several embedded objects [17]. Four parameters, including inter-arrival time, reading time $T_{reading}$, main object size S_m, embedded object size S_e, number of embedded objects per packet call N_e, and packet inter-arrival time T_p are used in this model.

The background class service is modeled as a sequence of file downloads [17] and is shown in Fig. 2.9. We denote the size of each file by S_f, and the inter-arrival time between each file by T_f.

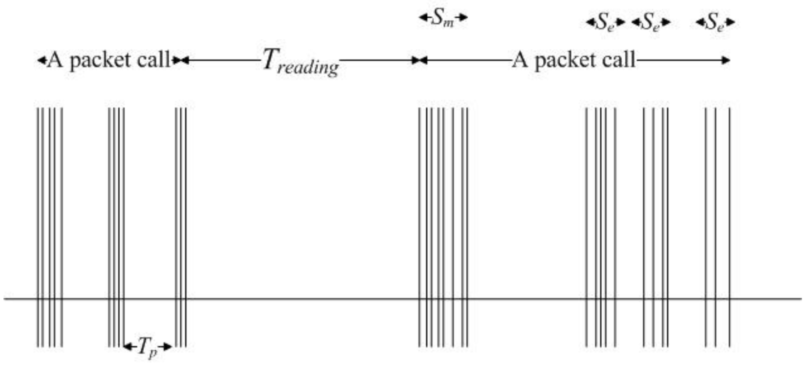


Fig. 2.9: HTTP traffic model

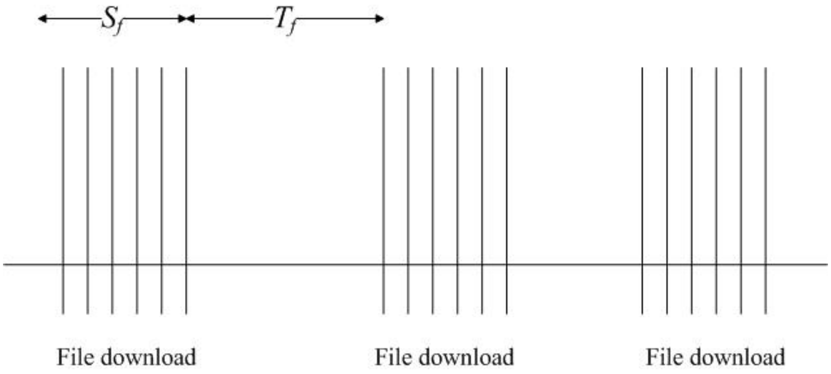


Fig. 2.10: FTP traffic model

2.5 Group Handover Scenario

In our group handover scenario, there have many ongoing MTs in a moving vehicle. While the vehicle is driving away from a serving BS of a serving network, the received signal power of all MTs that are served by this serving BS is decreasing. A handover group will be formed by these MTs. If a MT detects that its received signal power is less than a threshold, the MT will trigger the handover decision by sending the handover trigger message to the serving BS. Due to that MTs are close to each other in the vehicle, which means that the long-term channel condition of each MT is similar. The MTs in the handover group will send the handover trigger message at the same time or nearly simultaneously. After receiving these handover trigger messages, the serving BS is going to make a group handover decision.

We define the handover group set as $\mathcal{M}$. For the MT i, $i \in \mathcal{M}$, the required mean data rate is c_i kbps, the increased load of network j which caused by c_i is denoted by $\Delta\rho_{i,j}$, the communication life time is X_i which is modeled as an exponential distribution with mean $E[X_i]$, the head-of-line (HOL) packet delay is d_i ms, the packet dropping rate is p_i, and the average transmission rate is R_i kbps. The handover trigger message sent by MT i includes the candidate network set of MT *i*, which is denoted by $\mathcal{K}_i$. The MT *i* will scan and select one BS with largest pilot signal power in each network to represent each network. After selecting one BS in each network, the MT *i* will decide its candidate network set $\mathcal{K}_i$ by satisfying

$$PI_j \geq PI_j^* \tag{2.24}$$

where PI_j is the pilot signal power of network *j*, and PI_j^* is a required pilot signal power of network *j*. Besides, GSM network (*j*=1) can only serve the MTs with voice traffic class (*h*=1). In other words, GSM network cannot be included in the candidate network set of MTs with video, HTTP, and FTP traffic class (*h*=2, 3, 4). Therefore, the candidate network set of each MT *i* in the vehicle is given by

$$\mathcal{K}_i = \begin{cases} \{j \mid PI_j \geq PI_j^*, 1 \leq j \leq 4\}, & \text{for the MT } i \text{ with voice traffic,} \\ \{j \mid PI_j \geq PI_j^*, 2 \leq j \leq 4\}, & \text{for the MT } i \text{ with video, HTTP, FTP traffic} \end{cases} \tag{2.25}$$

During the group handover decision, the serving network will decide the target network for each MT from the candidate network set of each MT in the handover group. The target network decision result of MT *i* is denoted by s_i, $1 \leq i \leq |\mathcal{M}|$. All the s_i forms a target network decision vector which is denoted by $\vec{s}$, where $\vec{s} = [s_1\ s_2\ \ldots\ s_{|\mathcal{M}|}]$. If the serving network decides to let MT *i* handover to network *j*, the value of s_i will be set to *j*, *j*=1, 2, 3, 4. If the serving network decides to block the MT *i*, the value of s_i will be set to 0. The reason why a MT will be blocked by the serving network are that, there is no network that can satisfy the QoS requirements of the MT in the handover group, or there is no network that can provide enough loading capacity to accept the MT in the handover group.

Chapter 3

Simulated Annealing Based Group Handover Decision (SA-GHOD) Scheme

In this chapter, we propose a simulated annealing based group handover decision (SA-GHOD) scheme to solve the group handover decision problem. The goals of SA-GHOD scheme are to minimize the group handover blocking ratio, improve the load balancing of heterogeneous network, reduce the number of handovers, and provide a good QoS for each MT. The group handover decision problem will be formulated to an optimization equation first. Then we will introduce the simulated annealing (SA) algorithm. Finally, the simulated annealing based group handover decision (SA-GHOD) scheme will be proposed to solve the problem.

3.1 Problem Formulation

In the group handover decision problem, we need to select the appropriate target network for each MT in the handover group. Firstly, we consider the group handover blocking ratio in the handover group. The handover blocking will occur when all the candidate networks of a MT cannot provide enough loading capacity, or all the candidate networks of a MT cannot afford the QoS requirements of this MT. The serving network will calculate the value of group handover blocking ratio before making the group handover decision. If the group handover blocking ratio is high, many MTs in the handover group will be blocked by the serving network.

Secondly, we consider the balance of loading between every target networks. From the view point of one operator who own the four different wireless networks, if a network is overloaded and cannot accept any new MT while other network at the same area is underutilized, it will affect the operator's revenue. Load balancing among networks will improve radio resource availability and provide better quality of service for MTs as well.

Then, frequently performing handover may cause some severe disadvantages in the perspective of single MT. For example, it leads to heavy processing loads, causes more delay in the processing of handover requests, and increases the chance that a call will be denied access by a BS. In order to reduce the circumstance of frequent handover, a BS with large cell range will be considered in the group handover decision.

Finally, we consider that guaranteeing the quality of service (QoS) for each MT. Because there are four different traffic classes of MTs in the handover group, the group handover decision needs to satisfy the requirements of different kind of MTs. The serving BS will communicate with the candidate networks of each MT to get the quality of candidate networks, including the average packet delay, the average packet dropping rate, and the average transmission rate. Therefore, the serving BS can select a target network which may provide a better QoS according to these quality parameters.

The group handover decision problem is formulated as follows

$$\vec{s}^{*} = \arg\min_{\vec{s}} \ f(\vec{s}), \tag{3.1}$$

subject to the QoS constraints:

$$\rho_{s_i} + \Delta\rho_{i,s_i} \le \rho_{s_i}^{th}, \quad \forall i \in \mathcal{M}, s_i \in \mathcal{K}_i,$$
$$d_{s_i,h} \le d_h^*, p_{s_i,h} \le p_h^*, \quad \text{for MT} \, i \text{ with traffic type} \, h = 1 \text{ or } 2, s_i \in \mathcal{K}_i,$$
$$R_{s_i,h} \ge R_h^*, \quad \text{for MT} \, i \text{ with traffic type} \, h = 3 \text{ or } 4, s_i \in \mathcal{K}_i,$$

and the system constraint:

$$\rho_j + \sum_{i \in \mathcal{M}_j} \Delta\rho_{i,j} \le \rho_j^{th}, \; j = 1, 2, 3, 4,$$

where $f(\vec{s})$ is the cost function, which will be defined in section 3.1.1. $\vec{s}^*$ is the optimal solution of the group handover decision problem. The $d_{s_i,h}$, $p_{s_i,h}$, and $R_{s_i,h}$ are the average voice/video packet delay, the average voice/video packet dropping rate, and the average HTTP/FTP transmission rate of network s_i, respectively. The ρ_{s_i} is the load intensity of network s_i. During the group handover decision, the serving network will communicate with all candidate networks to get the above four information. The d_h^* is the upper bound of the HOL packet delay of the MT with traffic class *h*. The p_h^* is the upper bound of the packet dropping rate of the MT with traffic class *h*. The R_h^* is the minimum required average transmission rate of the MT with traffic class *h*. The $\mathcal{M}_j$ is the handover group subset, where $\mathcal{M}_j = \{i | s_i = j, i \in \mathcal{M}\}$. The handover group subsets $\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3$, and $\mathcal{M}_4$ contain the MTs that are going to handover to network 1, 2, 3 and 4, respectively. The MTs in subset $\mathcal{M}_0$ will be blocked by the serving network.

The constraint $\rho_{s_i} + \Delta\rho_{i,s_i} \le \rho_{s_i}^{th}$ is to assure that the target network of the MT *i* in the handover group must have enough capacity to accommodate it. For the MT *i* with voice or video traffic (*h* = 1 or 2) in the handover group, we hope that the MT have enough QoS after group handover. Therefore, the average packet delay of the target network needs lower than the delay requirement of the MT ($d_{s_i,h} \le d_h^*$) and the average packet loss rate of the target network needs lower than the requirement of packet loss rate of the MT ($p_{s_i,h} \le p_h^*$). For the MT with HTTP or FTP traffic in the handover group, the joinable network constraint $R_{s_i} \ge R_h^*$ is to assure that the average transmission rate of the target network is larger than the minimum transmission rate requirement of the MT. The network loading constraint illustrates that the load intensity of every network cannot overloaded after the group handover.

3.1.1 Cost Function

For a given $\vec{s}$, it has a corresponding cost function $f(\vec{s})$ which is defined as

$$f(\vec{s}) = B(\vec{s}) + W(\vec{s}) + L(\vec{s}), \tag{3.2}$$

where $B(\vec{s})$ denotes a group handover blocking factor, $W(\vec{s})$ denotes a dwelling time factor, and $L(\vec{s})$ denotes a load balancing factor. The contents of these three factors will be described below.

(1) Group Handover Blocking Factor

For a given target network decision vector $\vec{s}$, we can get the group handover blocking ratio directly by counting how many MTs in the handover group are decided to be blocked over the total number of MTs in the handover group. Hence, the group handover blocking ratio, denoted by $R_B(\vec{s})$, is defined as

$$R_B(\vec{s}) = \frac{|\mathcal{M}_0|}{|\mathcal{M}|}, \tag{3.3}$$

where $\mathcal{M}_0 = \{i | s_i = 0, i \in \mathcal{M}\}$. If the value of $R_B(\vec{s})$ is higher than a predefined upper bound R_B^*, we want that the effect of $R_B(\vec{s})$ is larger than that of the other three parameters in the cost function. Therefore, a group handover blocking factor is defined as

$$B(\vec{s}) = \begin{cases} \dfrac{R_B(\vec{s}) - R_B^*}{1 - R_B^*} + 3, & R_B^* < R_B(\vec{s}) \le 1, \\ R_B(\vec{s}), & 0 \le R_B(\vec{s}) \le R_B^*, \end{cases} \tag{3.4}$$

where R_B^* is a predefined upper bound for the group handover blocking ratio. Fig. 3.1 shows the design concept of $B(\vec{s})$ when R_B^* is set to 0.05. When $R_B(\vec{s}) > R_B^*$, we hope that the effect of $R_B(\vec{s})$ to the cost function is higher than that of the other maximum value of $W(\vec{s})$, and $L(\vec{s})$.

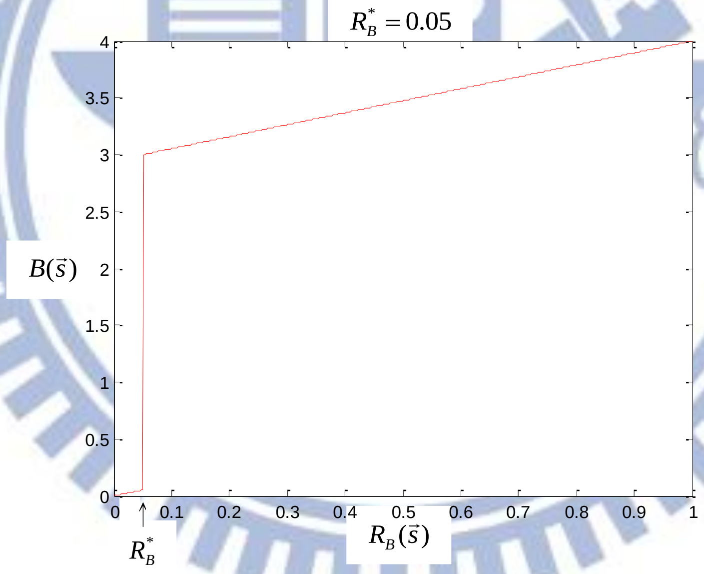


Fig. 3.1: Group Handover Blocking Ratio Function

(2) Dwelling Time Factor

For each target network decision result s_i in a given target network decision vector $\vec{s}$, the dwelling time ratio of MT *i*, denoted by α_{i,s_i}, is defined as

$$\alpha_{i,s_i} = \min\{\frac{T_{s_i}}{E[X_i - t | X_i > t]}, 1\} \tag{3.5}$$

where T_{s_i} is the dwelling time of the vehicle in network s_i. $E[X_i - t | X_i > t]$ is the mean residual communication life time of MT *i*, which is equal to $E[X_i] - t$ since the MT *i*'s communication life time X_i is modeled as an exponential distribution with mean $E[X_i]$. Considering all MTs in the handover group, the average dwelling time ratio, denoted by $\overline{\alpha}(\vec{s})$, is given by

$$\overline{\alpha}(\vec{s}) = \frac{1}{|\mathcal{M} - \mathcal{M}_0|} \sum_{i \in \mathcal{M} - \mathcal{M}_0} \alpha_{i,s_i} \tag{3.6}$$

and the variance of dwelling time ratio, denoted by $\beta(\vec{s})$, is given by

$$\beta(\vec{s}) = \sqrt{\frac{1}{|\mathcal{M} - \mathcal{M}_0|} \sum_{i \in \mathcal{M} - \mathcal{M}_0} (\alpha_{i,s_i} - \overline{\alpha}(\vec{s}))^2}. \tag{3.7}$$

We hope that MTs in the handover group can handover to a network with large cell range in order to reduce the number of handover. Therefore, the handover group tend to have a large value of $\overline{\alpha}(\vec{s})$. In addition, When there has many $\vec{s}$ with the same value of $\overline{\alpha}(\vec{s})$, we hope that the dwelling time ratio of each MT in the handover group is closed with each other. It means that the handover group tend to have a small value of $\beta(\vec{s})$. Namely, the large value of $\overline{\alpha}(\vec{s})$ with the small value of $\beta(\vec{s})$ can reduce the number of handover of the whole group. We define the dwelling time factor as

$$W(\vec{s}) = e^{-\overline{\alpha}(\vec{s}) + \beta(\vec{s})}. \tag{3.8}$$

(3) Load Balancing Factor

Based on a given target network decision vector $\vec{s}$, we can find the prospective load intensity of candidate network *j*, denoted by $\rho'_j(\vec{s})$, is calculated by

$$\rho'_j(\vec{s}) = \rho_j + \sum_{i \in \mathcal{M}_j} \Delta\rho_{i,j}, \tag{3.9}$$

where $\mathcal{M}_j = \{i | s_i = j, i \in \mathcal{M}\}$, $1 \le j \le 4$. Then we can calculate the average load intensity of all candidate networks, which is denoted by $\bar{\rho}'$ and is given by

$$\bar{\rho}' = \frac{1}{4}\sum_{j=1}^{4} \rho'_j(\vec{s}). \tag{3.10}$$

Finally, we define the load balancing factor as

$$L(\vec{s}) = \sqrt{\frac{1}{4}\sum_{j=1}^{4} (\rho'_j(\vec{s}) - \bar{\rho}')^2}. \tag{3.11}$$

The value of $L(\vec{s})$ is between 0 to1. If the value of load balancing factor is close to 0, it means that the loading between candidate networks are very close.

3.2 Simulated Annealing Algorithm

The annealing process is heating the solid until it is melted and then slowly cooling the solid to obtain a strong crystalline structure. In 1953, Metropolis et al. proposed a method for computing the equilibrium distribution of a set of particles of a solid in a heat bath [18]. Simulated annealing derives from the analogy between the process of annealing a solid and the process of finding minimal solutions for minimization problems, which is linked by Kirkpatrick et al. in 1983 [19].

The simulated annealing algorithm is formed by three parts: the initialization part (line 1 to line 2), the generation and acceptance part (line 5 to line 8), and the cooling schedule part (line 3, line4, and line11).

The initialization part is to set an initial value of the temperature which is denoted by T and generate an initial solution $\vec{s}$ by the GenerateInitialSolution() function. The temperature T is a control parameter which will affect the probabilistic decision in the generation and acceptance part. The initial temperature needs to be high enough to accept any new solution which is generated from a old solution. Under the initial temperature, the GenerateInitialSolution() will generate an initial solution which is common to be a random initial solution and let the simulated annealing process improve on that. However, due to the random solution might has a long searching time, it is better to start with a useful solution which can reduce the searching time. For a solution $\vec{s}$, there must be some way of measuring the quality of the solution, which is the cost function $f(\vec{s})$. If possible, the cost function should be designed so that it can lead the search. One way of achieving this is to avoid cost functions where many states return the same value.

In the generation and acceptance part, first is to generate a new solution which is denoted by $\vec{s}'$ from the old solution $\vec{s}$ by the function Generate(). In order to reduce the time of calculation, it is common that generate a new solution by changing the result of some elements of old solution. After generating a new solution, the function Accept() will calculate the difference of the cost functions between the new solution and the old solution, and determine whether the new solution will be accepted or not. Here we use the common criteria called the Metropolis criteria [18]. If $f(\vec{s}')$ is lower or equal to $f(\vec{s})$, the old solution will be replaced by the new solution. However, if $f(\vec{s}')$ is larger than $f(\vec{s})$, it still may accept the solution with a probability $\exp\{-[f(\vec{s}')-f(\vec{s})]/T\}$. The significant advantage of the Metropolis criteria is that any neighbor solution of the old solution can be accepted as a new solution when the temperature is high. It could avoid the new solution trapping in a local minimum solution. For example, it can first generate and accept a worse new solution from an old solution, then generate a best solution by the worse solution because the best solution is the neighbor of the worse solution, but it is not a neighbor of the old solution.

The cooling scheduling part is in charge of deciding the iteration criterion, deciding how to decrease the temperature, and deciding the stopping criterion. The iteration criterion is the number of iterations at each temperature. A common method will be adopted by setting the number of iterations to a constant which is denoted by L. Then the geometric decrement will be used in the function Decrease(), that is, Decrease(T)=$\gamma \cdot T$, where $0 < \gamma < 1$. The way in which we decrease the temperature and decide the value of L are important due to that they decide whether the algorithm can find the minimum solution and how long it will take. We can either doing a large number of iterations at a few temperatures, a small number of iterations at many temperatures or a balance between the two. Therefore, the L and γ will be tested and decided in simulation. The stopping criterion can either be a suitably low temperature or when the system is "frozen" at the current temperature, i.e., no better or worse moves are being accepted. The pseudo code of the simulated annealing algorithm is shown below.

Simulated Annealing Algorithm

```
output: $\vec{s}_{best}$
set initial temperature $T$.
$\vec{s} \leftarrow$ GenerateInitialSolution(),
$\vec{s}_{best} \leftarrow \vec{s}$.
while stopping criterion do
    while iteration criterion do
        $\vec{s}' \leftarrow$ Generate($\vec{s}$).
        $\vec{s} \leftarrow$ Accept($T, \vec{s}, \vec{s}'$).
        if $f(\vec{s}) < f(\vec{s}_{best})$ then // Record the output solution
            $\vec{s}_{best} \leftarrow \vec{s}$ .
        end if
```

```
        end while
        $T \leftarrow$ Decrease($T$).
end while
return $\vec{s}_{best}$
end
```

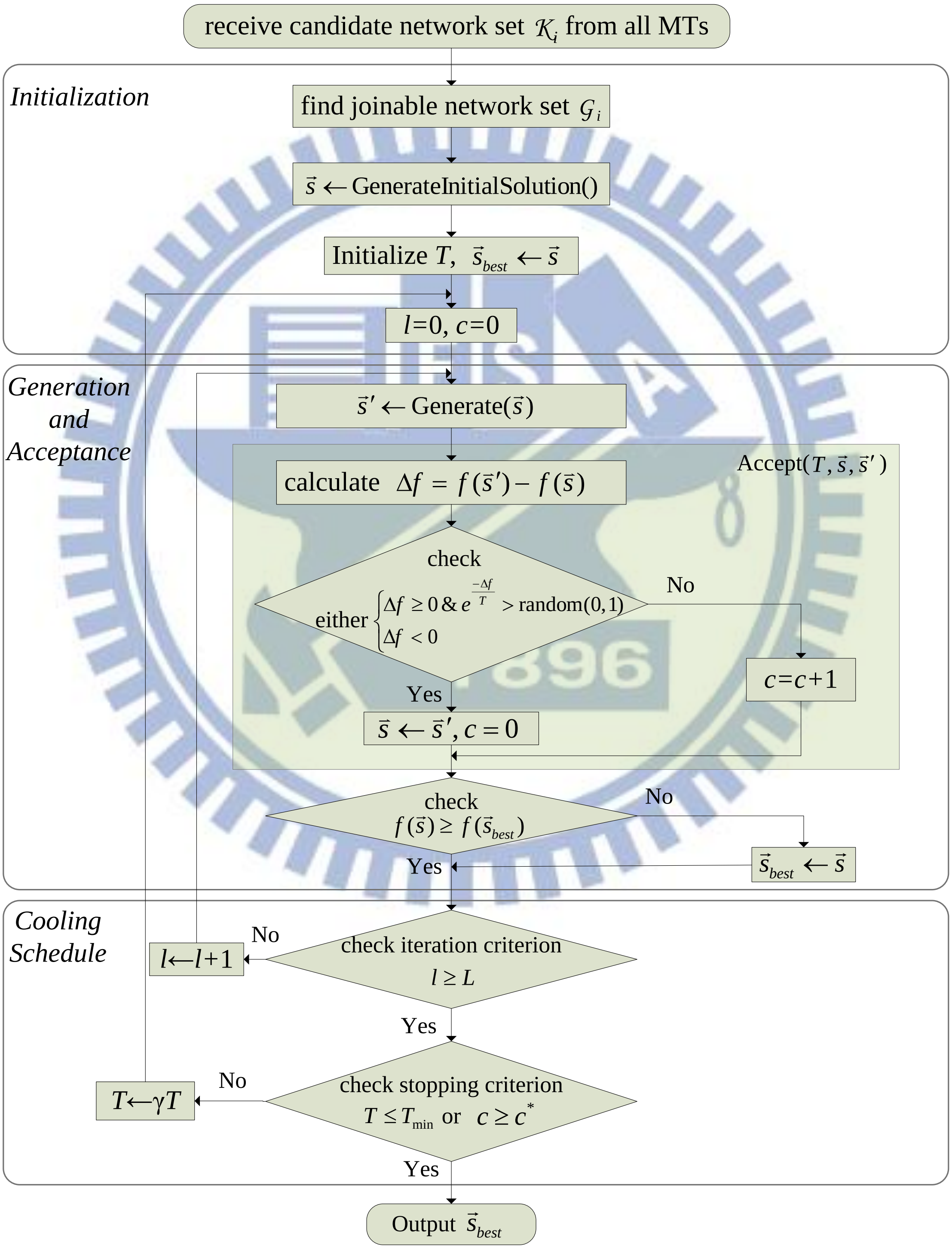


Fig. 3.2: Flow Chart of the SA-GHOD Scheme

3.3 SA-GHOD Scheme

Fig. 3.2 shows a flowchart of the SA-GHOD scheme. The scheme includes three parts: initialization, generation and acceptance, and cooling schedule. The initialization is to find a good initial solution which satisfies the joinable network constraints of the optimization equation (3.1). The generation and acceptance shows how to find and accept a new solution in our problem. The cooling schedule is used to let the final solution with minimal value of the cost function. The detail will be described in the next three subsections.

3.3.1 Initialization

The initialization part is going to initialize the initial solution $\vec{s}$ and the starting temperature $T_{\max}$.

First of all, the serving network will find a joinable network set $\mathcal{G}_i$ of MT *i* from the candidate network set $\mathcal{K}_i$ by satisfying the joinable network constraints (line 1). The joinable network constraints are calculated by

$$\mathcal{G}_i = \begin{cases} \{ j \mid d_{j,h} \le d_h^*,\ p_{j,h} \le p_h^*, \text{and } \rho_j + \Delta\rho_{i,j} \le \rho_j^{th},\ j \in \mathcal{K}_i, h = 1, 2\}, \\ \{ j \mid R_{j,h} \le R_h^*, \text{and } \rho_j + \Delta\rho_{i,j} \le \rho_j^{th},\ j \in \mathcal{K}_i, h = 3, 4\}. \end{cases} \tag{3.15}$$

When a candidate network of MT *i* satisfies the joinable network constraints, it means that this candidate network can guarantee the QoS for the MT *i* and afford the increased load of MT *i*. In addition, it can reduce the searching space of $\vec{s}$.

Then the function GenetateInitialSolution() will generate the initial solution $\vec{s}$ by repeating the following three steps until all MTs in the handover group match their target networks.

Step 1: Selecting a MT from the handover group by the MT selection strategy (line 2). The MT selection strategy selects a MT *i* with the smallest required mean data rate c_i. If there are more than one MT have the same smallest required mean data rate in the handover group, selecting a MT from these MTs randomly. The reason why selecting the MT with smallest c_i is that, it tend to let more MTs in the handover group handover successfully in order to have a minimal value of the group handover blocking ratio $B(\vec{s})$ when the total resource of all joinable networks is not enough.

Step 2: For the selected MT $\hat{i}$, deciding its target network $s_{\hat{i}}$ by the network decision strategy (line 9). The network decision strategy selects a network with the smallest value of a price function from the joinable network set of MT $\hat{i}$. The price function of MT $\hat{i}$ is composed by three factors of the cost function (3.2) and is denoted by

$$P_{\hat{i}}(s_{\hat{i}}) = W(s_{\hat{i}}) + L(s_{\hat{i}}) \tag{3.16}$$

where the $W(s_{\hat{i}})$ is the dwelling time factor when there is only one MT $\hat{i}$ in the handover group. The $L(s_i)$ is the load balancing factor when there is only one MT $\hat{i}$ in the handover group. If there is only one MT $\hat{i}$ in the handover group and is decided to handover to the network $s_{\hat{i}}$ ($\vec{s} = s_{\hat{i}}$), the dwelling time factor $W(s_{\hat{i}})$ in (3.8) will be the dwelling time ratio $\alpha_{\hat{i}, s_{\hat{i}}}$ duo to that the average dwelling time ratio $\overline{\alpha}(s_{\hat{i}})$ is $\alpha_{\hat{i}, s_{\hat{i}}}$ and the variance of dwelling time ratio $\beta(s_{\hat{i}})$ is 0. The load balancing factor will be $L(s_{\hat{i}})$ in (3.11), where $\overline{\rho}' = \frac{1}{4}\sum_{j=1}^{4} \rho_j'(s_{\hat{i}})$ and $\rho_j'(s_{\hat{i}}) = \begin{cases} \rho_j + \Delta\rho_{\hat{i}, j}, & j = s_{\hat{i}} \\ \rho_j, & j \neq s_{\hat{i}} \end{cases}$. Due to that we hope to find a solution to minimize the cost function, therefore, for a MT *i*, selecting a network *j* with minimum value of

$P_i(j)$ will tend to minimize the cost function in the initialization part. The network selection result s_i of MT *i* will be

$$s_i = \arg\min_{j\in\mathcal{G}_i} P_i(j), \tag{3.17}$$

Step 3: Updating the load intensity of the selected network in step 2. If there have MT that haven't decide their target network, go back to step 1. However, in the process of generating the initial solution, matching a MT to a network at a time cannot assure that the value of the cost function is minimal. For this reason, the simulated annealing algorithm is used to check whether there have better solutions or not.

In the end of the initialization, One way to set the initial temperature T to be the high value of the cost function (line 4). In addition, two index l and c will be set. The index l is the current number of iterations and is set to 1 initially. The index c is the cumulative number of times that the current solution is not replaced by a new solution and is also set to be zero initially. The pseudo code of this part is described as below.

Initialization

```
input: $\mathcal{K}_i$
output: $l$, $c$, $T$, $\vec{s}$, $\vec{s}_{best}$
1: Calculate $\mathcal{G}_i$ for all MT $i$ by (3.15)
2: $\vec{s} \leftarrow$ GenerateInitialSolution()
3: $\vec{s}_{best} \leftarrow \vec{s}$.
4: Set $T$= 5, $c$=0, $l$=1.
5: end
```

GenerateInitialSolution()

```
Output: $\vec{s}$
1: for $a \leftarrow 1$ to $|\mathcal{M}|$ do   // Generate $\vec{s}$; $a$ is a index
```

```
    $\hat{i} \leftarrow \arg\min_{i\in\mathcal{M}} c_i$.  // Select a MT with minimum $c_i$.
        for $j \leftarrow$ 1 to $|\mathcal{G}_{\hat{i}}|$ do  // Update the joinable network set of MT i.
            if $\rho_j + \Delta\rho_{\hat{i},j} \geq 1$ then
                $\mathcal{G}_{\hat{i}} \leftarrow \mathcal{G}_{\hat{i}} - \{j\}$.
            end if
        end for
    if $\mathcal{G}_{\hat{i}} \neq \phi$ then
        $s_{\hat{i}} \leftarrow \arg\min_{j\in\mathcal{G}_i} P_i(j)$  // MT $\hat{i}$ Select a network with minimum $P_{\hat{i}}(j)$.
        $\rho_{s_{\hat{i}}} \leftarrow \rho_{s_{\hat{i}}} + \Delta\rho_{\hat{i},s_{\hat{i}}}$.  //Update the load intensity of the selected network.
    else  // if MT $\hat{i}$ has no joinable network.
        $s_{\hat{i}} = 0$.  // MT $\hat{i}$ will be blocked
    end if
    $\mathcal{M} \leftarrow \mathcal{M} - \{\hat{i}\}$.
end for
```

3.3.2 Generation and Acceptance

In this part, a new solution will be generated from the old solution (initial solution at the first time), compare the value of cost function between the new solution and the old solution, and decide whether to accept the new solution or keep the old solution.

First of all, the function Generate() will generate the new solution $\bar{s}'$ from the old solution $\bar{s}$ by selecting two MTs to exchange their network decision with each other. The detail is shown in the following steps.

Step 1: Selecting a MT i_1 from the handover group set $\mathcal{M}$ randomly.

Step 2: Keep selecting another MT i_2 from the rest of MTs in the handover group set $(\mathcal{M} - \{i_1\})$ randomly until satisfying $s_{i_2} \neq s_{i_1}$. Because changing the target network of two MTs with the same target network decision result in the old solution is useless.

Step 3: Calculating a changing index which is denoted by *k* and is defined as

$$k=\begin{cases}1,\ \text{if } (s'_{i_1}=0,\ s'_{i_2}\in\mathcal{G}_{i_1},\text{ and } \rho_{s'_{i_2}}-\Delta\rho_{i_2,s'_{i_2}}+\Delta\rho_{i_1,s'_{i_2}}\le 1)\text{ or}\\ \qquad (s'_{i_2}=0,\ s'_{i_1}\in\mathcal{G}_{i_2},\text{ and } \rho_{s'_{i_1}}-\Delta\rho_{i_1,s'_{i_1}}+\Delta\rho_{i_2,s'_{i_1}}\le 1)\text{ or}\\ \qquad (s'_{i_2}\in\mathcal{G}_{i_1},\ s'_{i_1}\in\mathcal{G}_{i_2},\ \rho_{s'_{i_2}}-\Delta\rho_{i_2,s'_{i_2}}+\Delta\rho_{i_1,s'_{i_2}}\le 1,\text{ and } \rho_{s'_{i_1}}-\Delta\rho_{i_1,s'_{i_1}}+\Delta\rho_{i_2,s'_{i_1}}\le 1),\\ 0,\ \text{otherwise.}\end{cases}\tag{3.18}$$

When $k=1$, it means that the two selected MTs can exchange their network selection results. When $k=0$, it means they cannot exchange network selection results. There are three situations that the exchange will happen. Two of the situations are that if one of the two selected MTs is set to be blocked by the serving network ($s'_{i_1}=0$ or $s'_{i_2}=0$), they will exchange their network decision if the nonzero value of network decision result is a joinable network of the MT with zero value of network decision result ($s'_{i_2}\in\mathcal{G}_{i_1}$ or $s'_{i_1}\in\mathcal{G}_{i_2}$) and the network load intensity is not full after the exchange ($\rho_{s'_{i_2}}-\Delta\rho_{i_2,s'_{i_2}}+\Delta\rho_{i_1,s'_{i_2}}\le 1$ or $\rho_{s'_{i_1}}-\Delta\rho_{i_1,s'_{i_1}}+\Delta\rho_{i_2,s'_{i_1}}\le 1$). The other situation is that if the two selected MTs select different networks in the old solution, they will exchange their network decision if their network decision results is the candidate network of each other ($s'_{i_2}\in\mathcal{G}_{i_1}$ and $s'_{i_1}\in\mathcal{G}_{i_2}$) and the load intensities of two networks are not full after exchange ($\rho_{s'_{i_2}}-\Delta\rho_{i_2,s'_{i_2}}+\Delta\rho_{i_1,s'_{i_2}}\le 1$ and $\rho_{s'_{i_1}}-\Delta\rho_{i_1,s'_{i_1}}+\Delta\rho_{i_2,s'_{i_1}}\le 1$). If the two selected MTs cannot exchange their target network decision results, go back to step 2 to select another MT. If all the MTs in step 2 cannot match the MT which is selected in step 1, go back to step 1 to select another MT.

Step 4: Find the blocked MT in old solution except i_1 and i_2, then try to join the network s'_{i_1} and s'_{i_2}.

After that, the function Accept() will calculate the difference of the cost function between the new solution and the old solution, which is denoted by Δf. Then it will use the Metropolis criteria [18] to determine whether the new solution $\bar{s}'$ will be accepted or not. That is, the new solution $\vec{s}'$ will be accepted directly if $\Delta f < 0$ and the new solution $\vec{s}'$ will also be accepted with probability $e^{-\frac{\Delta f}{T}}$ if $\Delta f \geq 0$. If the new solution is accepted, the value of index c will be set to zero. Otherwise, the value of index c will plus 1.

In the end of this part, the output solution $\vec{s}_{best}$ will be set to the new solution if the value of the cost function of the new solution is less than the old solution. On the contrary, the output solution $\vec{s}_{best}$ will keep the old solution if the value of the cost function of the new solution is larger than the old solution. The pseudo code of this part is described as below.

Generation and Acceptance

```
Input: $c$, $T$, $\vec{s}$, $\vec{s}_{best}$
Output: $c$, $T$, $\vec{s}_{best}$
$k \leftarrow -1$.
$\mathcal{M}_1 \leftarrow \mathcal{M}, \mathcal{M}_2 \leftarrow \mathcal{M}$.
$\vec{s}' = [s'_1\ s'_2 \ldots s'_{|\mathcal{M}|}] \leftarrow \vec{s} = [s_1\ s_2 \ldots s_{|\mathcal{M}|}]$.
$\vec{s}' \leftarrow$ Generate( $\vec{s}$ )
$\vec{s} \leftarrow$ Accept( $T, \vec{s}, \vec{s}'$ )
if $f(\vec{s}) < f(\vec{s}_{best})$ then
    $\vec{s}_{best} \leftarrow \vec{s}$ .
end if
```

Generate()

```
Input: $\vec{s}$
Output: $\vec{s}'$
for $i \leftarrow 1$ to $|\mathcal{M}|-1$ do
    $i_1 \leftarrow$ random_select ( $\mathcal{M}_1$ ). // Select a MT $i_1$ from $\mathcal{M}_1$ randomly
    $\mathcal{M}_1 \leftarrow \mathcal{M}_1 - \{i_1\}, \mathcal{M}_2 \leftarrow \mathcal{M}_1$.
    while $\mathcal{M}_2 \neq \phi$ do
        $i_2 \leftarrow$ random_select( $\mathcal{M}_2$ ). // Select a MT $i_2$ from $\mathcal{M}_2$ randomly
        $\mathcal{M}_2 \leftarrow \mathcal{M}_2 - \{i_2\}$.
        if $s'_{i_1} \neq s'_{i_2}$ then
```

```
            calculate $k$ by (3.18).
            break
        end if
    end while
    if $k = 1$ then
        $temp \leftarrow s'_{i_1}$.
        $s'_{i_1} \leftarrow s'_{i_2}$.
        $s'_{i_2} \leftarrow temp$.
    end if
end for
```

Accept()

```
Input: $T, \vec{s}, \vec{s}'$
Output: $\vec{s}$
$\Delta f \leftarrow f(\vec{s}') - f(\vec{s})$.
if $\Delta f < 0 \,\|\, (\Delta f \geq 0 \,\&\, e^{-\frac{\Delta f}{T}} > random(0,1))$ then
    $\vec{s} \leftarrow \vec{s}'$.
else
    $c \leftarrow c + 1$.
end if
```

3.3.3 Cooling Schedule

In this part, it will decide and check the iteration criterion and the stopping criteria. The iteration criterion is that if the current number of iteration l exceeds a upper bound L, the iteration will stop. The value of L is set to 100 and is same in every temperature. The stopping criteria are that if the cumulative number of times that the current solution is not replaced by a new solution is larger than a upper bound value c^* or the current temperature is lower than a lower bound temperature which is denoted by $T_{\min}$, where $T_{\min}$ = 0.1. The function Decrease() is set to Decrease(T)= γT. where γ =0.8. If the current number of iteration l at temperature T reaches L, the temperature will decrease by the decrement rate γ until the temperature lower the temperature $T_{\min}$. If the stopping criteria are satisfied, output the best solution $\vec{s}_{best}$. The pseudo code is described as following:

Cooling Schedule

```
Input: $c$, $l$, $\vec{s}_{best}$
Output: $c$, $l$, $T$, $\vec{s}_{best}$
If $l \le L$ then
    $l \leftarrow l + 1$.
    go to Generation and Acceptance
else
    If $T \le T_{\min} \| c \le c^*$ then
        $T \leftarrow \gamma T$ . // Decrease($T$) = $\gamma T$
        go to Initialization
    else
        output the solution $\vec{s}_{best}$
    end if
end if
```

Chapter 4
Simulation Results

4.1 Simulation Environment

The simulation environment is shown in Fig. 2.1 which includes 7 GSM/EDGE cells, 19 WCDMA cells, and 27 LTE cells. Table 4.1 shows the system parameters of this heterogeneous environment.

Table 4.1: System parameters for GSM/EDGE, WCDMA, and LTE

Parameters	GSM/EDGE	WCDMA	LTE
Inter Site Distance (ISD)	1 Km	0.675 Km	0.5 Km
Frame duration	4.615 ms	10 ms	10 ms
Carrier frequency	1.8 GHz	2.1 GHz	2 GHz
Total bandwidth per cell	3.7 MHz	5 MHz	3 MHz
Capacity	4 Mbps	7.2 Mbps	13 M bps
Number of cells	7	19	27
Maximum load intensity ρ_j^{th}	1/0.9	0.75	0.9
Number of channels	56		
Number of channels in a cell	18		
Chip rate (*W*)		3.84 Mbps	
Ratio of inter-cell interference to the total		0.55	
Number of subchannels			15
Number of resource blocks in a frame			150

4.2 Traffic Source Model and QoS Requirements

As described at chapter 2, there are four traffic classes considered. The source model parameters for voice, video, http, and ftp traffic classes are shown in Table 4.2, 4.3, 4.4, and 4.5, respectively.

Table 4.2: Parameters for voice traffic class

Component	Distribution	Parameters
ON time	Exponential	Mean = 1 sec
OFF time	Exponential	Mean = 1.35 sec
Packets per second	Deterministic	50
Packet size	Deterministic	28 Bytes(B)
Call holding time	Normal	Mean = 90 sec
Data during active period		11.2 kbps
Active rate		0.426
Mean data rate		4.77 kbps

Table 4.3: Parameters for video traffic class

Component	Distribution	Parameters
Inter-arrival time between each video frame	Deterministic	100 ms
Number of packets in each video frame	Deterministic	8
Inter-arrival time between packets in a frame	Truncated Pareto	Min = 2.5 ms, Max = 12.5 ms, Mean = 6 ms, α = 1.2
Packet size	Truncated Pareto	Min = 40 B, Max = 250 B, Mean = 100 B, α = 1.2
Call holding time	Normal	Mean = 120 Variance = 30 (sec)
Data during active period		133.33 kbps
Active rate		0.48
Mean data rate		64 kbps

Table 4.4: Parameters for HTTP traffic class

Component	Distribution	Parameters
Main object size (S_m)	Truncated lognormal	Min = 100 B, Max = 2 MB, Mean = 10.7 kB, STD = 25 kB
Embedded object size (S_e)	Truncated lognormal	Min = 50 B, Max = 2 MB, Mean = 7758 B, STD = 126 kB
Number of embedded objects per page (N_e)	Truncated Pareto	Min = 5.64, Max = 53
Inter-arrival time between each page ($T_{reading}$)	Exponential	Mean = 30 sec
Packet inter-arrival time (T_p)	Exponential	Mean = 0.13 sec
Packet size	Deterministic	1500 B
Call holding time	Normal	Mean = 120 Variance = 30 (sec)
Data during active period		92.3 kbps
Active rate		0.136
Mean data rate		12.55 kbps

Table 4.5: Source model parameters for FTP traffic class

Component	Distribution	Parameters
File size (S_f)	Truncated lognormal	Min = 50 B, Max = 5 MB, Mean = 2 MB, STD = 722 kB
Inter-arrival time between each file (T_f)	Exponential	Mean = 180 sec
Packet size	Deterministic	3000 B
Call holding time	Normal	Mean = 180 Variance = 40 (sec)
Data during active period		88.9 kbps
Active rate		1
Mean data rate		88.9 kbps

The MTs with different traffic classes have different QoS requirements. The QoS requirements of each traffic class are listed in Table 4.6. [20]

Table 4.6: The QoS Requirements of each traffic class

Traffic class	Requirement	Value
Voice	Required BER	10^{-3}
	Required E_b / N_0	4 dB
	Max. delay tolerance	40 ms
	Max. allowable packet dropping rate	1%
Video	Required BER	10^{-4}
	Required E_b / N_0	3 dB
	Max. delay tolerance	100 ms
	Max. allowable packet dropping rate	1%
HTTP	Required BER	10^{-6}
	Required E_b / N_0	2 dB
	Min. transmission rate	5 kbps
FTP	Required BER	10^{-6}
	Required E_b / N_0	1.5 dB
	Min. transmission rate	20 kbps

4.3 Parameters Setting of SA

In the thesis, the SA parameters are determined by criterions suggested by [19], which states as follows:

(1) The value of $T_{\max}$

The value of $T_{\max}$ is determined to satisfy the criterion that the probability of accepting a worse new neighbor solution should be larger than 80% at the initial stage.

It is set to 5 according to (4.1)

$$\exp(-1/T_{\max}) \geq 80\%. \tag{4.1}$$

(2) The value of $T_{\min}$

The value of $T_{\min}$ is determined by the criterion that the probability of accepting a worse new neighbor solution should be less than 0.01% at the final stage, and it is set to 0.1 according to (4.2).

$$\exp(-\Delta f_{\max} / T_{\min}) \geq 0.01\%. \tag{4.2}$$

(3) The number of iterations at each temperature, L

It is suggested that the number of iterations should be set to 100 times of the number of variables that the SA algorithm try to optimize. In the thesis, it is set to 100.

(4) The decrement rate of temperature, γ

In order to escape from being trapped at the local optimum, it is suggested that γ should be set between 0.8 and 0.9. In the thesis, it is set to 0.8.

4.4 AB-GHOD Scheme and RSP-HO Scheme

In the simulation, the proposed SA-GHOD scheme is compared with the available bandwidth preferred group handover decision (AB-GHOD) scheme [4] and the received signal power based handover (RSP-HO) scheme [6]. In AB-GHOD scheme, each MT will have a probability min(M^* / M_{remain}, 1) to execute the group handover decision when the group handover is happening in a certain frame time, where M^* is the maximum number of MTs that can let the GHO blocking ratio less than 0.05, and M_{remain} is the remaining number of MTs that haven't execute the group handover in this frame. To get M^*, it is assumed that each MT knows the total number of MTs and the number of each traffic types in the handover group.

After obtaining the executing probability, each MT will decide whether to execute the group handover decision or not. If the MT decide to do the group handover in this frame, it uses another probability $C_k(t)/\sum_{k=1}^{K} C_k(t)$ to decide its target network, where $C_k(t)$ is the amount of the available bandwidth that BS k has at time t. If the MT decide to not to do the group handover in this frame, it will recalculate the executing probability after 100 ms. During this 100 ms, the author assume that the BS that will cause GHO blocking ratio higher than 0.05 can borrow a certain amount of bandwidth from another BS. In our simulation environment, there is no borrow mechanism between each BSs so that the handover latency is useless. Therefore, the executing probability is set to be 1 in our simulation. Besides, the original scheme does not consider different traffic types of MT. In order to compare fairly, the target network decision probabilities of different traffic types are calculated by their own perspective.

The RSP-HO scheme is a conventional single MT handover decision scheme. Each MT will detect and compare the received pilot signal power from BSs of each network. After finding the BS with largest signal power in each network, the MT will handover to one of the network randomly.

4.5 Performance Evaluation

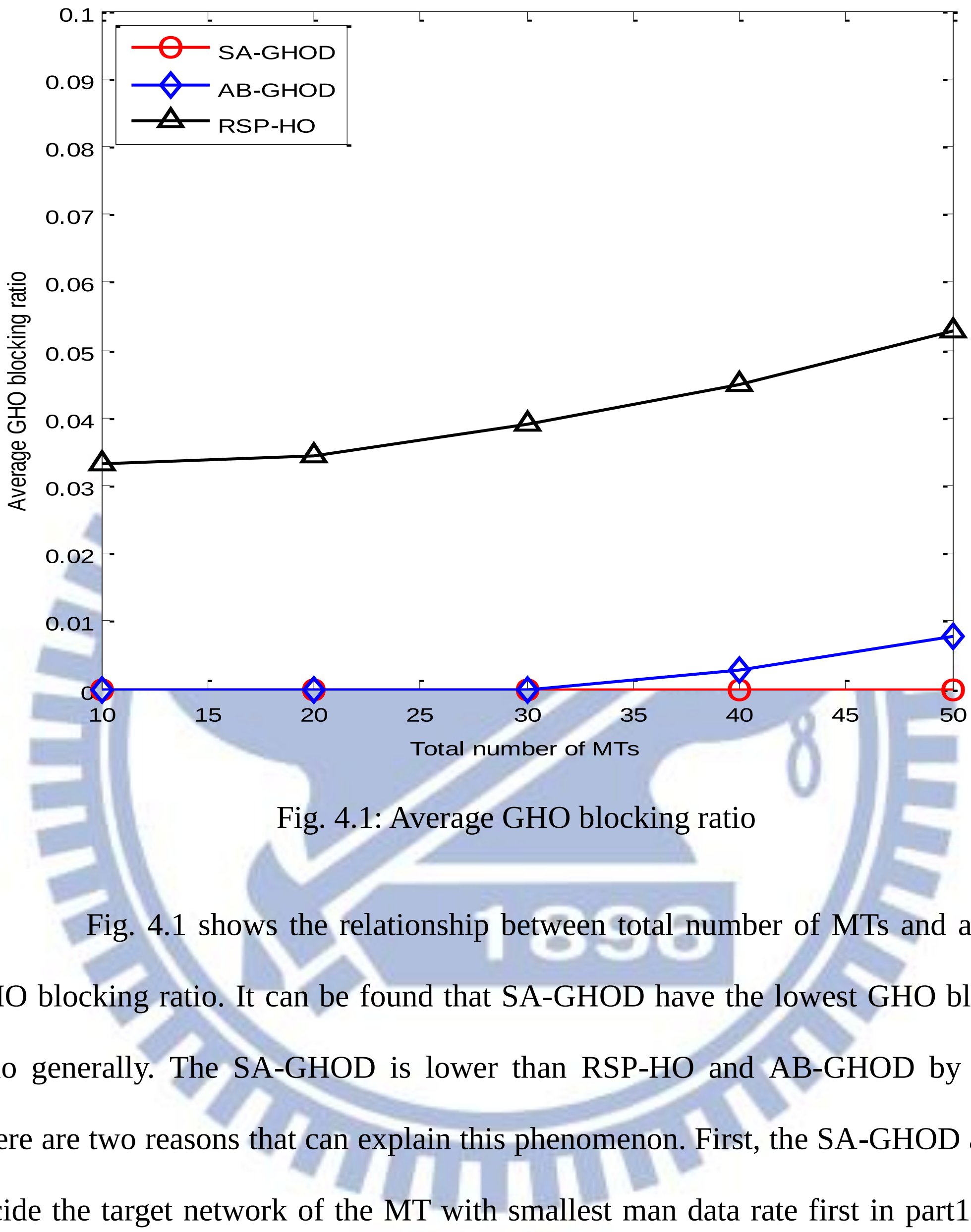


Fig. 4.1: Average GHO blocking ratio

Fig. 4.1 shows the relationship between total number of MTs and average GHO blocking ratio. It can be found that SA-GHOD have the lowest GHO blocking ratio generally. The SA-GHOD is lower than RSP-HO and AB-GHOD by 100%. There are two reasons that can explain this phenomenon. First, the SA-GHOD always decide the target network of the MT with smallest man data rate first in part1 of the scheme. Second, SA-GHOD will exchange the networks of two MTs and trying to let those MTs that are blocked in old solution to join one of the network also can reduce the average GHO blocking ratio. The reason why RSP-HO is much higher than the other two schemes is because of that the MTs does not consider the influence of other handover MTs. The AB-GHOD is lower than RSP-HO. It is because MTs will tend to select a network with larger available bandwidth.

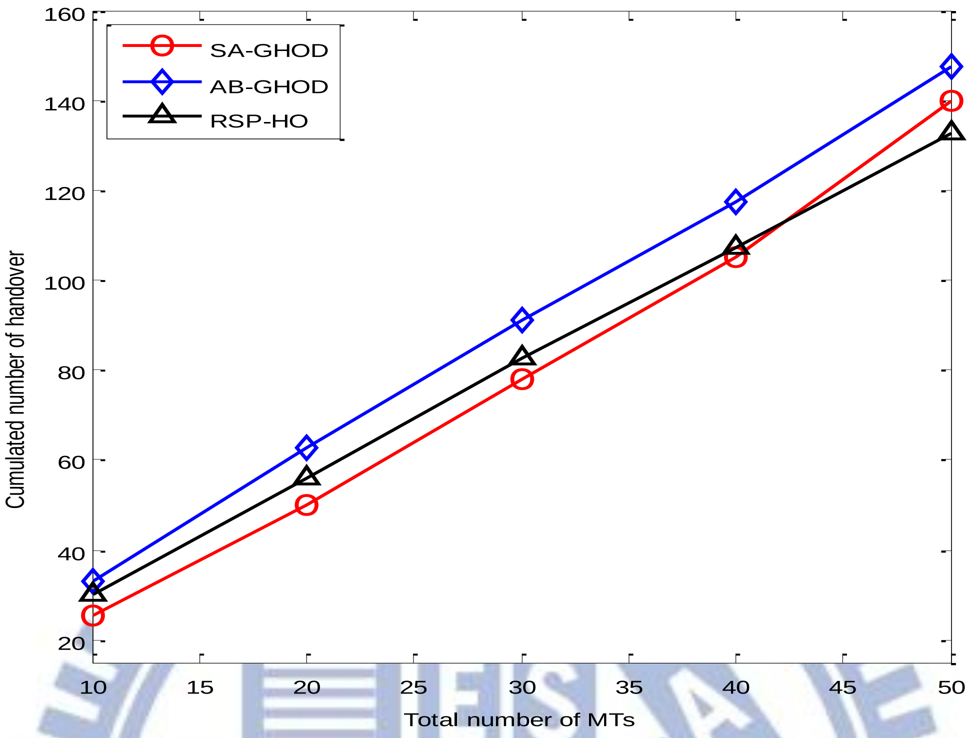


Fig. 4.2: Cumulated number of handover

Fig. 4.2 shows the cumulate number of handover during the simulation time. When the total number of MTs is lower than 50, SA-GHOD is lower than AB-GHOD and RSP-HO. The SA-GHOD is lower than RSP-HO and AB-GHOD by 15.7% and 9.7% when the total number of MTs is 40. When the total number of MTs rise to 50, the cumulated number of handover of SA-GHOD become higher. Due to that the GHO blocking ratio is designed to be the most important factor, it will keep the GHO blocking low by raising the total number of handover. However, it still better than the DP-GHOD. When there are more MTs in a vehicle, SA-GHOD needs more handover times.

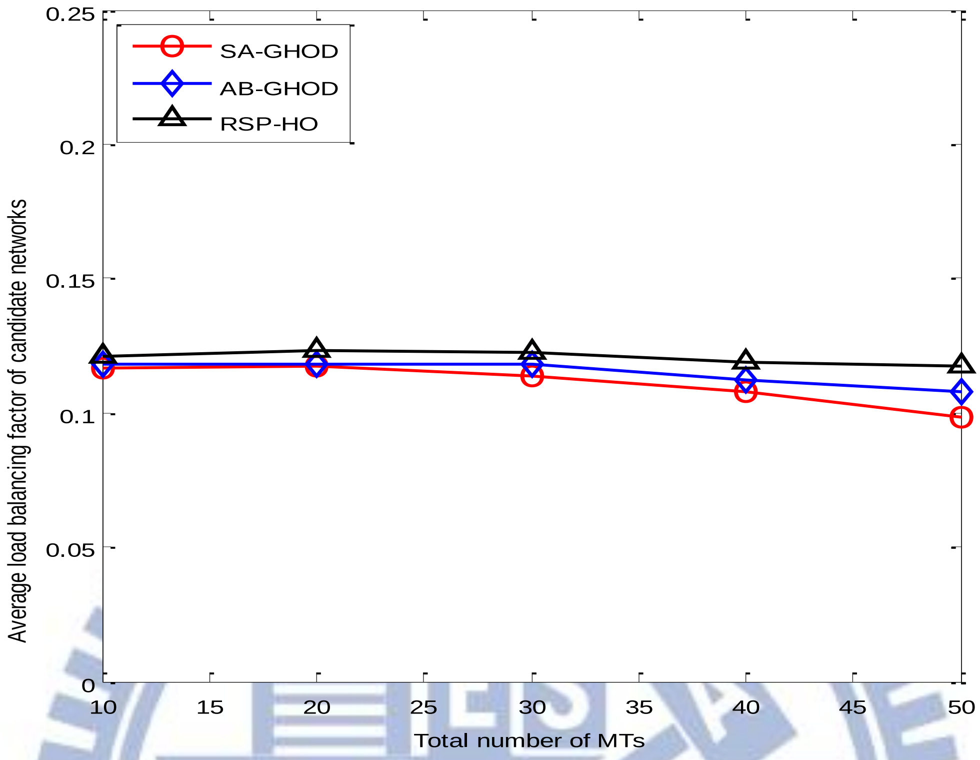


Fig. 4.3: Average load balancing factor of candidate networks

Fig. 4.3 shows the average load balancing factor of candidate networks in every GHO. The SA-GHOD is close to AB-GHOD when the total number of MTs is lower. The main reason is caused by the design of cost function. In the cost function, the dwelling time factor is consider the standard deviation, which means we want that every MT can have a larger dwelling time in the target network. Therefore, the average load balancing factor would be higher in order to have a lowest cumulated number of handover. When the total number of MTs become large, SA-GHOD has the lowest average load balancing factor, which means that the traffic load of candidate networks are the most balance. It is due to that the GHO blocking factor interact with the load balancing factor . In other words, balancing the load between networks can also reduce the GHO blocking ratio.

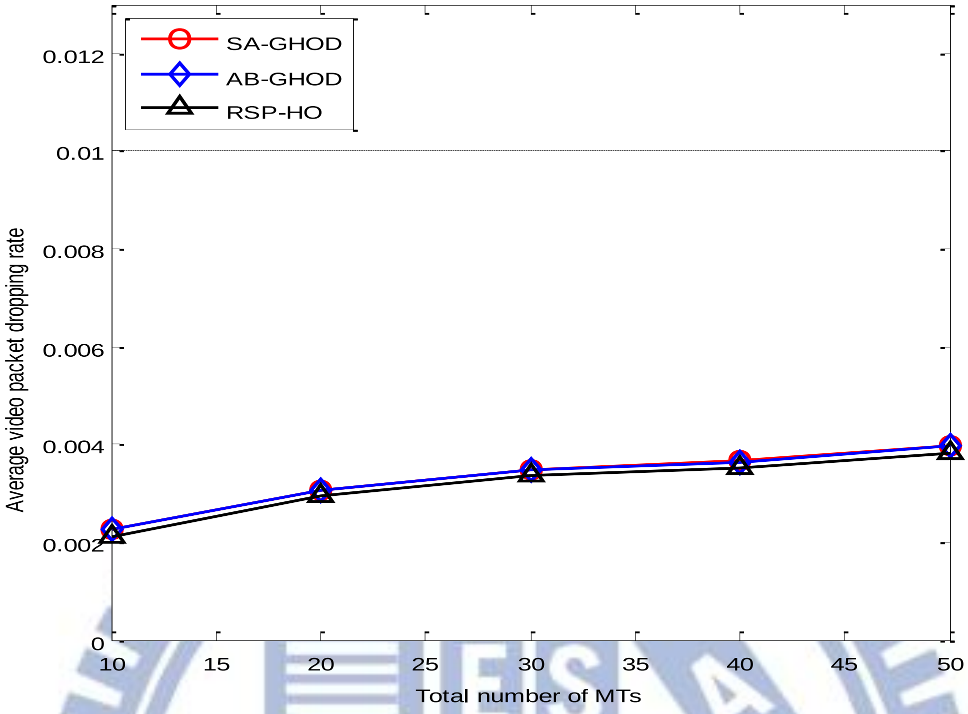


Fig. 4.4: Average video packet dropping rate

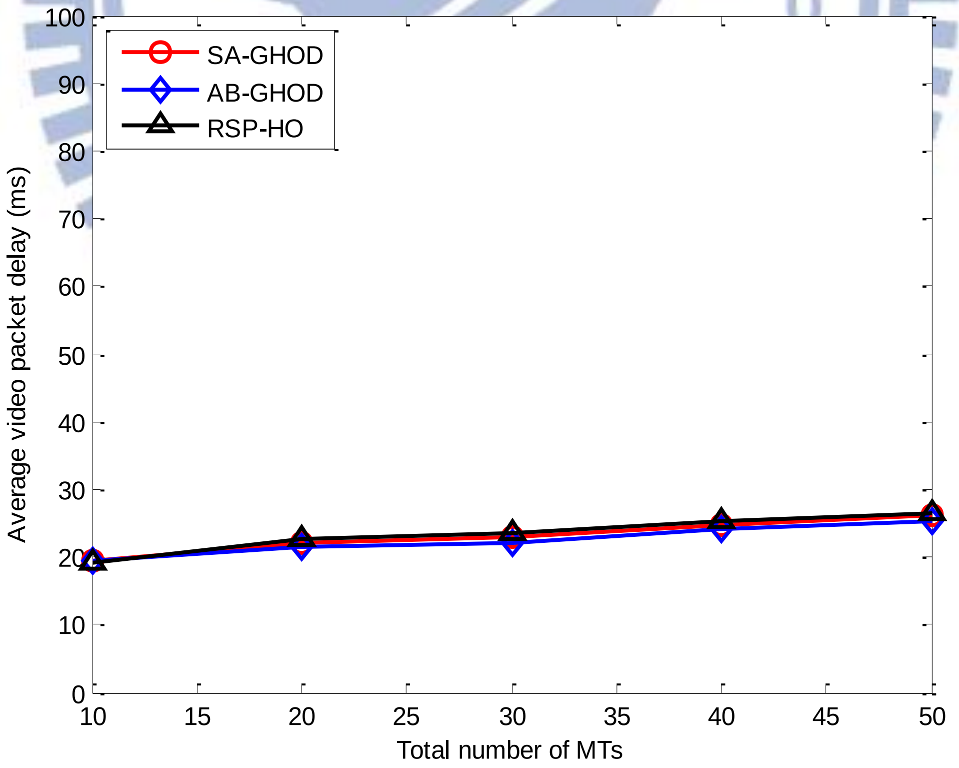


Fig. 4.5: Average video packet delay

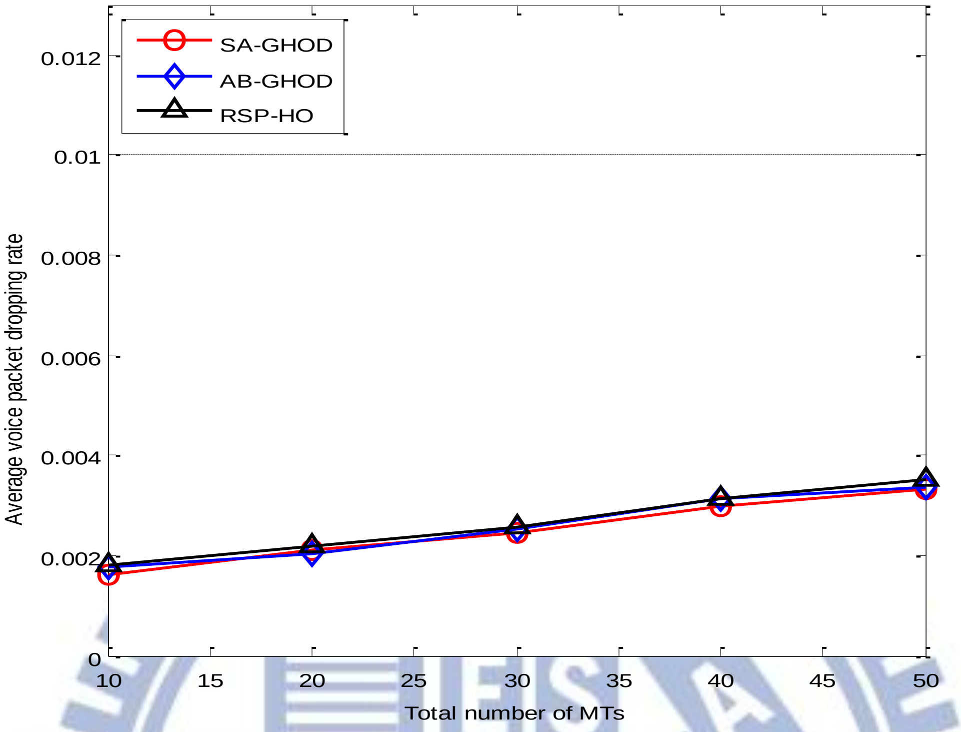


Fig. 4.6: Average voice packet dropping rate

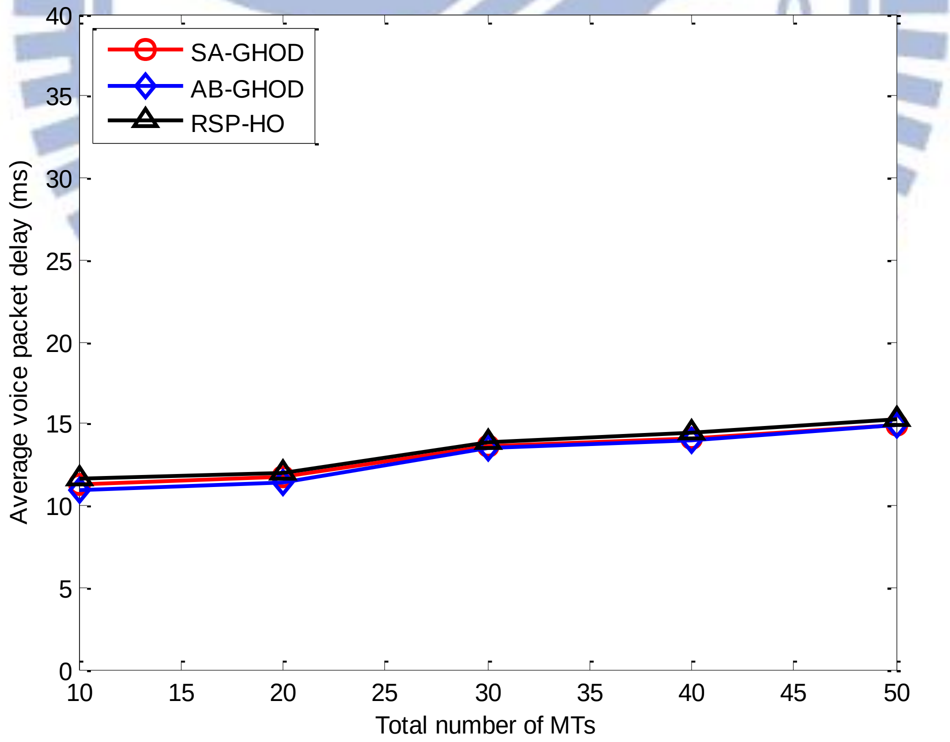


Fig. 4.7: Average voice packet delay

Fig. 4.4 to Fig. 4.7 shows the average video/voice packet dropping rate and the average video/voice packet delay of the MTs in the vehicle. Due to that we assume that each scheme cares about the QoS, which means the MTs in the three schemes only join the network in joinable network set, using any of the three scheme can satisfy the packet dropping rate requirement. When the total number of MTs is lower than or equal to 60, the average packet are low due to that the networks completely have enough resources to serve the MTs in the vehicle.

Chapter 5
Conclusions

In this thesis, a simulated annealing based group handover decision (SA-GHOD) scheme is proposed for heterogeneous wireless network environment, which considers voice, video, HTTP, and FTP services. A candidate network set will be found first by checking the pilot signal power. A cost function is designed, which consider group handover blocking ratio, dwelling time, and load balance to represent the degree of cost of the decision.

The proposed SA-GHOD scheme can be divided into three parts. The first is initialization, which finds the network with better QoS and the initial target network of each MT in the vehicle. The second is generation and acceptance, which try to get a new target network vector by switching target network of two MTs in the existing target network vector. The third is cooling schedule, which decreases the temperature in order to find a better solution after each round of iteration.

In the simulation results, the SA-GHOD scheme is compared to the DP-GHOD and RSP-HO scheme. The results shows that SA-GHOD has lowest average GHO blocking ratio. the cumulated number of handover has a trade-off relationship with the load balancing factor when compare to RSP-HO, the low balancing factor is better than the DP-GHOD, and all three schemes working under the satisfaction of QoS requirements. Finally, we can conclude that the performance of SA-GHOD is better than the DP-GHOD, especially when the total number of MT is not too large. Then the use of RSP-HO scheme in the group handover scenario is not suitable due to that the main problem in GHO is the group handover blocking ratio.

Bibliography

[1] F. A. Zdarsky and J. B. Schmitt, “Handover in Mobile Communication Networks: Who is in Control Anyway?,” *IEEE Computer Science EURMIC.*, pp.205-212, Sep. 2004.

[2] X. Cai and F. Liu, “Network Selection for Group Handover in Multi-Access Networks,” *IEEE ICC,* pp.2164-2168, May 2008.

[3] S. Lei, T. Hui, and H. Zheng, “Group Vertical Handover in Heterogeneous Radio Access Networks,” *IEEE VTC-Fall,* pp.1-5 2010.

[4] W. Lee and D.H. Cho, “Enhanced Group Handover Scheme in Multiaccess Networks,” *IEEE Trans. Veh. Technol.*, Vol.60, pp.2389-2395, Jun. 2011.

[5] L. Shan, F. Liu, L. Wang, and Y. Ji, “Predictive Group Handover Scheme with Channel Borrowing for Mobile Relay Systems,” *IEEE IWCMC*, pp.153-158, Aug. 2008.

[6] G. Zhang and F. Liu, “An Auction Approach to Group Handover with Mobility Prediction in Heterogeneous Vehicular Networks,” *IEEE ITST*, pp.584-589, Aug. 2011.

[7] M. Abid, T. A. Yahiya, and G. Pujolle, “A Utility-based Handover Decision Scheme for Heterogeneous Wireless Networks,” *IEEE CCNC*, pp.650-654, Jan. 2012.

[8] C. Wan, “Efficiency comparison of simulated annealing algorithm and Genetic Algorithm in Feature Selection,” *IEEE ICNC*, Vol.5, pp.2330-2333, Aug. 2010.

[9] J. Eberspächer, H.J. Vögel, C. Bettstetter, and C. Hartmann, “*GSM Architecture, Protocols, and Services*,” 3rd Ed., John Wiley and Sons, Ltd, 2009.

[10] H. Holma and A. Toskala, “*WCDMA for UMTS*,” 3rd Ed., John Wiley and Sons,

Ltd, 2002.

[11] F. Khan, “*LTE for 4G Mobile Broadband*,” Cambridge University Press, 2009.

[12] T.S. Rappaport, “*Wireless Communications Principles and Practice*,” 2nd. Ed., Prentice Hall, 2002.

[13] G.L. Stüber, “*Principles of Mobile Communication*,” 2nd Ed., Springer, 1996.

[14] A. Rojas, “Shadow Margin, Overlapping and Transceivers per Cell when making Handovers with Hysteresis Margin,” *IEEE PIMRC*, Vol. 1, pp.340-344, Sep. 1998.

[15] E. Hyytia, P. Lassila, and J. Virtamo, “Spatial Node Distribution of the Random Waypoint Mobility Model with Applications,” *IEEE Trans. Mobile Computing*, Vol. 5, No. 6, pp.680-694, Jun. 2006.

[16] UMTS Std. 30.03, “Selection procedures for the choice of radio transmission technologies of the UMTS,”1998.

[17] 3GPP TR 25.892, “Feasibility study for orthogonal frequency division multiplexing (OFDM) for UTRAN enhancement,” 3rd Generation Partnership Project, Tech. Rep., Jun. 2004.

[18] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluh, and A. H. Teller, “Equation of State Calculations by Fast Computing Machines,” *The Journal of Chemical Physics*, Vol.21, No. 6, pp.1087-1092, Jun.1953.

[19] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, “Optimization by Simulated Annealing,” *Science*, Vol. 220, No. 4598, pp.671-680, May 1983.

[20] 3GPP TS 23.107, “QoS concept and architecture,” 3rd Generation Partnership Project, Tech. Spec., Dec. 2008.

Vita

Heng-Hao Pu was born in 1985 in Taoyuan County, Taiwan. He received the B.E. degree in engineering science from National Cheng Kung University, Tainan, Taiwan, in 2010, and the M.E. degree in the institute of communication engineering, college of electrical and computer engineering from National Chiao Tung University, Hsinchu, Taiwan, in 2012. His research interests include wireless communication and radio resource management.