The University Of Sheffield.

Electronic & Electrical Engineering

Data Traffic Analysis and Small Cell Deployment in Cellular Networks

A THESIS SUBMITTED TO

THE UNIVERSITY OF SHEFFIELD

IN THE SUBJECT OF TELECOMMUNICATIONS

FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY.

By

Nan E

2016

Abstract

In this thesis, the study of small cell deployment in heterogeneous networks is presented. The research work can be divided into three aspects. The first part is user data traffic analysis for an existing 3G network in London. The second part is the deployment of additional small cells on top of existing heterogeneous networks. The third part is small cell deployment based on stochastic geometry analysis of heterogeneous networks.

In the first part, an analysis of 3G network user downlink data traffic is presented. With the increasing demands for high data rate and energy-efficient cellular service, it is important to understand how cellular user data traffic changes over time and in space. A statistical model of time-varying throughput per cell and the distribution of instantaneous throughput per cell over different cells based on throughput measurements from a real-world large-scale urban cellular network are provided. The model can generate network traffic data that are very close to the measured traffic and can be used in simulations of large-scale urban-area mobile networks.

In the second part of the work, three different small-cell deployment strategies are proposed. As the mobile data demand keeps growing, an existing heterogeneous network composed of macrocells and small cells may still face the problem of not being able to provide sufficient capacity for unexpected but reoccurring hot spots. The proposed strategies avoid replanning the overall

network while fulfilling the hot spot demand by optimizing the deployment of additional mobile small cells on top of the existing HetNet. By simplified the optimisation problem, we first proposed a fixed number deployment algorithm and then extend it into deployment over existing network algorithm to solve the joint optimisation problem. The simulation results show that these two proposed algorithms require less small cells to be deployed while providing higher minimum user throughput. Moreover, a reduced-complexity iterative algorithm is proposed. The simulation results show that it significantly outperforms the random deployment of new small cells and achieves performance very close to numerically solving the joint optimisation in terms of minimum user throughput and required number of new small cells, especially for a large number of unexpected hot-spot users.

In the third part, a stochastic geometry analysis is provided for a heterogeneous network affected by a large hot spot. Based on the analysis, the optimal numbers of additional small cells required in the HS and non-HS areas are obtained by minimizing the difference between the numbers of macrocell users after and before the HS occurs. Then an algorithm is proposed to maximize the average user throughput by jointly optimizing the locations of additional small cells and user associations of all cells. Simulation results show that the proposed algorithm can maintain the average user throughput above a threshold with excellent fairness among all users even for a very high density of HS users.

Acknowledgements

In the researching process of this work, I have received unconditional support and love from my parents Juan Cheng and Jingping E. I am also indebted to my wife Wenjing Liu, who always supports me and helps me to overcome all the difficulties I have met.

I would also like to acknowledge the guide and help provided by Dr. Xiaoli Chu. With her patient guidance and encouragements, I have made more achievements than expected. I am also thankful to Prof. Jie Zhang, who always support my research. During the course of my research, Dr. Weisi Guo and Dr. Wei Liu have also given me valuable suggestions. I am particularly grateful Prof. Richard Langley for the excellent research environment in my department. My special thanks are extended to Dr. Simon Bai, Dr. Hector Shi, Dr. Yue Wu, Dr. Yang Liu, Dr. Siyi Wang, Dr. Xirui Zhang and Dr. Bo peng. I would also like to thank my colleague Long Li, Wuling Liu, Dehua Li, Zhan Qin, Ken Deng, Tian Feng, Mengdi Jiang and Wenfei Yin for their friendship and assistance during my doctorate study.

Finally, I would like to offer my special thanks to Hilary J Levesley for her assistance and guidance. I would also want to send my gratitude to my faculty for all the supports throughout my doctorate process.

Nan E

2016

List of Figures

List of Tables

Notations

C- Capacity

B- Bandwidth

P_t- Transmit power

AP- Antenna pattern

PL- Path loss

k- Boltzman's Constant

T- Absolute temperature

f_c- Frequency in GHz

h_1- Base station height

h_2- user equipment height

$h_{building}$- Average height of buildings

W_{street}- Width of street

d -Distance

A_H- Horizontal antenna pattern

A_V- Vertical antenna pattern

θ- Azimuth angle

bw- Bin width

T_{max} - Maximum values of the instantaneous throughput per cell

T_{min} - Minimum values of the instantaneous throughput per cell

λ_{weekday} - Rate parameter for weekdays

λ_{weekend} - Rate parameter for weekend days

N_F - Number of small cells

N_{eBS} - Number of existing base stations

N_{nBS} - Number of new base stations

N_{BS} - Total number of base stations

N_{nhs} - Number of non-hot spot users

N_{hs} -Number of hot spot users

N_U - Total number of users

N_U - Set of users

N_{BS} -Set of base stations

N_{RB} - Set of resource blocks

γ_i - Throughput of i th user

$a_{i,j}^k$ - Joint user allocation and resource allocation indicator

P_j^k - Downlink transmit power of the j th base station

$g_{i,j}^k$ - Channel power gain

$g_{f,ij}^k$ - Exponentially distributed fading gain

$g_{pl,ij}$ - Path loss gain

(x_i, y_i) - Location coordinates of the i th user

(x_j, y_j) - Location coordinates of the j th base station

α_{pl} - Path loss distance exponent

N_0 - Additive white Gaussian noise power

$I_{i,j}^k$ - Interference power

$\mathbf{A}$ - Joint user association and resource allocation joint matrix

$(\mathbf{x}, \mathbf{y})$ - Coordinate vectors

γ_{th} - user throughput threshold

$|H|$ - Feasible deployment area

$\mathbf{B}$ - User association matrix

$b_{i,j}$ - User association indicator

$\mathbf{C}$ - Resource allocation matrix

$N_{i,j}^{RB}$ - Number of resource blocks allocated by the j th base station to the i th user

N_{j}^{RB} - Number of resource blocks per user in j th base station

N_{nBS}^{max} - Maximum number of new small cells

W - Bandwidth of a resource block

α - Non-hot spot user intensity

β - Hot spot user intensity

ν - Existing small cell intensity

μ - Intensity of new small cells in non-hot spot area

ρ - Intensity of new small cells in hot spot area

σ_1^2 - User to small cell distance correlation factor

R_M - Macrocell coverage area

R_H - Hot spot area

$R_{H'}$ - Hot spot area excluding the coverage area of existing small cells in the hot spot

R_{nH} - Non-hot spot area

χ_{AC} - Expected number of non-hot spot users in the coverage of an existing small cell base station

χ_{C} - Expected number of existing small cells in the macrocell coverage area

$E\{N_U\}$ - Expected total number of non-hot spot users in the macrocell coverage area before hot spot occurs

χ_{ABE} - Expected number of users in a new small cell coverage in the hot spot area

χ_{AB} - Expected number of users in an existing small cell coverage that overlaps with the hot spot area

χ_{AD} - Expected number of non-hot spot users served by new small cells in non- hot spot areas

χ'_{C} - Expected number of existing small cells in non-hot spot areas

χ_{E} - Expected number of additional small cells in the hot spot area

χ_{D} - Expected number of additional small cells in non-hot spot area

$E\{N'_U\}$ - Expected total number of UEs in the macrocell coverage area with hot spot

F - Fairness index

HetNet- Heterogeneous Network

LTE- Long Term Evolution

LTE-A- Long Term Evolution advanced

3GPP- 3rd Generation Partnership Project

MIMO- Multiple Input Multiple Output

AP- Access Points

QoS- Quality of Service

UE- User Equipment

HS- Hot Spot

OFDMA- Orthogonal Frequency Division Multiple Access

OFDM- Orthogonal Frequency Division Multiplexing

ISI- Inter Symbol Interference

RB- Resource Block

SINR- Signal-to-Interference-Plus-Noise Ratio

NLOS- Non-Line-of-Sight

LOS- Line-of-Sight

PDF- Probability Density Function

RMSE- Root Mean Squared Error

GMM- Gaussian Mixture Model

SPPP- Spatial Poisson Points Process

CDMA- Code-Division Multiple Access

HSPA- High Speed Packet Access

PPP- Poisson Point Process

CDF- Cumulative Distribution Function

FNDA- Fixed Number Deployment Algorithm

DOENA- Deployment over Existing Network Algorithm

DL- Downlink

B&B- Branch and Bound

GRG- Generalized Reduced Gradient

SD- Standard Deviation

Content

Chapter 1

Introduction

1.1 Background

Mobile phones are much more widely used all over the world since it has been developed in 1973 [1]. Nowadays, mobile phones are not only used for voice services but also data services in last decades. The mobile communication networks as one of the greatest changes provide a brand-new life style for people. One of the most significant revolutions of mobile communications is that users can use their devices to access the Internet as long as they are in the operator's coverage area. However, this revolution results in an exponential increase of data demands. Compared with 2014, the wireless data traffic has grown by 74 percent worldwide in 2015 [2]. This huge amount of data requirement gives operators a challenge of satisfying subscribers while

reducing power consumption and interference. From the third generation (3G) to fourth generation (4G) and fifth generation (5G), the users' data traffic demands continue increasing. In order to reduce the pressure of macro base stations, small cells are introduced to provide higher quality signal to users.

One of the core concepts to solve the conflictions between huge data demands and limited radio resource is to increase the number of base stations in order to get a higher total capacity. Those base stations which have coverage of kilo-meter order of magnitude have been changed to hundreds of meters' coverage which is known as macrocells. Another kind of base stations have also been developed as an additional tier of the networks which provide only dozens of meters' coverage which is known as small cells. Such multi-tier networks are known as heterogeneous networks (HetNets).

HetNets as an important concept in LTE-A is widely used by operators and network designers nowadays. It is anticipated that HetNets would mitigate the conflict between the rapid growth of data demand and limited radio resources by increasing the area spectral efficiency through densely deploying low-power small cells such as femtocells [3]. As a trend that mobile data demands increasing in an exponential rate and will become more and more important to people's daily lives, the management of limited network resource is vital to HetNets. For example, a HetNet can be constructed by overlaying low-power small base stations (BSs) on top of the existing macrocell network to increase the network capacity [4].

Since Long-Term Evolution (LTE) was launched in March of 2009 [5], the 3rd Generation Partnership Project (3GPP) has been devoted to improving the performance of LTE via advanced Multiple Input Multiple Output (MIMO), carrier aggregation and HetNets, as evident from the LTE Advanced (LTE-A). HetNets use mixes of microcells, femtocells, picocells, relay BSs and other kinds of small access points cooperate with macrocells to establish flexible and low-cost networks to provide high QoS to users [3]. Fig.1.1 shows an example of architecture of a HetNet.

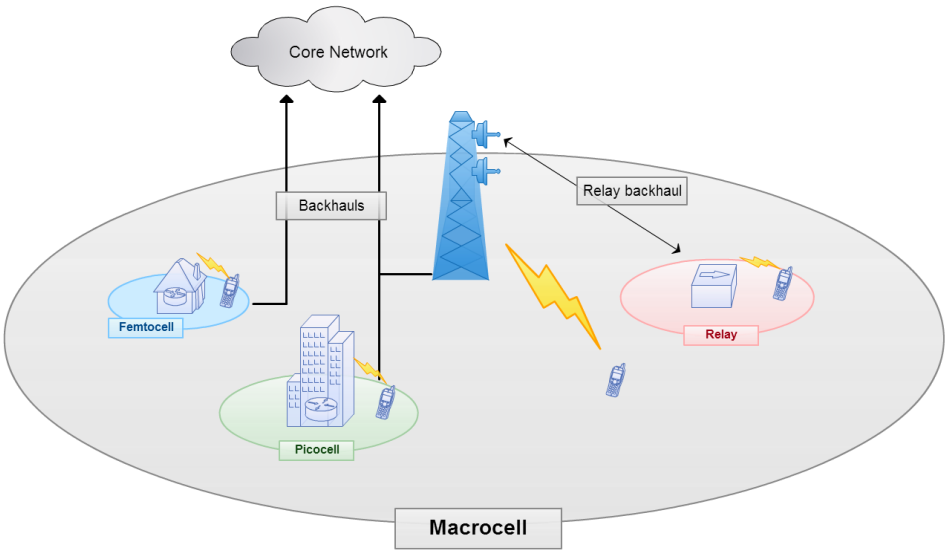


Fig. 1.1 Architecture of a heterogeneous mobile network.

Small cells are low-power access points (APs) for users to access mobile networks and get services. According to the scale of coverage area and different subscribers, small cells encompass femtocell, picocell, microcell and relays.

Among these small cells, femtocell is used in home or a small business. The coverage radius of a femtocell normally is from 10 meters to 20 meters. The subscribers of femtocell made it become randomly in deployment. There are two kinds of femtocells: open access and closed access small cells. The former can serve any user belongs to its operator located in its coverage area, while the later only allows the user that has authorized to access. According to recent researches [6][7], open access femtocells have a better performance in both uplink (UL) and downlink (DL) capacity. As a result, open access femtocells will be used widely by operators in the real network environment. As a result, these kind of femtocells have capability to support users that nearby their coverage and reduce stress of macrocells loads they are located in.

Picocells are also a kind of small coverage APs, however, they have a larger coverage radius than femtocells. Normally, a picocell can provide service to a user up to 200 meters away. Picocells are often deployed in an office building, shopping mall or a stadium. The deployment of a picocell are usually participated by operators in order to control the influence of interference into an acceptable level during deploying.

Relay base stations can increase the density and coverage area of the LTE networks. Due to no separate backhaul required by a relay base station, the deployment is easier than a small cell (femto or pico). The backhaul directly

connects to the macrocell base station. Therefore, relays have slight improvement on increasing spectrum efficiency in the same area but have significant effect on increasing the QoS and extending coverage.

1.2 Motivation

Usually, operators plan HetNets are based on the expected user distribution and mobile traffic pattern which is obtained from long-term observations and big data collections. It is important to achieve both good service quality and low cost in HetNet planning and deployment. In [8], the deployment of small cells is optimised to obtain the best trade-off between user Quality of Service (QoS) and operators' costs. The dynamic small cell deployment strategy in [9] can be used to find out when and where small cells need to be deployed. However, in an existing HetNet, persistent clusters of user equipments (UEs), a.k.a., hot spots (HSs), which were not been expected in the original network planning may occur, causing extra traffic demand. Therefore, the main target of my research work is focus on the solution of additional data requirements caused by HSs.

In order to study and solve the HS problem in HetNet, it is important to understand the existing mobile network. Therefore, a data traffic analysing for existing 3G network has been undertaken first. After, three different small cell deployment strategies are developed to solve the extra data demand problem caused by HSs in 4G HetNet. These three strategies are developed in the

scenario that close to reality in 4G networks, therefore, a small cell deployment strategy that based on a theoretical analysis has been studied by using stochastic geometry analysis.

1.3 Thesis Research and Contributions

Aa a result, my research includes three aspects: 1) data traffic analysis; 2) small cell deployment in HetNets; 3) stochastic geometry analysis and small cell deployment of HetNet. Statistical modelling of time-varying throughput per cell has been implemented in the study of 1) in order to understand the existing traffic demands for 3G network. The result is efficient and could be used in related researches. In the research of 2), the deployment strategy for 4G has been studied first with a joint optimisation of the number and locations of additional mobile small cells and the user associations of all cells carried out in order to maximizing the minimum UE throughput. Furthermore, a reduced-complexity iterative algorithm is devised to solve the joint small cell deployment optimisation problem has also been carried out. In the research of 3), small cell deployment is based on stochastic geometry analysis, a HetNet affected by a large HS and for the additional small cells that need to be deployed based on the spatial bivariate Poisson point process has been provided first and the optimal numbers of additional small cells are obtained by minimizing the difference between the numbers of macrocell users after and before the HS occurs based on the stochastic geometry analytical results. An

algorithm to maximize the average user throughput by jointly optimizing the locations of additional small cells and user associations of all cells has been developed.

The research work is presented in this thesis by the following layout:

Chapter 2 Preliminaries: This chapter introduces the background knowledge and related works of data traffic and HetNet. Concepts of HetNets architecture, femtocells, small cell deployment are introduced in this chapter. Moreover, literature reviews on the traffic load analysis, including mathematical methods to analyse a measurement data set obtained from an existing network, and small cell deployment has been presented as well.

Chapter 3 User Data Traffic Analysis: The analysis of an existing 3G network and proposes a statistical method to predict the instantaneous throughput per cell has been presented. This is followed by a complete mathematical analysing approach to the measured data and evaluation of the result.

Chapter 4 Small Cell Deployment for 4G heterogeneous Networks: The small cell deployment for 4G HetNet has been presented in this chapter. Firstly, a strategy to optimise the number and locations of additional mobile small cells on top of an existing HetNet to fulfil the excess traffic demand of recurring HSs that were not expected in the original HetNet planning has been proposed. Secondly, a strategy based on reduced-complexity iterative algorithm to solve the joint deployment optimisation problem has been presented. The simulation result for both strategies have been presented as well.

Chapter 5 Small Cell Deployment Based on Stochastic Geometry Analysis: An algorithm to maximize the average user throughput by jointly optimizing the locations of additional small cells and user associations of all cells has been presented that based on the stochastic geometry analysis of the HetNet with a large HS. The simulation results have been presented.

Chapter 6 Conclusion and Future Work: The summary of this research has been presented with recommended future work.

1.4 Publications

- N. E, X. Chu and J. Zhang, "Small-cell deployment over existing heterogeneous networks," IET Electronics Letters, vol. 52, iss. 3, pp. 241-243, Feb 2016.

- N. E and X. Chu, "Stochastic geometry analysis and additional small cell deployment for HetNets affected by hot spots," Mobile Information Systems, Volume 2016 (2016), Article ID 9727891, pp. 1-9, Jan 2016.

- N. E, X. Chu and J. Zhang, "Mobile small-cell deployment strategy for hot spot in existing heterogeneous networks," IEEE GLOBECOM'15 WS, San Diego, CA, USA, 6-10 Dec 2015.

- N. E, X. Chu, W. Guo and J. Zhang, "User data traffic analysis for 3G cellular networks," International Conference on Communications and Networking in China (ChinaCom'13), pp. 468-472, Guilin, China, 14-16 Aug 2013.

Chapter 2

Preliminaries

2.1 Introduction

This chapter reviews concepts of 3G and 4G networks in order to support the design in the following chapters. Fundamental concepts and elements of HetNets and related techniques are listed in this chapter. Moreover, mathematical concepts and random processes are also presented, including exponential random process, density evaluation, curve-fitting, goodness of curve-fitting evaluation and mixture models. The literature review of related works is presented as well.

2.2 Fundamental Concepts

2.2.1 Orthogonal Frequency Division Multiple Access

Orthogonal Frequency Division Multiple Access (OFDMA) has been considered as the multiple access technology for mobile networks in LTE. The concept of Orthogonal Frequency Division Multiplexing (OFDM) is to divide the bandwidth into multiple orthogonal frequency sub-carriers. Therefore, the Intersymbol Interference (ISI) is completely eliminated [10]. By using time-frequency plane, the resource of OFDMA is divided into resource blocks (RBs). Each RB has 180kHz bandwidth for subcarriers. The number of subcarriers in each RB depends on the spacing of subcarriers, for example, if the subcarrier spacing is 15kHz, there are 12 subcarriers in each RB. In small cell deployment researching, RBs are considered as an essential element for resource allocating.

2.2.2 Small cells

Small cells were first introduced in 1984 [11]. The idea is to use multiple smaller cells instead of a macrocell. As mentioned in Section 2.2, a small cell needs to be self-adaptive to networks environment. For example, femtocells as a kind of small cell that has IP-based backhaul interface directly connect to core networks. The reason for small cells to be promoted is the demands of mobile

data use is increasing. In 2015, the wireless data traffic grew by 74 percent worldwide as compared with 2014, which indicates there is a large potential group of users who will require more mobile data in the future [2]. One solution of meeting this demand is increasing spatial frequency reuse by reducing the coverage of cells. Therefore, small cells are more and more popular in recent year's research [12].

However, the advocating of small cells brings an additional challenge. A macrocell user near a closed-access small cell AP will experience high interference [13]. There are some recent researches to solve this issue. In [14], sectorized antennas are added to femtocell APs to limit the number of interference source so as to reduce the interference in uplink. Another approach using dynamic selection of antenna patterns can also reduce the interference by reducing the power leakage outdoor [15]. Moreover, in [16], a channel switching scheme based on interference level of each available channel is provided to reduce co-channel interference to uplink.

2.2.3 Standards in LTE-A

The simulation parameters and formulas to calculate important elements in LTE-A are listed in this section. 3GPP updates new contents and revises previous

concepts frequently. It is very important to review these concepts and use appropriate methods in simulation.

The main target of optimisation is to increase the capacity of users. Shannon formula is the basic method to calculate the capacity of a radio channel:

$$\mathrm{C} = \mathrm{B} \times \log_2(1 + \mathrm{SINR}) \tag{2.1}$$

where C is capacity, B is bandwidth. Signal-to-interference-plus-noise ratio (SINR) is used to measure the quality of wireless connections in telecommunications. It can be presented as:

$$\mathrm{SINR} = \frac{\text{Received Power}}{\text{Noise} + \text{Interference Power}} \tag{2.2}$$

where received power in this formula can be calculated by:

$$\text{Received Power} = \frac{P_t \cdot \mathrm{AP}}{\mathrm{PL}} \tag{2.3}$$

where P_t is transmit power, AP is antenna pattern and PL is path loss. Interference is the sum of received power from all co-channel interferers. If we assume there are n interferers, then we have:

$$\text{Interference} = \sum_{i=1}^{n} \frac{P_t \cdot \mathrm{AP}_i}{\mathrm{PL}_i} \tag{2.4}$$

Noise is considered as a constant value which is presented as follows:

$$\text{Noise} = k \times \text{T} \times \text{B} \tag{2.5}$$

where k is Boltzman's Constant, $k = 1.38 \times 10^{-23} \text{Joules / Kelvin}$, T is absolute temperature in Kelvin ($0^{\circ}\text{C} = 273\text{Kelvin}$).

2.2.4 Path loss

A latest path loss model can be found in [17]. This model includes non-line-of-sight (NLOS) formulas and line-of-sight (LOS) formulas for urban microcells and macrocells.

These models are suitable for bandwidth from 2 GHz to 6 GHz in different antenna heights. Moreover, the real world street map is also considered in this model. Therefore, simulation results will be closer to real propagation.

The microcell path loss for traditional hexagonal cell layout is presented as:

$$\text{PL}_{LOS} = 22\log_{10}(d) + 28 + 20\log_{10}(f_c) \tag{2.6}$$

$$\text{PL}_{NLOS} = 36.7\log_{10}(d) + 22.7 + 26\log_{10}(f_c) \tag{2.7}$$

where d is the distance in meters from user to base station and f_c is bandwidth in GHz. If taking user equipment height into consideration, this model could be presented as:

$$\mathrm{PL}_{LOS} = 40\log_{10}(d_1) + 7.8 + 18\log_{10}(h_1)$$

$$-18\log_{10}(h_2) + 2\log_{10}(f_c) \qquad (2.8)$$

Where $d_1 < 5000\mathrm{m}$, h_1 is base station height, $h_1 = 15\mathrm{m}$, h_2 is user equipment height, $h_2 = 1.5\mathrm{m}$.

And the macrocell path loss LOS model is represented is same as (2.6).

For NLOS model, the macrocell path loss is more complicated:

$$\mathrm{PL}_{NLOS} = 161.04 - 7.1\log_{10}(\mathrm{W})$$

$$+7.5\log_{10}(h) - \left(24.37 - 3.7\left(\frac{h}{h_{BS}}\right)^2\right)\log_{10}(h_1)$$

$$+(43.42 - 3.1\log_{10}(h_1))(\log_{10}(d) - 3) + 20\log_{10}(f_c)$$

$$-(3.2(\log_{10}(11.75h_{UT}))^2 - 4.97) \qquad (2.9)$$

where h is the average height of buildings, $h = 20\text{m}$, W is the width of street, $\text{W} = 20\text{m}$, h_1 in this case is 25 meters. The LOS model for macrocell in urban environment could be changed if consider height is same with microcells.

In [18], the path loss that for general use is given by:

$$PL = -15.3 - \alpha_{pl} \cdot 10log_{10}(d)\text{dB} \tag{2.10}$$

where α_{pl} is the path loss exponent and d is the distance. Since the generality of this path loss model, it has been implemented in this research work.

2.2.5 Antenna Gain

Antenna gain is an important element in radio propagation. Different antenna main-lobe pointing direction will have various signal strength. In [18], antenna gain has been defined. An example of using it can be found in [19].

In HetNet, the horizontal antenna pattern can be presented as follows:

$$\text{A}_H(\theta) = \text{A}_M^{max} - \min\left\{12\left(\frac{\theta}{\varphi_{3\text{dB}}}\right)^2, \text{A}_m\right\} \tag{2.11}$$

where θ is the azimuth angle relative to the main-lobe pointing direction, $\mathrm{A}_M^{max} = 15\mathrm{dB}$, $\varphi_{3\mathrm{dB}} = 70^\circ$, and $\mathrm{A}_m = 25\mathrm{dB}$.

The vertical antenna pattern can be presented as follows:

$$\mathrm{A}_V(\theta) = \mathrm{A}_M^{max} - \min\left\{12\left(\frac{\theta - \theta_{etilt}}{\theta_{3\mathrm{dB}}}\right)^2, \mathrm{SLA}_v\right\} \qquad (2.12)$$

Where $\theta_{3\mathrm{dB}} = 10^\circ$, $\mathrm{SLA}_v = 20\mathrm{dB}$ and $\theta_{etilt} = 15^\circ$ or 6° based on different simulation environment. 15° is used in 3GPP case 1 and 6° is used in 3GPP case 3. Both 3GPP case 1 and 3 have the same carrier frequency, but case 1 have a smaller inter-site distance which is 500 meters and the inter-site distance in case 3 is 1732 meters.

These formulas can be used in 3-sector antenna simulations. However, there is a set of measured antenna gain which could provide more accurate simulation result. Fig. 2.1 illustrates the difference between calculated and measured antenna gain should not be ignored.

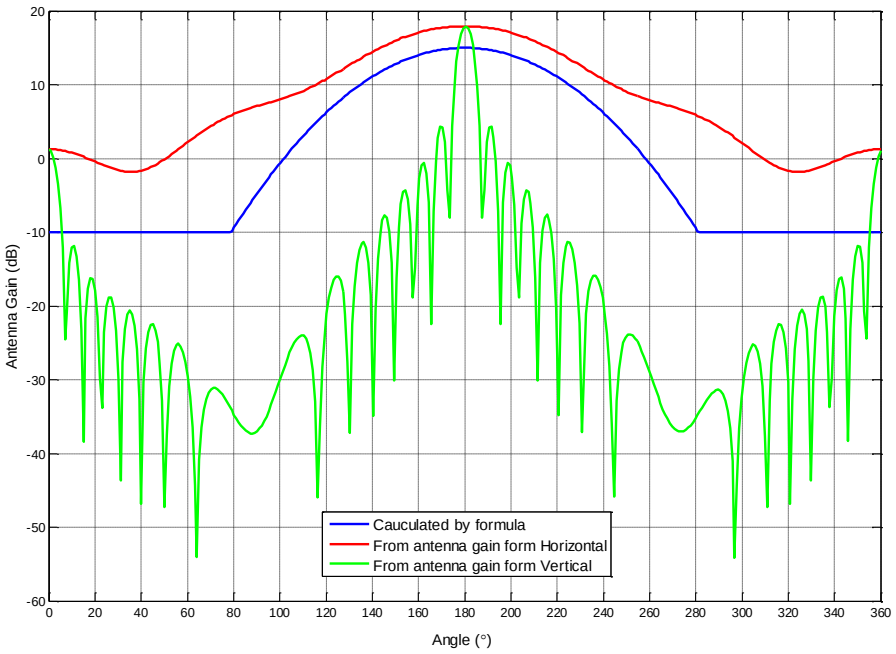


Fig. 2.1 Comparison of calculated and measured antenna gain.

2.3 Review of Mathematical Methodology

2.3.1 Probability Density Estimation

Probability density estimation is a popular method investigating data set and its distribution. [20] introduces density estimation in detail. The primary goal of this method is to construct a density function of measured (observed) data

samples. It is an efficient approach to obtain Probability Density Function (PDF) when studying observed data.

There are different methods to process density estimation, histograms, the naive estimator, the kernel estimator, the nearest neighbouring method, variable kernel method, orthogonal series estimators, maximum penalized likelihood estimators and so on. This research on data analysis uses histograms to estimate density distribution. Therefore, histograms method is introduced in the rest of this section.

Histogram is the oldest and famous way in density estimation. To construct a histogram, we need to first set an origin and a bin width, both of which will control the amount of smoothing inherent in the procedure and will have an effect on the obtained PDF. For example, if the observed data set is $\{y_1, y_2, ..., y_n\}$, the bin width is bw and origin is ori, from origin to the maximum value of this data set, there are *n* levels, where *n* is calculated by:

$$n = \frac{\text{Maximum data value} - ori}{bw} \tag{2.13}$$

For each sample that belongs to this region, it can represent 1 occurrence to the level. The total occurrence of each level illustrates the distribution of this data set. Histogram can be seen as a rough density distribution.

2.3.2 Curve Fitting

Curve fitting as a part of distribution estimation is used to construct the closest distribution function for observed data. Generally, curve fitting can be summarized to three situations of implementing which is based on the awareness of distribution function of original data [21]:

Situation 1: The distribution function is empirically known, and then the target function can be calculated directly.

Situation 2: The distribution function is empirically known but the parameter in this function is unknown, then the parameter values need to be calculated by curve fitting methods.

Situation 3: Both distribution function and its parameters are unknown, the estimation function needs to be modelled by multiple experiments, and parameters need to be calculated.

It is known that the ultimate goal of curve fitting is to minimize the different between observed data and corresponding data. For this goal, a method known as least-squares is used. For instance, if the observed data set is $\{y_1, y_2, \ldots, y_n\}$, the corresponding calculated function values are $f(x_i)$, the residual should be minimized as:

$$\min\{(y_i - f(x_i))^2\} \tag{2.14}$$

The reason to use square value is to make sure the positive and negative residuals are treated equally.

Evaluation of fitted result is essential to curve fitting procedure. There are two metrics I used to evaluate the goodness of curve fitting in my research. The first one is the square of correlation between the measured or empirical value and the predicted value, and is referred to as R-square [22]. If $\{y_1, y_2, ..., y_n\}$ are n measured or empirical data and $\{\hat{y}_1, \hat{y}_2, ..., \hat{y}_n\}$ are the n curve-fitted data, R-square can be calculated as follows:

$$\text{R-Square} = 1 - \frac{\sum_{i=1}^{n} w_i (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} w_i (y_i - \bar{y})^2} \tag{2.15}$$

where $\bar{y}$ is the mean value of $\{y_1, y_2, ..., y_n\}$, and w_i ($i = 1, ..., n$) are the weighting factors. R-square takes value in the interval from 0 to 1. For a R-square value larger than 0.9, the curve fitting is usually considered to be good enough. However, it is not always the higher R-square value the better in practice. Implementation and complexity issues also need to be considered.

The second curve fitting metric is Root Mean Squared Error (RMSE) [23], which is defined as:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{2.16}$$

where n is the number of measurement samples. RMSE represents the total deviation of the response values from the fit to the response values. A value closer to 0 indicates that the model has a smaller random error component, and that the fit will be more useful for prediction. For a good fitting, the value of RMSE is usually less than 0.2.

2.3.3 Gaussian Mixture Model

Gaussian Mixture Model (GMM) is one of the mature mixture models in statistics. Mixture models are the combination of multiple density functions to construct a multimodal density model [24]. GMM in my research is used to construct a function to fit a multi-peak curve. A GMM can be expressed as:

$$\mathrm{GMM} = \sum_{i=1}^{n} a_i e^{-\left(\frac{x-b_i}{c_i}\right)^2} \tag{2.17}$$

where a_i is the amplitude, b_i is known as the centroid location, c_i relates to the peak width, and n represents the number of peaks of the data series. For example, Fig. 2.2 shows a four-peak GMM and its component Gaussian functions.

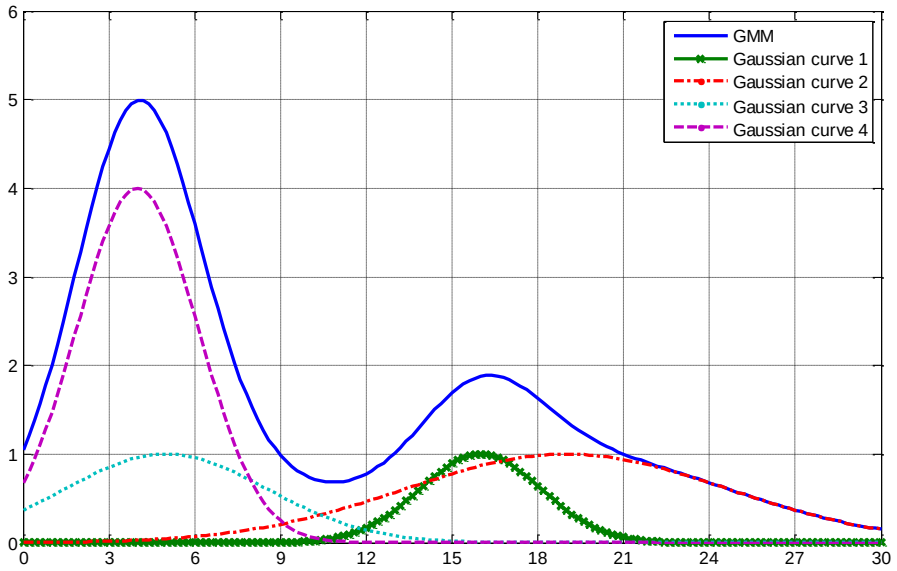


Fig. 2.2 An example of GMM.

The reason GMM is used in my research is because the real measured data is multi-peak and multi-density. GMM can construct an appropriate model to fit it. Moreover, the characteristics of a GMM are controlled by the parameters of each component Gaussian function, which is easy in operation.

2.3.4 System Modelling Method

There are two types of system models that can be used in HetNet analysis and simulation. One is the traditional grid model, which is considered as a well-planned network environment. The other one is a completely random model. Generally, these two types of system model can be used in most simulation conditions. Fig 2.3 is an example the graph of traditional grid model. A widely used model of this is only drawing three layers of cells which contain nineteen cells.

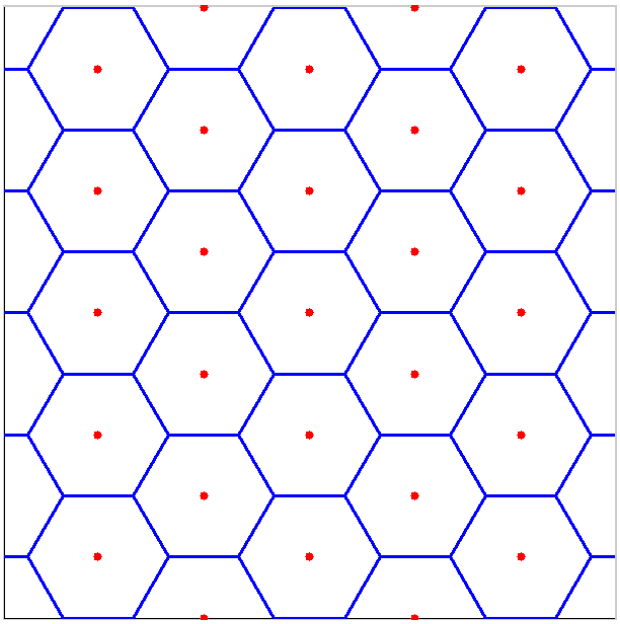


Fig. 2.3 Traditional grid model.

Traditional grid model is very popular in simulation, where the femtocell APs are distributed randomly in each macrocell. As the location of each macrocell base station is fixed, interference from macrocells in this model is calculated relatively simple. Traditional grid model provides an ordered environment which can be used in preliminary scheme study. However, traditional grid model is significantly distinct from real network, especially in urban-areas. Due to competition among multiple operators and complicated environment effect, the deployment of base stations is anomalous. As a result, it is insufficient only applying traditional grid model in simulation.

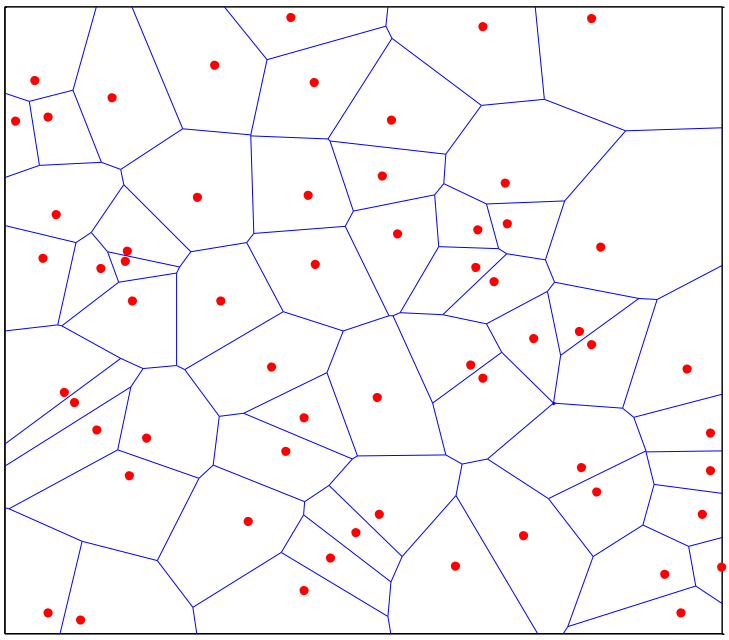


Fig. 2.4 Completely random model

Fig 2.4 shows an example of completely random model. Completely random model is acknowledged closer to real network. The location of each base station follows Spatial Poisson Points Process (SPPP), special cases such as overlaid macrocell base stations will be illustrated in this model. Due to the simulation result in this model will be closer to real world, unexpected issues have a higher possibility to be found in advance before implement.

2.4 Related work

2.4.1 Data Traffic Analysis

Data traffic analysis for cellular networks is an essential method to explore the limitation and issues that may occur. Several recent papers have reported data traffic analysis for different purposes. In [25] and [26], data traffic analysis was used to develop dynamic spectrum access approaches by analysing three weeks' voice traffic in an urban area in northern California. These two papers discuss call durations and exponential call arrival model. In [27], the authors modelled each cell as an independent M/G/∞ queue, with the number of users in each cell modelled as an independent Poisson random process, in order to predict the traffic capacity of code-division multiple access (CDMA) networks. User-behaviour measurements of a High Speed Packet Access (HSPA) network were presented in [28], where measurements made at network, cell and user-session

levels were studied separately. The analysis of the network-level data was based on the measured throughputs of only one day. For cell-level data analysis, the empirical cumulative distribution function (CDF) of the throughput per cell was obtained by using Poisson fitting. Moreover, the mean throughput per user-session was modelled as a log-normal distributed random process. In [29], large-scale Third Generation (3G) network data traffic was studied based on CDF of data traffic measurements to provide insights on traffic load, but there lacked an analysis on cell level throughput. Large-scale cellular data traffic measurements are difficult to obtain, as a result, little effort has been made in the literatures to develop statistical models based on real cellular network data to predict user data traffic distributions in time and in space. In [30] the user data traffic analysis is based on a nationwide 3G network data. Spatial correlation has been characterized by using 10 minutes and 1 hour granularities in order to using cross-correlation method to investigate the spatial correlation. Another research based on the location and traffic records data from radio access network and core network in a 3G data network of United States has been carried out in [31]. The analysis shows that the data traffic is significant different for the BS clusters in different geographical regions.

2.4.2 Small Cell Deployment in 4G HetNets

It is known that a HetNet is usually planned based on the expected user distribution and mobile traffic pattern obtained from long-term observations and statistical data. It is important to achieve both good service quality and low cost in HetNet planning and deployment. In [7], the deployment of small cells is optimised to obtain the best trade-off between user QoS and operators' costs. The dynamic small cell deployment strategy in [8] can be used to find out when and where small cells need to be deployed. Another study in 2015 presents a small cell deployment scheme that considers user location varying [32]. An iterative location deployment updating algorithm has been developed to optimise the deployment and resource allocation. Also, in [33], the small cell deployment is optimised by an algorithm based on weighted k-mean mechanism. By a given number of small cells to be deployed, k-means algorithm is used to optimise the locations of the small cells based on the different data demands. The study of small cell deployment is still popular, [34] [35] [36] [37] are recent studies for small cell deployment in different scenarios. In [34], a two-tier network with micro and pico tier has been analysed by the authors in order to optimise the energy efficiency problem. A small cell deployment scheme has been developed in [35] to solve the problem for cells are placed in biased manner. The work in [36] studies the small cell deployment problem by maximizing service time. The authors of [37] use joint optimisation to solve the deployment and small cell management together by maximizing the utility of

users and network operators. The localisation system for two-tier network planning is discussed in [38].

Once a HetNet has been deployed, its radio resources need to be reused among neighbouring cells and the network capacity is limited by inter-cell interference. In [39], interference management is achieved by controlling the number of RBs that can be used by small cells. In [40], the resource allocation and user association strategies are investigated for the orthogonal deployment, co-channel deployment and partially shared deployment of small cells.

When unexpected HS occurs in an existing HetNet causes mobile traffic demand goes beyond the network capacity, additional small cells may need to be deployed on top of the existing HetNet. In this case, the strategies in [7] [8] [41] and the structured deployment strategy in [42] would require the redesign of the overall HetNet to achieve the optimised deployment. In [43], the optimised deployment locations of new small cells are selected from a set of candidate locations, which need to be obtained before the optimisation process, making it less effective for unexpected but reoccurring HSs. Moreover, without constraint of minimum UE throughput, maximizing the sum throughput or average UE throughput [44] might still leave some UEs with dissatisfied QoS.

2.4.3 Small Cell Deployment Based on Stochastic Geometry Analysis

Small cell deployment in HetNets has been considered as an efficient solution to the rapid growth of mobile data demand under limited radio resources. It is anticipated that deploying low-power small cells will increase the area spectral efficiency [3]. In a HetNet, some small cells are deployed by the users, with their locations uncontrollable by the operators. Moreover, once a HetNet has been deployed, persistent HSs of UEs that were not expected in the original network planning may occur, causing extra traffic demand. As the mobile traffic demand goes beyond the network capacity, the QoS of UEs in the HetNet will be affected. In this case, deploying additional small cells on top of the existing HetNet by the operators would become necessary.

Considering cost effectiveness for network operators, it is desirable to optimise the number and locations of additional small cells to be deployed without changing the existing HetNet infrastructure. This relies on a thorough analysis of the HetNet. Stochastic geometry and the theory of random geometric graphs have been used in the analysis and design of wireless networks [45]. In [46], the stochastic geometry is used to analyse the outage probability for multi-cell cooperation. Also, stochastic geometry is used to analyse the coverage probability of cellular system [47] [48] [49]. Stochastic geometry has also been used in deployment strategies, for instance, a recent research in [50] proposes a deployment strategy by applying tools from stochastic geometry in a multi-

tier HetNet. The optimised deployment is obtained by minimizing the area power consumption. Another deployment strategy that considers energy efficiency based on stochastic geometry tools is carried out in [51]. The relation between the average coverage probability and deployment strategy is derived by the stochastic geometry tools.

Poisson point process (PPP) as a popular model tool in HetNet is applied in the research of this work. The inter-nodal distances were modelled using a spatial bivariate PPP in [52]. A recent research on HetNet use independent PPP to model the locations of BSs [53]. Moreover, in [54], the locations of BSs and UEs are modelled by two independent PPPs in HetNet. In [55], Spatial Poisson Point Process (SPPP) is also considered to be used in the urban and suburban regions in HetNets. An isotropic SPPP are introduced in a recent research [56] is also used in two-tier femtocell networks.

Although HS mitigation has been studied for wireless sensor networks [57], most existing works on small cell deployment strategies focused on optimizing the deployment of small cells on top of a conventional macrocell network or the deployment of a whole new HetNet from scratch [7] [8] [41] [42] [44]. There is a lack of strategies for deploying additional small cells on top of an existing HetNet in response to extra traffic demand from HSs that were not expected in the original network planning. It is especially challenging to optimise both the number and the locations of additional small cells to be deployed. Small cell

deployment for different scenarios are also popular, for instance, a study of dense small cell deployment is introduced in [58].

Chapter 3

User Data Traffic Analysis

3.1 Introduction

In this chapter, an analysis of 3G network user data traffic is presented. It is important to understand the operation of an existing cellular network. Investigating cellular user data traffic changes over time in an existing network is an efficient method to understanding the characteristics of cellular networks. In this work, a statistical modelling of time-varying throughput per cell and the distribution of instantaneous throughput per cell over different cells based on throughput measurements from a real-world large-scale urban cellular network is provided.

3.2 Measurement Data Set and Methodology

3.2.1 Measurement Data Set

The data set is collected from 1,668 NodeBs of a 3G cellular network located in central London, United Kingdom. A series of real-time DL throughput was recorded every 15 minutes at each of these NodeBs over exactly one week (168 hours). The data traffic measurements are given in terms of DL throughput per cell in Mbit/s. Over a million data samples have been recorded in this period of time as a result. The information of each data sample includes the identity of cell, time of measurement and DL throughput value which represents the instantaneous throughput per cell.

However, the measurement data set does not include any information about the number of active users in each cell, user session duration, or traffic type (e.g., voice or data). Consequently, the research focuses on data analysis at the network level rather than the cell level, to provide insights for macroscopic management of an existing 3G network.

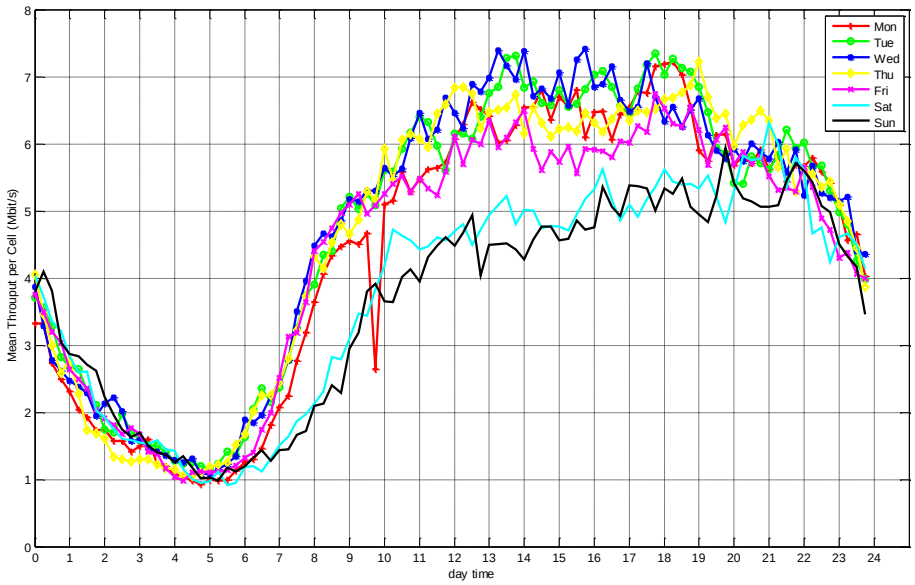


Fig. 3.1 Mean throughput per cell over 24 hours of measured data set from Monday to Sunday.

It is easily visible that the mean throughput per cell on Saturday and Sunday is significantly different from Monday to Friday. According to the observed difference in user traffic demands, the measurement data set has been divided into two groups, weekdays and weekends, respectively, for the subsequent statistical analysis. One of the reasons to separate the data set is people activities in weekdays and weekends have their different manifestations on data demands. And the second reason is that taking the sensitivity of the curve-fitting result into consideration, modelling each day of a week could reduce the generality of the random process. Moreover, the possibility of outliers existing in process of recording the data set cannot be excluded, increasing the number of samples in given time of a weekday or a weekend day can reduce the effect these outliers to the result. In consequence, Fig. 3.2 shows the weekday and weekend mean throughput per cell over 24 hours.

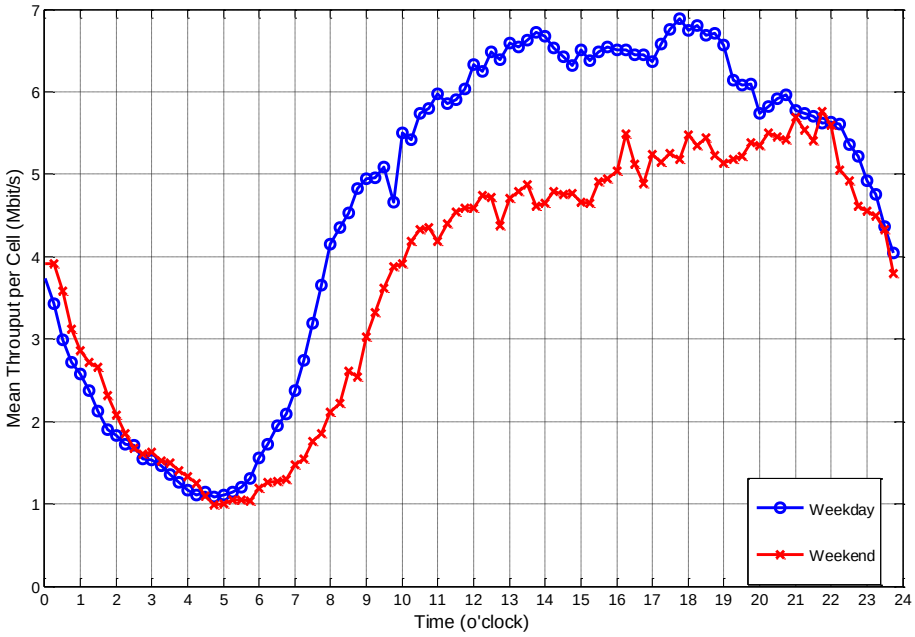


Fig. 3.2 Weekday and weekend mean throughput per cell over 24 hours.

3.2.2 Methodology

The methodology of user data analysis can be divided into following aspects:

I. Density estimation for PDF of the instantaneous throughput of a cell at a given time in a given day.

II. Curve-fitting from PDF and estimate the parameter of fitted density function.

III. Get the distribution over 24 hours.

IV. Curve-fitting the parameters over time and get appropriate function using time as variable.

V. Evaluate the result by reproducing the instantaneous throughput per cell using the fitted distribution model.

Density estimation is the first step in this research. Histograms will be the main method in this process as mentioned in Chapter 2, Section 2.4.1. The origin for histogram is set as 0, and the bin width is present as follows:

$$\text{bin width} = \frac{(T_{max}\text{-}T_{min})}{(T_{max}\text{-}T_{mean}) \times \frac{T_{mean}}{T_{day\ mean}}} \tag{3.1}$$

where T_{max} and T_{min} are the maximum and minimum values of the instantaneous throughput per cell at a given time instant among all NodeBs from Monday to Friday (for a weekday) or over Saturday and Sunday (for a weekend day), T_{mean} is the mean value of all NodeBs' instantaneous throughput at a given time instant average from Monday to Friday (for a weekday) or over Saturday and Sunday (for a weekend day), and $T_{day\ mean}$ is the mean throughput per cell over all sampling time instants from Monday to Friday (for a weekday) or over Saturday and Sunday (for a weekend day) and among all NodeBs.

A varying bin width is used in this density estimation for different time instants of a weekday or a weekend day is because the maximum values of instantaneous throughput per cell over all the NodeBs are significantly different for different time instants of a weekday or a weekend day. For example, the highest mean throughput among all given time is blow 7 Mbit/s, and the highest instantaneous throughput of a single NodeB can be over 300 Mbit/s. For a curve, mean value, somehow, represents the importance of each sample in curve-fitting. Therefore, a weight needs to be added to each sample to maintain the curve-fitting process to limit deviation between the results and measured data on an acceptable level. The bin width given in the function adapts to the maximum instantaneous throughput of the corresponding sampling time instant of a weekday or a weekend day, so that the influence of the maximum instantaneous throughput value to the PDF estimation process can be reduced.

According to the estimated PDF of each NodeB, curve fitting process has been carried out. In this process, the main method of curve fitting is least-squared. This method involves both linear and non-linear curve fitting models. The curve fitting procedure of each given time is evaluated by the two metrics mentioned in Chapter 2, Section 2.4.2, R-square and RMSE.

The PDF approximately follows the exponential distribution, which only have one parameter. This represents each given time will have its own parameter. As a result, these parameters are addressed into a time-varying model which is a multi-peak curve. In order to get a function using time as variable, GMM is applied to construct the time-varying distribution function for parameters for the fitted PDF of instantaneous throughput per cell.

The evaluation of fitted result is executing by comparing of reproduced and measured throughput per cell in cell-level Cumulative Distribution Function (CDF) and mean throughput over 24 hours for a weekday or a weekend day, respectively.

3.3 Modelling and Results

The first step is to investigate how cellular DL throughputs change over time, which is illustrated in Fig. 3.2. The mean throughputs per cell versus different time instants over 24 hours for weekday and weekend are showed separately. For any given time-instant, the weekday mean throughput per cell was obtained by averaging the throughputs measured at that time over all the 1,668 NodeBs and over 5 workdays; while the weekend mean throughput per cell was obtained by averaging the throughputs at that time instant over all the 1,668 NodeBs and over Saturday and Sunday.

From the mean throughput varying over time, there are several features reveals. For weekdays, the mean throughput per cell has the minimum value at 5:00 and hits the maximum at around 17:45. For weekends, the minimum mean throughput per cell occurs at 5:00 while the maximum occurs at 22:00. Both weekday and weekend have the lowest mean throughput per cell of around 1 Mbit/s, but the maximum mean throughput per cell is nearly 7 Mbit/s for weekdays as compared with 5.7 Mbit/s for weekends. Moreover, from 6:00 to 21:00 the mean throughput per cell in a weekday is much higher than that in weekends, which is in accordance with the office hours from Monday to Friday.

According to the observed difference in user traffic demands between weekdays and weekends, the measurement data set has been divided into two parts for weekdays and weekends, respectively. In addition to the time-varying nature of cellular traffic load, it is anticipated that different cells would have different traffic loads at any given time. Fig. 3.3 shows the histograms of throughput per cell at 5:00 and 17:45 on a weekday, which were obtained based on the throughput measurements of all the 1,668 NodeBs at these two time-instance from Monday to Friday.

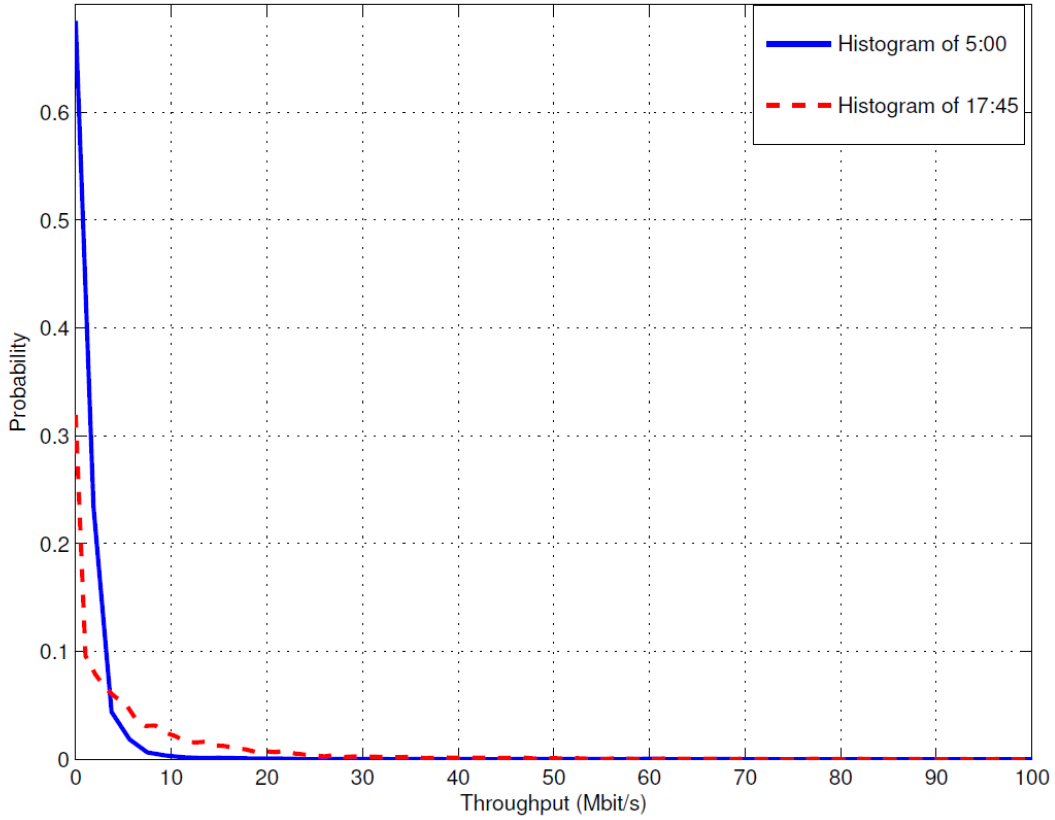


Fig. 3.3 Histogram of throughput per cell at 5:00 and 17:45 on weekdays.

As mentioned in Section 3.2.2 the probability distribution of throughput per cell at the considered time instants approximately follows the exponential distribution. The PDF of an exponential distribution is given by:

$$f(x) = \lambda e^{-\lambda x}, x \geq 0. \tag{3.2}$$

where λ is called the rate parameter for exponential distribution.

In order to find the PDF of exponential distribution that closely matches the probability distribution of throughput per cell at a given time, curve fitting is performed for the histogram of instantaneous throughput per cell using the least-squares method where the values of λ are calculated.

As mentioned above, the measurement data set contains throughput samples recorded at 1,668 NodeBs every 15 minutes over one whole week, all the available throughput samples have been group into two sub-sets: one for the five weekdays, and the other for the two weekend days. There are 96 sampling time instants each day. We assume that the probability distribution of throughput per cell at any time of a day is the same for each weekday (i.e., Monday to Friday), and is the same for each weekend day (i.e., Saturday and Sunday), respectively. We generate histograms and then perform curve fitting to get 96 PDFs for the 96 sampling time instants of a weekday based on the throughput samples of the five weekdays, and do the same for the 96 sampling time instants of a weekend day based on the throughput samples of Saturday and Sunday. R-square and RMSE have been evaluated for all the 192 PDFs obtained from curve fitting, where each obtained PDF has a R-square value higher than 0.9 and a value of RMSE lower than 0.2. Fig.3.4 compares the real-data-based histogram and the curve-fitting obtained exponential probability distribution of throughput per cell for 5:00 and 17:45 of a weekday, where $\lambda = 0.9891$ for 5:00, while $\lambda = 0.1451$ for 17:45.

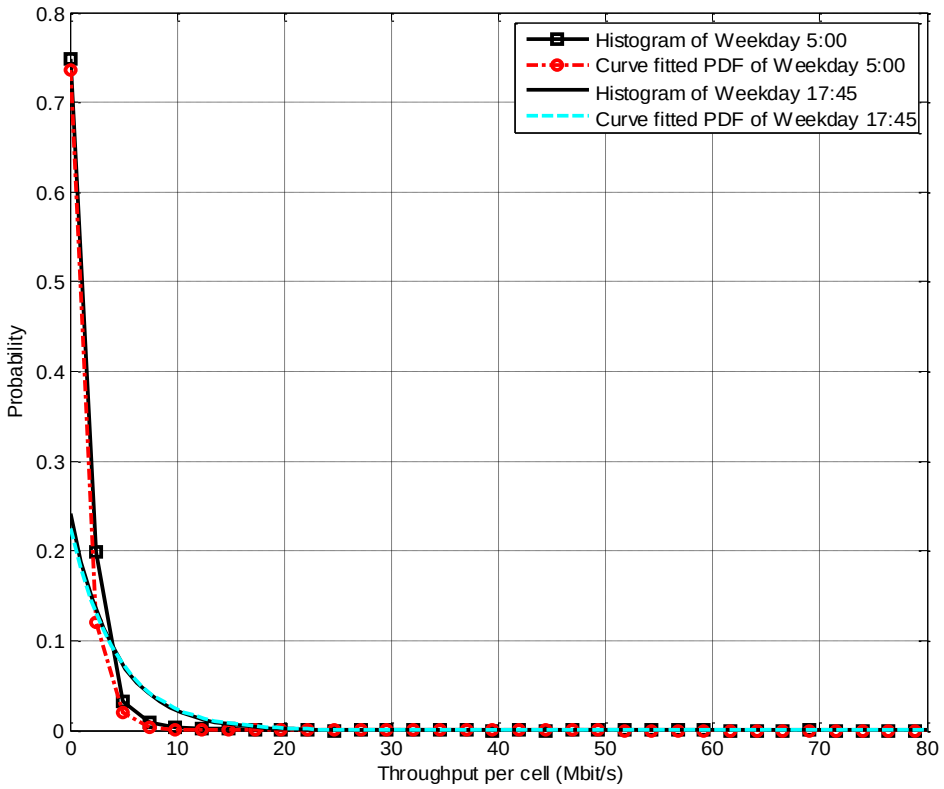


Fig. 3.4 Comparison of real-data-based histogram and curve-fitted exponential probability distribution for a weekday.

Similar comparisons are presented in Fig.3.5 for 5:00 and 22:00 of a weekend day, where λ= 0.8673 for 5:00, while λ= 0.1951 for 22:00. From both figures we can see that the curve-fitting obtained exponential distribution closely matches the corresponding real-data histogram of throughput per cell.

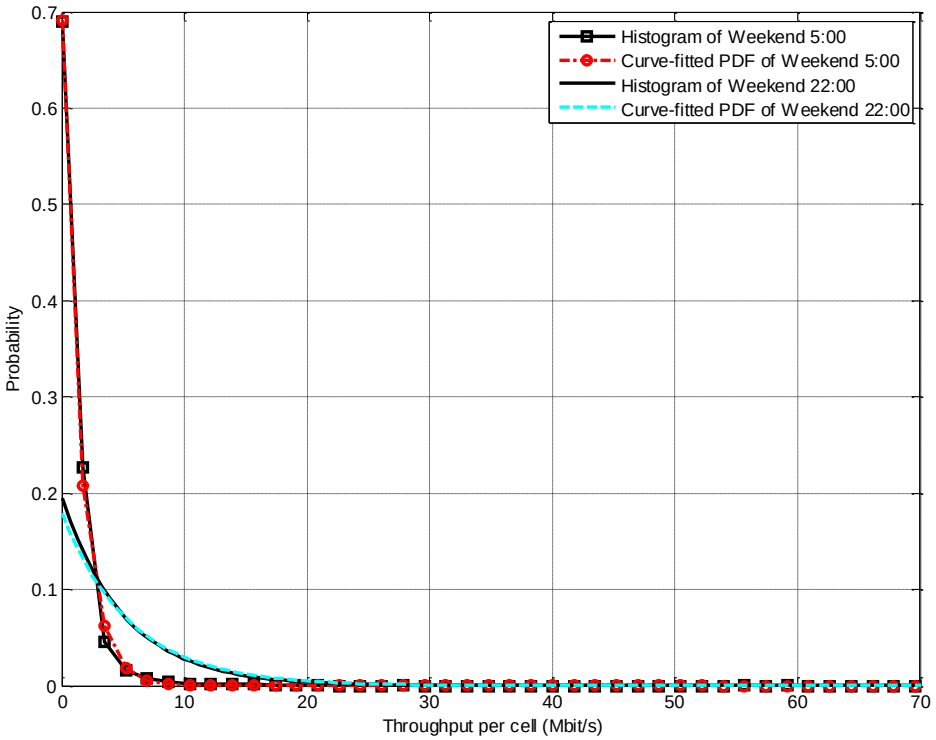


Fig.3.5 Comparison of real-data-based histogram and curve-fitted exponential probability distribution for a weekend day.

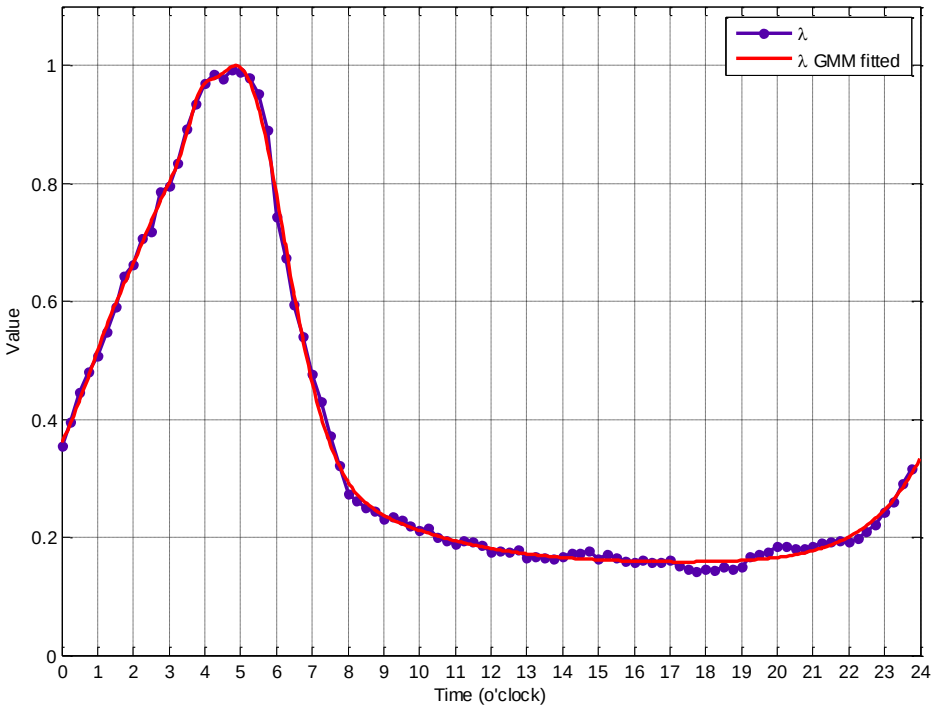
The curve-fitting results show that the throughput per cell follows a different exponential distribution (with different values of λ) for each different time of a day (weekday or weekend). Since the 3G network traffic data set contains instantaneous throughputs measured at each cell every 15 minutes over a whole week, the process of curve-fitting based on the data set can produce 96 exponential PDFs for the 96 sampling time instants of a weekday and another 96 exponential PDFs for the 96 sampling time instants of a weekend day. As a result, the statistical results need to be presented in 192 exponential PDFs, i.e., 192 different values of λ. This is not convenient to regenerate cellular throughput data, e.g., for system-level simulations that require time-varying instantaneous throughputs at all cells of a network. In order to facilitate the use of our curve-fitting results, the function that can model how λ varies with time

is constructed, so that the exponential distribution PDF of throughput per cell can be easily generated for any given time of a weekday or a weekend day.

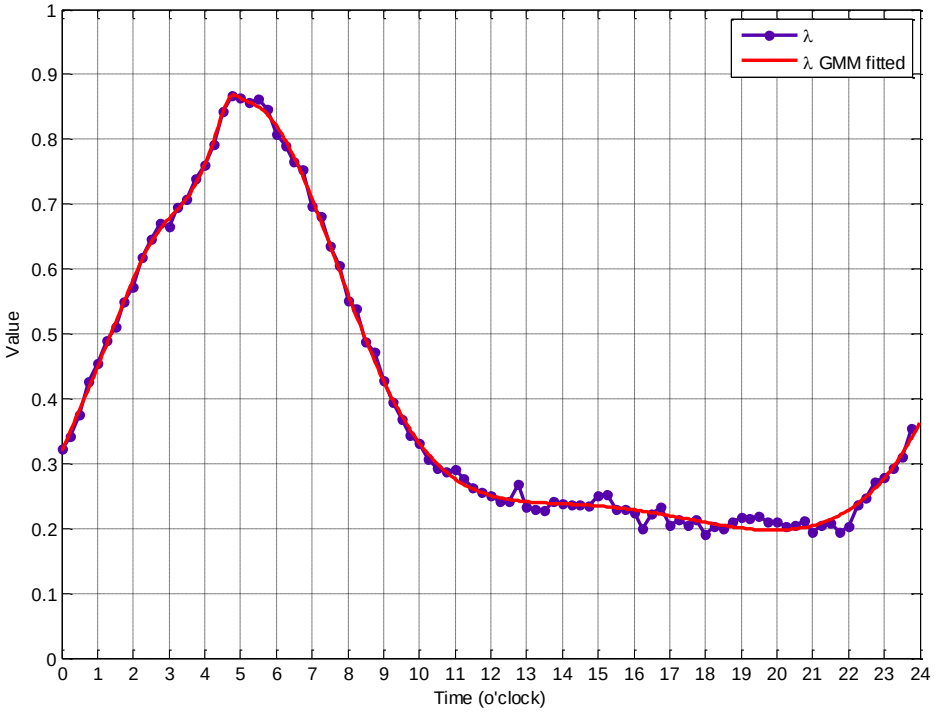
In Fig.3.6, the previously obtained λ is plotted as a function of time for both weekday and weekend. According to the λ curves plotted in Fig.3.6, GMM in (2.18) is used to model the time varying behaviour of λ:

$$\lambda = \sum_{i=1}^{n} a_i \times e^{-\left(\frac{x-b_i}{c_i}\right)^2} \tag{3.3}$$

for which the variable x in (3.3) represents time and can be any time instant in the 24 hours of a day. The results of curve fitting for λ on weekday and weekend are also included in Fig. 3.6, Table I presents the fitted values of a_i, b_i, c_i and *n*. The evaluation of this curve-fitting is similar with previous PDF curve-fitting process. R-squares and RMSE values are calculated to evaluate the curve fitting for λ. For both weekday and weekend, R-square values are larger than 0.99 and RMSE values are less than 0.1. As mentioned in Chapter 2, Section 2.4.2, this indicates that the fitted results are extremely accurate.



(a) λ curve fitting result for a weekday



(b) λ curve fitting result for a weekend day

Fig. 3.6 λ curve fitting results.

TABLE 3.1: Curve-fitting parameters for λ.

i	$\lambda_{weekday}$				$\lambda_{weekend}$			
	n	a_i	b_i	c_i	n	a_i	b_i	c_i
1		3.263×10^6	75.27	12.56		0.02779	4.687	0.3328
2		0.0781	3.85	0.4829		0.1302	5.841	2.755
3		0.6616	4.971	1.77		0.2285	14.23	8.285
4	7	0.1097	2.996	0.862	7	-0.06644	3.656	0.9236
5		0.2584	1.868	1.543		2.698	3.777	4.321
6		0.1822	3.221	5.972		-2.02	3.629	4.178
7		0.1652	-2.871	84		8.16×10^6	106.8	20.03

In order to evaluate the usefulness of our statistical models for cellular network simulations, the DL throughputs of 1,668 cells has been reproduced by using the statistical models and compare them with the DL throughput measurements. The evaluation is separated in two parts, CDF comparison and mean throughput over 24 hours' comparison.

The comparison between the empirical CDFs of measured throughput per cell and statistical-model reproduced throughput per cell at 17:45 on weekdays is depicted in Fig.3.7. It can be seen that the constructed statistical model slightly underestimates (but still closely matches) the measured throughput pre cell.

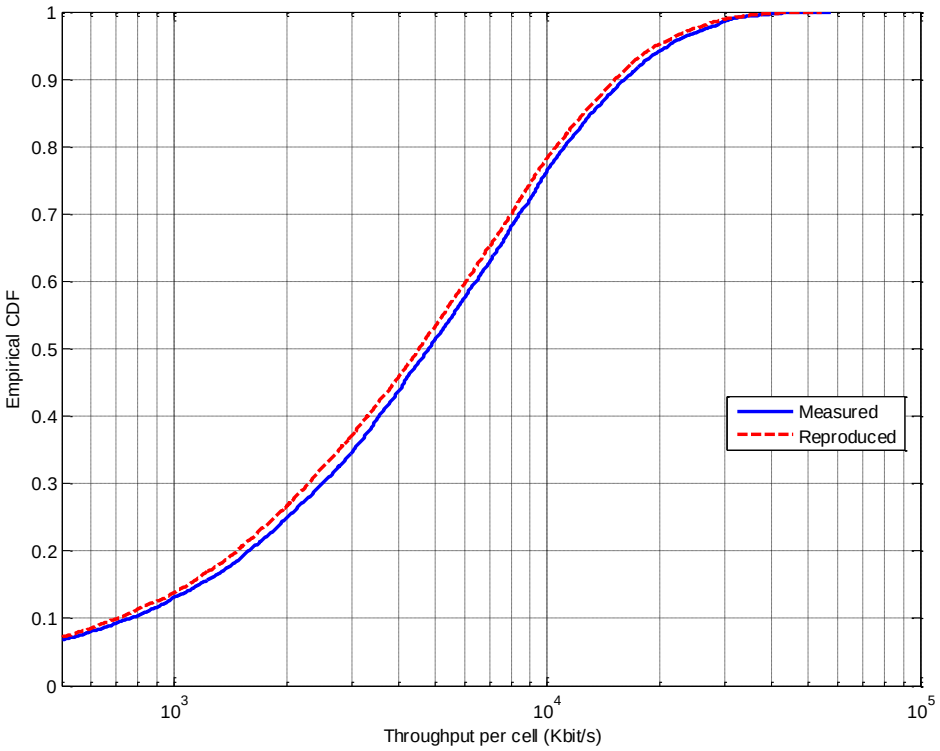
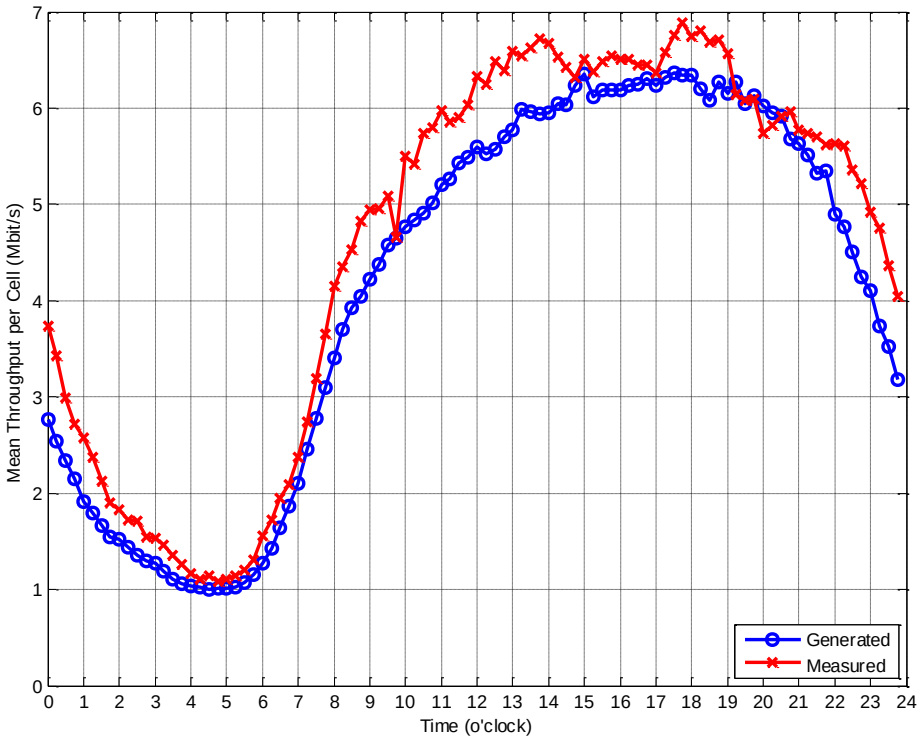
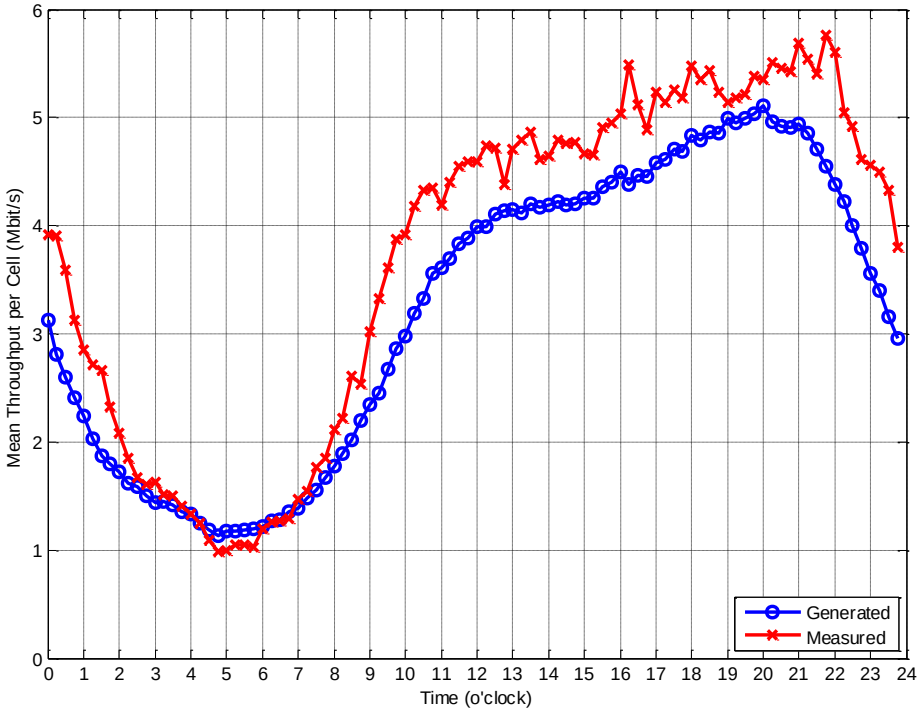


Fig. 3.7 Empirical CDFs of measured throughput per cell and reproduced throughput per cell at 17:45 on weekdays.

The comparison of mean throughput between reproduced throughput and measured throughput over 24 hours in a weekday and a weekend day are illustrated in Fig.3.8 (a) and Fig.3.8 (b), respectively. From the result, we can see that although there are gaps between the mean throughputs per cell of reproduced data and that of the measured data, the model-reproduced data traffic catches the time-varying trend and spatial distribution of the measured data traffic. Hence, the statistical models could be used in simulations of urban-area large-scale cellular networks.



(a) Comparison of measured and reproduced mean throughputs per cell of a weekday.



(b) Comparison of measured and reproduced mean throughputs per cell of a weekend day.

Fig. 3.8: Comparison of reproduced mean throughput and measured mean throughput.

3.4 Conclusion

In this work, the DL throughput per cell distributions over time and over different cells of an existing 3G network based on real network throughput data have been analysed and statistically modelled. By inspecting the large set of real cellular throughput data, a characteristic that the throughput per cell during office hours of a weekday is much higher than that of a weekend has been found. Accordingly, statistical models of cellular throughput for weekdays and weekends are separately constructed. Another characteristic that the instantaneous throughput per cell follows exponential distribution has also been found. The exponential distribution PDFs of instantaneous throughputs per cell at different sampling time instants and the time-varying function of exponential PDF parameter λ have been provided. The accuracy of our statistical models is verified by evaluating the corresponding R-square and RMSE values, and by comparing the model reproduced throughputs with real network measurements. The results show that the proposed statistical models can be used to simulate the time-varying and location-varying throughputs of cells in an urban cellular network as part of my future research on optimisation of HetNet.

Chapter 4

Small Cell Deployment for 4G Heterogeneous Networks

4.1 Iterative Deployment Algorithm of Small-cells in Heterogeneous Networks

4.1.1 Introduction

In this chapter, a strategy has been proposed which is used to optimise the number and locations of additional mobile small cells on top of an existing HetNet to fulfil the excess traffic demand of recurring HSs that were not expected in the original HetNet planning. The mobile small cells can be mounted on vehicles so that they can be deployed by the operator in response to the HSs in a timely manner. It is assumed that the number and locations of the HS UEs are statistically known by the operator. Since the time scales of

deploying mobile small cells and user association are much longer than that of resource allocation, round-robin and even resource allocation are applied to UEs in each cell for simplicity. The objective of the strategy is maximizing the minimum UE throughput through a joint optimisation of the number and locations of additional mobile small cells and the user associations of all cells. The simplified optimisation problem is still NP-hard. Hence, a Fixed Number Deployment Algorithm (FNDA) to optimise the locations of mobile small cells and user associations for a given number of mobile small cells is proposed first. Afterwards, FNDA is extended into the Deployment over Existing Network Algorithm (DOENA) to jointly optimise the number and locations of mobile small cells deployment together with the user associations of all cells. Performance of the proposed DOENA is evaluated in terms of minimum user throughput and number of mobile small cells required through simulations. The relationship between the number of required mobile small cells and the number of HS UEs is discussed. The comparison between deployments from maximizing minimum UE throughput and maximizing sum UE throughput [44] is also included in the simulation.

4.1.2 System Model

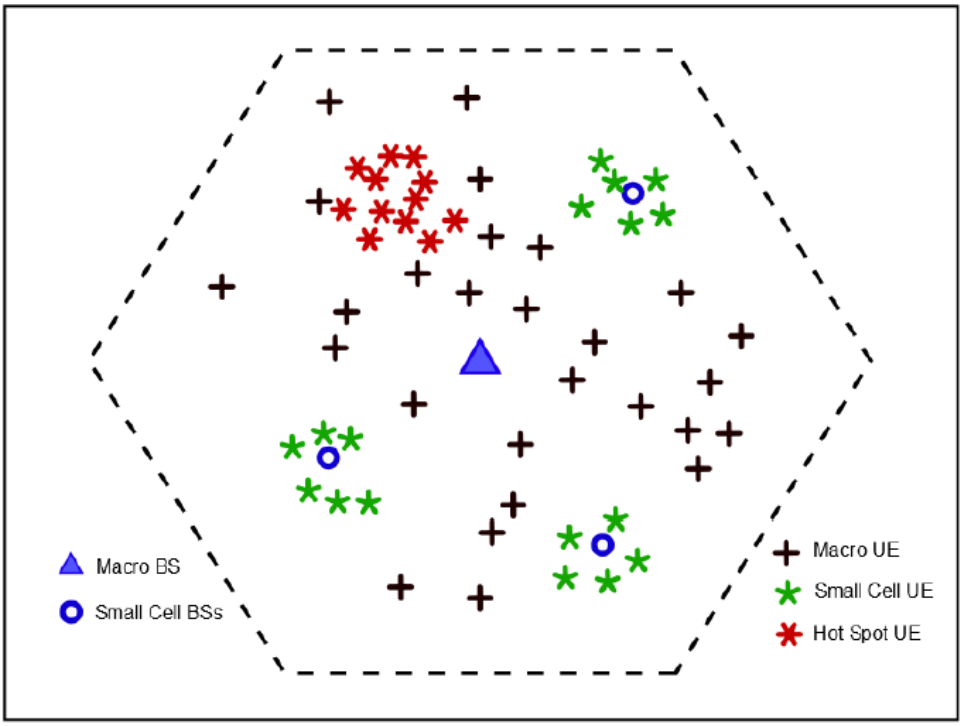


Fig. 4.1 An instance of the problem scenario.

This work studies the DL of a two-tier HetNet consisting of one central macrocell and N_F small cells, all sharing the same spectrum. With one HS randomly distributed in the macrocell coverage area, all non-HS UEs are uniformly distributed within the coverage area of their serving cell, and the HS UEs are uniformly distributed in the HS area. The coverage area of each macrocell or each small cell is assumed to be a disk area centred at the macro or small BS with a certain radius. The HS area is also assumed to be circular for simplicity. Denote the coverage area of the macrocell as H. The total number of existing BSs is $N_{eBS} = 1 + N_F$. Denote the number of mobile small cells to be deployed as N_{nBS}, and the total number of all BS would be $N_{BS} = N_{eBS} + N_{nBS}$. Each cell has access to the total of N_{RB} RBs. Let N_{nhs} denote the number of non-HS UEs and N_{hs} denote the number of not expected but reoccurring HS

UEs. The total number of UEs is given by $N_U = N_{nhs} + N_{hs}$. The set of all UEs is denoted as $N_U = \{1,2,...,N_U\}$, the set of BSs is denoted as $N_{BS} = \{1,2,...,N_{BS}\}$ and the set of RBs is denoted as $N_{RB} = \{1,2,...,N_{RB}\}$.

The throughput of the i th UE is given by:

$$\gamma_i = \sum_{k=1}^{N_{RB}} \sum_{j=1}^{N_{BS}} B \cdot \log_2(1 + \frac{P_j^k \cdot g_{i,j}^k \cdot a_{i,j}^k}{I_{i,j}^k + N_0}), \tag{4.1}$$

where $a_{i,j}^k = 1$ if the i th UE is served by the i th BS in the k th RB, $a_{i,j}^k = 0$ otherwise; B is the bandwidth of a RB, P_j^k is the DL transmit power of the j th BS in the k th RB, $g_{i,j}^k$ is the channel power gain of the link between the i th UE and the j th BS in the k th RB and can be expressed as:

$$g_{i,j}^k = g_{f,ij}^k \cdot g_{pl,ij} \tag{4.2}$$

where $g_{f,ij}^k$ is the exponentially distributed fading gain with unit mean, and $g_{pl,ij}$ is the path loss gain given by [18]:

$$g_{pl,ij} = -15.3 - \alpha \cdot 10 log_{10}(\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2})\text{dB} \tag{4.3}$$

where (x_i, y_i) are the location coordinates of the i th UE and (x_j, y_j) are the coordinates of the j th BS; α is the path loss distance exponent; N_0 is the additive white Gaussian noise (AWGN) power; and $I_{i,j}^k$ is the interference power received by UE i in the k th RB from BSs other than BS j i.e.,

$$I_{i,j}^k = \sum_{\substack{i'=1, \\ i' \neq i}}^{N_U} \sum_{\substack{j'=1, \\ j' \neq j}}^{N_{BS}} P_{j'}^k \cdot g_{i,j'}^k \cdot a_{i',j'}^k \tag{4.4}$$

4.1.3 Mobile Small Cell Deployment Optimisation

I Joint Optimisation

The joint optimisation is defined as maximizing the minimum UE throughput among all UEs over the number and locations of mobile small cells, user associations of all cells, and resource allocation in each cell. That is,

$$\arg\max_{\mathbf{A},(\mathbf{x},\mathbf{y})} \min_i \{\gamma_i\}, i \in N_U \tag{4.5}$$

$$\text{s.t.} \quad \gamma_i > \gamma_{th}, \forall i \in N_U \tag{4.6}$$

$$(\mathbf{x}, \mathbf{y}) \in |H| \tag{4.7}$$

$$\sum_{j=1}^{N_{BS}} \min(\sum_{k=1}^{N_{RB}} a_{ij}^k, 1) = 1, \forall i \in N_U \tag{4.8}$$

$$\sum_{k=1}^{N_{RB}} a_{ij}^k \le N_{RB}, \forall i \in N_U, j \in N_{BS} \tag{4.9}$$

$$\sum_{i=1}^{N_U} a_{ij}^k \le 1, \forall j \in N_{BS}, k \in N_{RB} \tag{4.10}$$

$$a_{i,j}^k \in \{0,1\}, \forall i \in N_U, \forall j \in N_{BS}, k \in N_{RB} \tag{4.11}$$

where **A** is the $N_U \times N_{BS} \times N_{RB}$ matrix that contains all $a_{i,j}^k$ as elements, (**x**, **y**) are the coordinate vectors that contain locations of all mobile small cells, **x** and **y** are each of size $N_{nBS} \times 1$, γ_i is the throughput of the i th UE as given in (4.1), without loss of generality, the macro BS is denoted as the first BS (i.e., $j = 1$),

(4.6) guarantees that each UE throughput is beyond the threshold γ_{th}, (4.7) limits the deployment area for small cells with respect to the macro BS, $|H|$ is the feasible deployment area for mobile small cells as presented in [59] with the exclusion of coverage overlap between any two small cells. (4.8) guarantees that each UE is served by one BS, (4.9) limits the number of RBs each UE can be allocated, (4.10) ensures that each RB can be allocated to at most one UE in each cell, and (4.11) is the binary indicator constraint.

II Decomposed Optimisation

In (4.5), **A** is a UE-to-BS and UE-to-RB joint association matrix. The joint optimisation over the number and locations of mobile small cells, user association to all cells, and resource allocations per cell has a very high complexity. Therefore, the joint association matrix decomposed into two matrixes: 1) user association matrix, and 2) resource allocation matrix. Accordingly, the joint association indicator $a_{i,j}^{k}$ is given by:

$$a_{i,j}^{k} = b_{i,j} \cdot c_{i,k} \tag{4.12}$$

where $b_{i,j} = 1$ if UE i is served by BS j, otherwise $b_{i,j} = 0$; $c_{i,k} = 1$ if UE i is allocated with RB k, otherwise $c_{i,k} = 0$.

The throughput of the i th UE can be rewritten as:

$$\gamma_i = \sum_{k=1}^{N_{RB}} \sum_{j=1}^{N_{BS}} B \cdot \log_2 (1 + \frac{P_j^k \cdot g_{i,j}^k \cdot b_{i,j} \cdot c_{i,k}}{I_{i,j}^k + N_0}) \tag{4.13}$$

The interference power received by the i th UE that is served by BS j in the k th RB can be rewritten as:

$$I_{i,j}^{k} = \sum_{\substack{i'=1,\\ i'\neq i}}^{N_U} \sum_{\substack{j'=1,\\ j'\neq j}}^{N_{BS}} P_{j'}^{k} \cdot g_{i,j'}^{k} \cdot b_{i',j'} \cdot c_{i',k} \tag{4.14}$$

The decomposed optimisation problem is defined as:

$$\arg\max_{\mathbf{B},\mathbf{C},(\mathbf{x},\mathbf{y})} \min_i \{\gamma_i\}, i \in N_U \tag{4.15}$$

$$\text{s.t. (4.6), (4.7), } \sum_{j=1}^{N_{BS}} b_{i,j} = 1, \forall i \in N_U \tag{4.16}$$

$$\sum_{k=1}^{N_{RB}} c_{i,k} \leq N_{RB}, \forall i \in N_U \tag{4.17}$$

$$\sum_{i=1}^{N_U} c_{i,k} \cdot b_{i,j} \leq 1, \forall j \in N_{BS}, k \in N_{RB} \tag{4.18}$$

$$b_{i,j}, c_{i,k} \in \{0,1\}, \forall i \in N_U, \forall j \in N_{BS}, k \in N_{RB} \tag{4.19}$$

where $\mathbf{B}$ is the $N_U \times N_{BS}$ UE-to-BS association matrix that contains all $b_{i,j}$ as elements, $\mathbf{C}$ is the $N_U \times N_{RB}$ UE-to-RB association matrix that contains all $c_{i,k}$ as elements, (4.16) guarantees that each UE is served by one BS, (4.17) guarantees that the number of RBs allocated to each UE is not beyond the total available RBs, (4.18) guarantees that each RB is allocated to at most one UE in each cell, and (4.19) is the binary indicators constraint.

III Simplified Optimisation for Mobile Small-cell Deployment

The deployment of additional mobile small cells is for meeting the excessive data demand in the period of HS occurrence, during which the HS UEs' locations are relatively stable. In contrary to resource allocation, the deployment optimisation of mobile small cells and user associations do not need to be updated frequently in this situation. Hence, the main problem can be converted into optimizing the number and locations of mobile small cells by solving for the matrix **B** under the following assumptions.

Assumption 1: It can be assumed that the RBs are allocated in each cell following the round-robin algorithm with full bandwidth allocation [60]. That is, all the N_{RB} RBs are allocated to UEs in each cell following the round robin algorithm at all times, and there will be inter-cell interference in each RB. This can be considered as the worst-case interference scenario.

The number of RBs allocated by the j th BS to the i th UE can be calculated as:

$$N_{i,j}^{RB} = b_{i,j} \sum_{k=1}^{N_{RB}} c_{i,k}, \forall i \in N_U, j \in N_{BS} \tag{4.20}$$

Without loss of generality, the total number of RBs are assumed evenly distributed among the UEs in each cell, and relax the constraint that the number of RBs per UE has to be an integer. Accordingly, the number of RBs per UE in cell j can be written as:

$$N_j^{RB} = \frac{N_{RB}}{\sum_{i=1}^{N_U} b_{i,j}}, \forall j \in N_{BS} \tag{4.21}$$

Assumption 2: Given the relatively long time scales of small cell deployment and user association, it can be assumed that the effect of fast fading has been averaged out.

Since power control in the DL is not considered, a BS uses the same transmit power in all RBs and P_j^k is expressed as:

$$P_j^k = P_j \tag{4.22}$$

where P_j is the DL transmit power per RB of the j th BS. If the DL transmit power per RB of a macro BS and a small BS is given by P_M and P_F, respectively, then $P_1 = P_M$, and $P_j = P_F$ for $j = 2,3,...,N_{BS}$.

In this case, the superscript k can be ignored hereafter and the UE throughput γ_i can be rewritten as:

$$\gamma_i = \sum_{j=1}^{N_{BS}} N_j^{RB} \cdot B \cdot \log_2(1 + \frac{P_j \cdot g_{pl,ij} \cdot b_{i,j}}{I_{i,j} + N_0}) \tag{4.23}$$

Accordingly, the interference power received by UE i in RB when it is served by BS j is given by

$$I_{i,j} = \sum_{j'=1, j' \neq j}^{N_{BS}} P_{j'} \cdot g_{pl,ij'} \tag{4.24}$$

Thus, the mobile small-cell deployment optimisation problem can be simplified as:

$$\arg\max_{\mathbf{B},(\mathbf{x},\mathbf{y})} \min_i \{\gamma_i\}, i \in N_U \tag{4.25}$$

$$\text{s.t. (4.6), (4.7), } \sum_{j=1}^{N_{BS}} b_{i,j} = 1, \forall i \in N_U \tag{4.26}$$

$$N_{RB} \geq N_j^{RB} > 0, \forall j \in N_{BS} \tag{4.27}$$

$$b_{i,j} \in \{0,1\}, \forall i \in N_U, \forall j \in N_{BS} \quad (4.28)$$

where (4.27) guarantees that the number of RBs each UE can be allocated is not beyond the total number of available RBs.

4.1.4 Solving the Optimisation Problem

In (4.25), the size of vectors **x** and **y**, i.e., the number N_{nBS} of mobile small cells to be deployed, is also an unknown variable to be determined. Note that the joint optimisation of the number and locations of mobile small cells together with the user association in (4.25) is a mixed integer programming problem, which is NP-hard. The difficulty of finding the global optimal solution is high due to the computational complexity. In this section, a simple algorithm to solve the optimisation problem (4.25) for a given feasible number of mobile small cells is proposed first, an extended algorithm is proposed to solve the joint optimisation problem in (4.25).

I Fixed Number Mobile Small-cell Deployment Algorithm

Given a feasible value of the number of mobile small cells to be deployed N_{nBS}, the objectives of (4.25) are reduced to finding the deployment locations of mobile small cells and user associations of all cells. In order to solve this simplified problem, a Fixed Number Deployment Algorithm (FNDA) based on the branch and bound (B&B) method is proposed, where the binary constraint of user association indicators is relaxed to $0 \le b_{i,j} \le 1$ [61].

Algorithm 1 FNDA

Initialisation: $b_{i,j}^r \leftarrow 0, \forall i,j$; $(\mathbf{x},\mathbf{y})_r \leftarrow \varnothing$; MAXVAL$\leftarrow 0$

Main Algorithm:

function FNDA

  $g(\mathbf{B}^*,(\mathbf{x},\mathbf{y})^*) \leftarrow$ B&B_SEARCH($\mathbf{B}_r,(\mathbf{x},\mathbf{y})_r$, MAXVAL)

  return $\mathbf{B}^*,(\mathbf{x},\mathbf{y})^*$

end function

B&B Algorithm:

function B&B_SEARCH($\mathbf{B}_r$,$(\mathbf{x},\mathbf{y})_r$, MAXVAL)

  $|\mathcal{H}| \leftarrow |\mathcal{H}|_{(\mathbf{x},\mathbf{y})_r}$

  $(\mathbf{x},\mathbf{y})_r \in |\mathcal{H}|$

  $(g,\mathbf{B}_r,(\mathbf{x},\mathbf{y})_r) \leftarrow$ Solve (4.25) for $\mathbf{B}_r,(\mathbf{x},\mathbf{y})_r$

  if (4.25) $\geq \gamma_{th}$, $\mathbf{B}_r \in \mathbb{Z}^+$ **then**

    if $g >$MAXVAL **then**

      MAXVAL$\leftarrow g(\mathbf{B}_r,(\mathbf{x},\mathbf{y})_r)$

      $\mathbf{B}^* \leftarrow \mathbf{B}_r$

      $(\mathbf{x},\mathbf{y})^* \leftarrow (\mathbf{x},\mathbf{y})_r$

    else if $g \leq$MAXVAL **then**

      return

    end if

  end if

  return MAXVAL, $\mathbf{B}^*,(\mathbf{x},\mathbf{y})^*$

  if (4.25) $\geq \gamma_{th}$, $\mathbf{B}_r \notin \mathbb{Z}^+$ **then**

    for all $b_{i,j}^r \notin \mathbb{Z}^+$ **do**

      B&B_SEARCH($(b_{i,j}^r = 0) \rightarrow \mathbf{B}_r,(\mathbf{x},\mathbf{y})_r$, MAXVAL)

      B&B_SEARCH($(b_{i,j}^r = 1) \rightarrow \mathbf{B}_r,(\mathbf{x},\mathbf{y})_r$, MAXVAL)

    end for

  else if (4.25) $< \gamma_{th}$ **then**

    return

  end if

end function

Denote g as the minimum UE throughput calculated by (4.23), $\mathbf{B}_r$ as the user association matrix with all elements relaxed to $0 \leq b_{i,j}^r \leq 1$, and $(\mathbf{x},\mathbf{y})_r$ as the

location vectors of mobile small cells each with known size of N_{nBS}. The FNDA is presented in Algorithm 1.

The optimal solutions of mobile small cells' deployment locations and user associations will be returned in $\mathbf{B}^*$ and $(\mathbf{x},\mathbf{y})^*$, respectively.

II Mobile Small-cell Deployment over Existing Network Algorithm

For a given operator, there would usually be a maximum number of mobile small cells N_{nBS}^{max} that can be deployed due to cost and infrastructure considerations. The set $\mathbf{S}=\{s_j\}$ is defined as containing all existing BSs and the maximum number of mobile small cells. The size of $\mathbf{S}$ is $N_{BS}^{max}=N_{eBS}+N_{nBS}^{max}$. If BS j has been deployed or is to be deployed, then $s_j=1$; otherwise, $s_j=0$. Obviously, $\forall j\in\{1,2,...,N_{eBS}\}, s_j=1$. The set of maximum numbers of mobile small cells is denoted as $N_{nBS}^{max}=\{N_{eBS},...,N_{eBS}+N_{nBS}^{max}\}$.

Accordingly, (4.23) can be rewritten as:

$$\gamma_i=\sum_{j=1}^{N_{BS}^{\max}} s_j\cdot N_j^{RB}\cdot B\cdot\log_2(1+\frac{P_j\cdot g_{pl,ij}\cdot b_{i,j}}{I_{i,j}+N_0}) \qquad (4.29)$$

In order to solve the joint optimisation in (4.25), the binary constraints on both $b_{i,j}$ and s_j is relaxed to $0\le b_{i,j}\le 1$ and $0\le s_j\le 1$, and put the relaxed indicators in matrixes $\mathbf{B}_r$ and $\mathbf{S}_r$, respectively. The DOENA is presented in Algorithm 2.

Algorithm 2 DOENA

Initialisation:

$b^r_{i,j} \leftarrow 0, \forall i,j; (\mathbf{x},\mathbf{y})_r \leftarrow \varnothing; s^r_j \leftarrow 0, \forall j \in \mathcal{N}^{max}_{nBS}$; MAXVAL$\leftarrow 0$

Main Algorithm:

function DOENA($\mathbf{B}_r, (\mathbf{x},\mathbf{y})_r, \mathbf{S}_r$,MAXVAL)
  $(\mathbf{B}_{\mathbf{S}_r}, (\mathbf{x},\mathbf{y})_{\mathbf{S}_r}) \leftarrow$ B&B_SEARCH($\mathbf{S}_r, \mathbf{B}_r, (\mathbf{x},\mathbf{y})_r$)
  $(g^*_{\mathbf{S}_r}, \mathbf{B}^*_{\mathbf{S}_r}, (\mathbf{x},\mathbf{y})^*_{\mathbf{S}_r}) \leftarrow$ FNDA($\mathbf{B}_{\mathbf{S}_r}, (\mathbf{x},\mathbf{y})_{\mathbf{S}_r}$)
  if (4.29) $\geq \gamma_{th}$, $\mathbf{S}_r \in \mathbb{Z}^+$ **then**
    if $g >$MAXVAL **then**
      MAXVAL$\leftarrow g^*_{\mathbf{S}_r}$
      $\mathbf{S}^* \leftarrow \mathbf{S}_r$
      $\mathbf{B}^* \leftarrow \mathbf{B}^*_{\mathbf{S}_r}$
      $(\mathbf{x},\mathbf{y})^* \leftarrow (\mathbf{x},\mathbf{y})^*_{\mathbf{S}_r}$
    else if $g \leq$MAXVAL **then**
      return
    end if
  end if
  return MAXVAL, $\mathbf{B}^*, (\mathbf{x},\mathbf{y})^*, \mathbf{S}^*$
  if (4.29) $\geq \gamma_{th}$, $\mathbf{S}_r \notin \mathbb{Z}^+$ **then**
    for all $s^r_j \notin \mathbb{Z}^+$ **do**
      DOENA($(s^r_j = 0) \rightarrow \mathbf{S}_r$, MAXVAL)
      DOENA($(s^r_j = 1) \rightarrow \mathbf{S}_r$, MAXVAL)
    end for
  else if (4.29) $< \gamma_{th}$ **then**
    return
  end if
end function

Based on the returned value $\mathbf{S}^*$, the optimal number of mobile small cells can be calculated by:

$$N_{nBS} = \sum_{j=N_{eBS}+1}^{N^{max}_{BS}} s^*_j \tag{4.30}$$

The mobile small cells' deployment locations and user associations will be returned in $(\mathbf{x}, \mathbf{y})^*$ and $\mathbf{B}^*$, respectively.

4.1.5 Simulation Results

In this section, simulation results are presented to evaluate the performance of the proposed deployment strategy of mobile small cells. In each run of the simulation, the N_F existing small cells are independently and uniformly distributed in the coverage area of the macrocell, the location of the HS is randomly generated within the macrocell coverage area following a uniform distribution as well, and the locations of all UEs are randomly generated following the system model in Section 4.1.2. The system parameters used in our simulation are given in Table 4.1.

TABLE 4.1: System Settings for FNDA and DOENA.

Parameter	Value
P_M	46 dBm
P_F	23 dBm
N_{nhs}	50
N_{hs}	[10,40]
N_{eBS}	4
N_F	3
N_{nBS}^{max}	10
N_{RB}	50
γ_{th}	4 Mbps
Macrocell Radius	300m
Small cell Radius	20m
HS Radius	60m
N_0	-174 dB/Hz
α	4

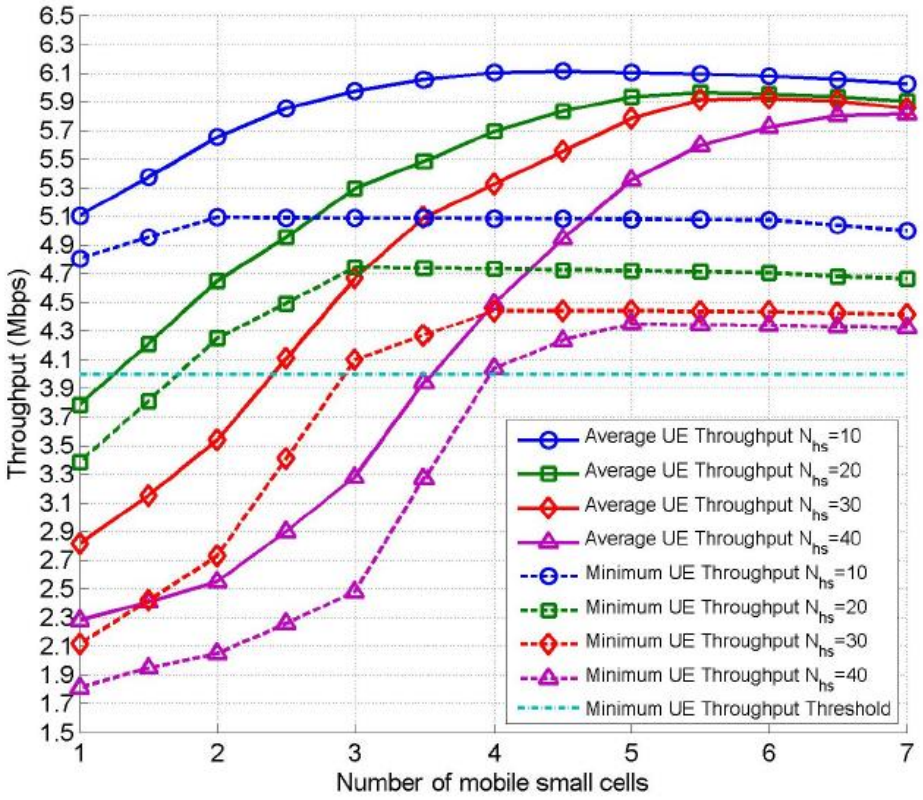


Fig. 4.2 The average UE throughput and minimum UE throughput versus the number of mobile small cells by FNDA.

The minimum UE throughput and average throughput achieved by the proposed FNDA versus the number of mobile small cells are shown in Fig. 4.2. We can see that both average UE throughput and minimum UE throughput increase with the increasing number of mobile small cells, and can exceed the UE throughput threshold when the number of mobile small cells is sufficiently large for a given number of HS UEs. However, as the number of mobile small cells goes beyond a certain value, the increase of UE throughput slows down, and the minimum UE throughput even starts to decrease. This is because, the inter-cell interference becomes dominant and diminishes the capacity gain offered by dense deployment of small cells.

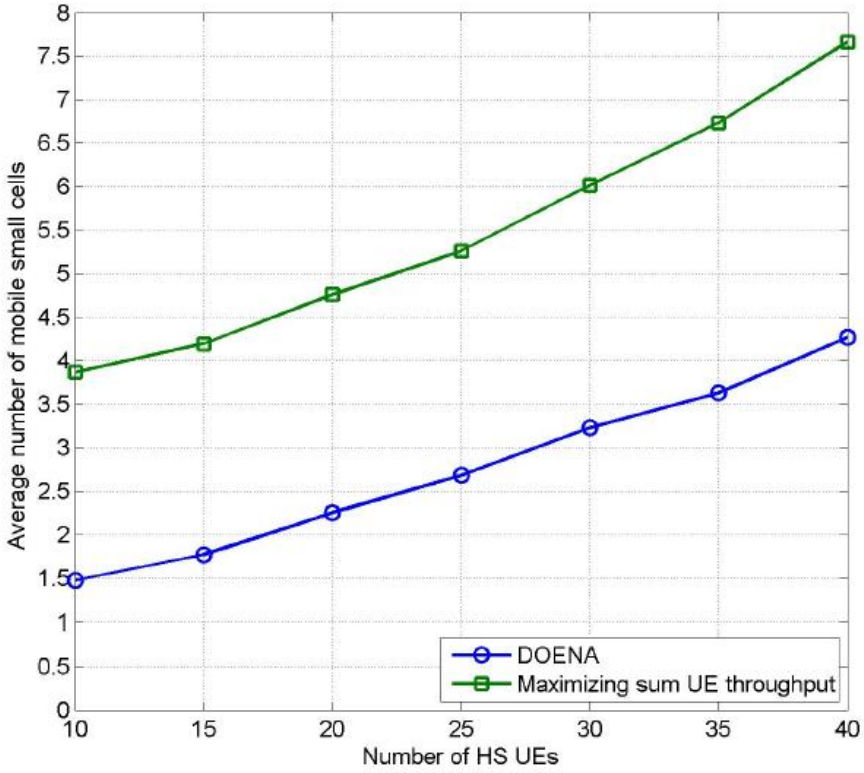


Fig. 4.3 The average number of mobile small cells required by DOENA and by the deployment optimisation based on maximizing sum UE throughput versus the number of HS UEs.

Fig. 4.3 shows the average number of mobile small cells required by the proposed DOENA and by the small cell deployment optimisation based on maximizing sum UE throughput [44] versus the number of HS UEs, under the same minimum UE throughput constraint. Since the scheme based on maximizing sum throughput in [44] optimises the number and locations of all small cells in an iterative manner for given user association, for a fair comparison in the simulation, the user association is optimised using the same method as in DOENA and initialize the iteration with the randomly generated locations of existing small cells. We can see that the optimal number of mobile small cells for maximizing sum UE throughput is larger than that from DOENA for all considered numbers of HS UEs. Deploying less mobile small cells while fulfilling the extra mobile traffic demand of HS UEs will reduce the capital and operational costs of the operator.

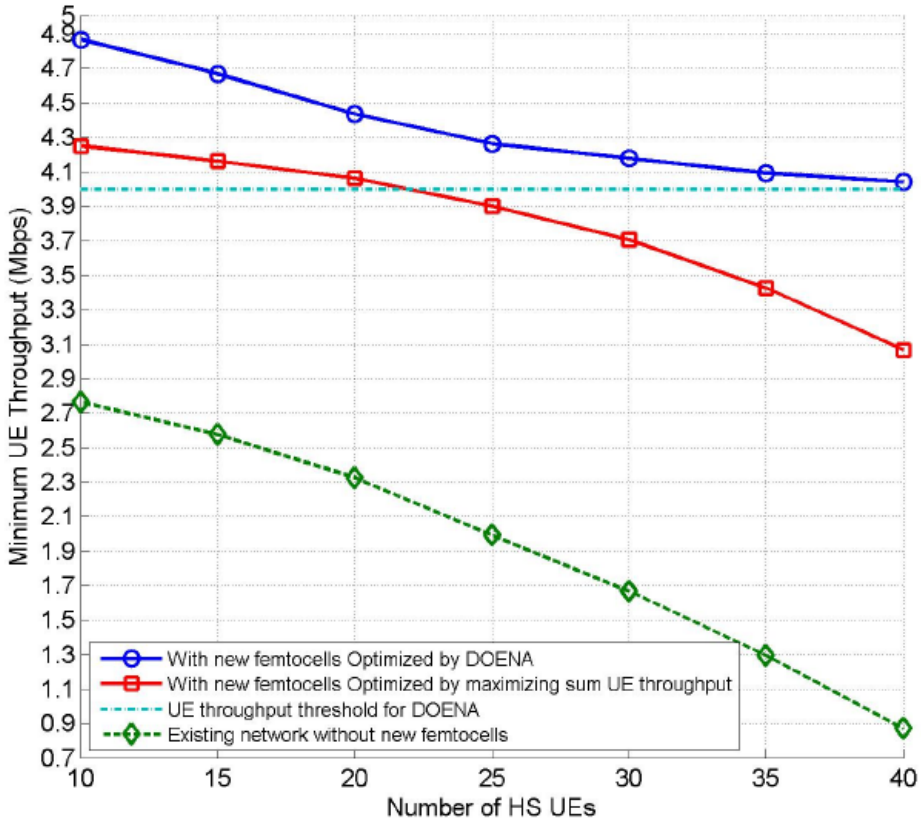


Fig. 4.4 The minimum UE throughput of DOENA and maximizing sum UE throughput versus the number of HS UEs.

Fig. 4.4 plots the minimum UE throughput versus the number of HS UEs before and after the deployment of mobile small cells for DOENA and for maximizing sum UE throughput. It shows that as the number of HS UEs increases, if without mobile small cells, the minimum UE throughput falls below the UE throughput threshold and decreases significantly, indicating that the existing HetNet can no longer fulfil the total traffic demand. The minimum UE throughput of the DOENA is higher than that of the deployment optimised by maximizing the sum UE throughput, and it can be kept above the threshold even for a large number of HS UEs. However, with the number of HS UEs increasing, the minimum UE throughput of deployment optimised by maximizing the sum UE throughput

falls below the threshold, which indicates that the QoS of some UEs cannot be satisfied.

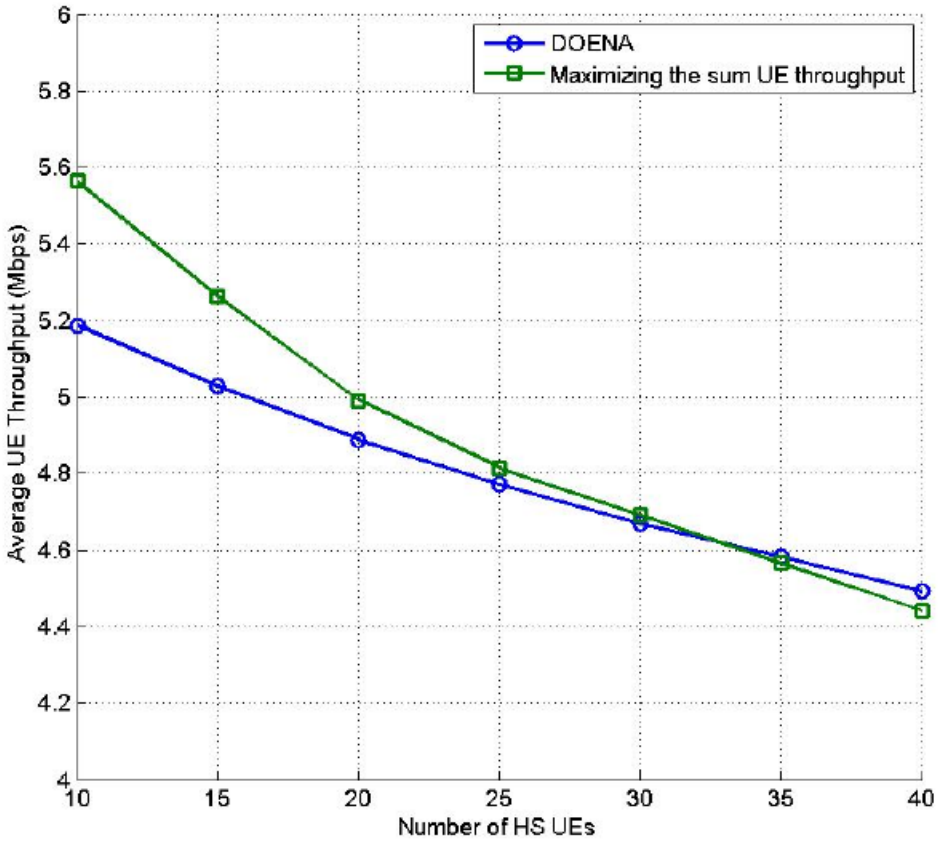


Fig. 4.5 The average UE throughput of DOENA and maximizing sum UE throughput versus the number of HS UEs.

Fig. 4.5 plots the average UE throughput achieved by DOENA and by small cell deployment optimisation based on maximizing sum UE throughput versus the number of HS UEs. We can see that the deployment optimised by maximizing sum UE throughput offers a higher average UE throughput than DOENA when the number of HS UEs is low, but its average UE throughput decreases faster with the number of HS UEs than DOENA. As a result, DOENA outperforms the deployment optimisation based on maximizing sum UE throughput in terms of

average UE throughput (and sum UE throughput) when the number of HS UEs is large.

4.1.6 Conclusion

In this work, a deployment strategy of additional mobile small cells on top of an existing HetNet is proposed in order to fulfil the extra traffic demand of recurring HSs that have not been considered in the original network planning. The optimisation problem is first formulated as maximizing the minimum UE throughput jointly over the number and locations of mobile small cells, user associations of all cells, and resource allocation in each cell. Then, the problem is simplified into a joint optimisation of mobile small cell deployment and user association. By solving the simplified optimisation problem, the FNDA is first proposed to optimise the deployment locations of mobile small cells and the user association for a given feasible number of mobile small cells. By extending FDNA into DOENA algorithm to jointly optimise the number and locations of mobile small cells together with user association. The simulation results have shown that compared with the deployment optimisation based on maximizing sum UE throughput, the proposed DOENA offers higher minimum UE throughput while requiring less mobile small cells to be deployed, and higher average UE throughput when the number of HS UEs is large.

4.2 Reduced-complexity Deployment Algorithm of Small-cells in Heterogeneous Networks

4.2.1 Introduction

The appropriate deployment of small cells is crucial to the success of HetNets. Existing deployment strategies mainly focus on designing the whole HetNet to offer good service quality at low cost [7], [8]. After a HetNet has been established, it is desirable to meet the extra traffic demands from recurring HSs of UEs, which were not expected in the original network planning, by deploying new small cells on top of the existing HetNet without replanning the whole network. Mobile small cells mounted on vehicles can be used in this situation, however, the operators need to know the optimal number and locations of new small cells to be deployed. In this section, an iterative optimisation algorithm is proposed to maximize the minimum UE throughput in a HetNet with unexpected recurring HSs by jointly optimizing the number and locations of new small cells and user associations of all cells. Due to the high computational complexity of the joint optimisation problem, a reduced-complexity iterative algorithm which differs from Section 4.1 is proposed to solve this optimisation. Performance of the iterative new small-cell deployment algorithm in terms of minimum UE throughput and required number of new small cells is evaluated through simulations, in comparison with numerically solving the joint optimisation and the random deployment of new small cells.

4.2.2 System Model

Same with Section 4.1, the DL of a two-tier HetNet consisting of one central macrocell and N_S small cells randomly distributed in the macrocell coverage area is considered as the scenario of this work. The number of existing BSs is

$N_{eBS} = 1 + N_S$. Each cell has access to the total of N_{RB} RBs. Denoting the number of new small cells to be deployed as N_{nBS} , the total number of BSs is given by $N_{BS} = N_{eBS} + N_{nBS}$. Let N_{nhs} denote the number of non-HS UEs and N_{hs} denote the number of HS UEs. The set of all UEs is denoted as $N_U = \{1, 2, ..., N_U\}$, where $N_U = N_{nhs} + N_{hs}$.

Assuming that the channel on each RB sees independent and identical Rayleigh fading, the channel power gain of the link between the i th UE and the j th BS in a RB is expressed as (4.2). And the path loss is given by (4.3).

Since the deployment of new small cells and UE-BS association can be updated at much lower frequencies than radio resource allocation, RBs allocating in each cell are assumed following the round robin algorithm with full bandwidth allocation [18]. That is, all the RBs are allocated to UEs in each cell following the round robin algorithm at all times. This can be considered as the worst-case interference scenario, as there will be inter-cell interference in each RB. The throughput of the i th UE is given by:

$$\gamma_i = \sum_{j=1}^{N_{BS}} N_j^{RB} \cdot W \cdot \log_2(1 + \frac{P_j \cdot g_{i,j} \cdot b_{i,j}}{I_{i,j} + N_0}) \tag{4.31}$$

where $b_{i,j}$ = 1 if the i th UE is served by the j th BS, $b_{i,j} = 0$ otherwise; W is the bandwidth of a RB, P_j is the DL transmit power of the j th BS in a RB, and $N_{RB} \cdot P_j = P_{M(S)}$ if it is a macro (small-cell) BS; N_j^{RB} is the number of RBs per UE in cell j and is given by:

$$N_j^{RB} = \left\lfloor \frac{N_{RB}}{\sum_{i=1}^{N_U} b_{i,j}} \right\rfloor, \forall j \tag{4.32}$$

where $\lfloor \cdot \rfloor$ denotes the floor function; N_0 is the additive white Gaussian noise (AWGN) power; and $I_{i,j}$ is the interference power received by UE i from BSs other than BS j, i.e.,

$$I_{i,j} = \sum_{j'=1, j' \neq j}^{N_{BS}} P_{j'} \cdot g_{i,j'} \tag{4.33}$$

4.2.3 Problem Formulation

For a given HetNet and extra traffic demand from HS UEs not expected in the original network planning, the object is maximizing the minimum UE throughput among all UEs by jointly optimizing the number and locations of new small cells and user associations of all cells. That is,

$$\arg\max_{\mathbf{B},(\mathbf{x},\mathbf{y})} \min_i \{\gamma_i\}, i \in N_U. \tag{4.34}$$

$$\text{s.t. } (\mathbf{x}, \mathbf{y}) \in |H| \tag{4.35}$$

$$\gamma_i \geq \gamma_{th}, \forall i \tag{4.36}$$

$$\sum_{j=1}^{N_{BS}} b_{i,j} = 1, \forall i \tag{4.37}$$

$$N_{RB} \geq N_j^{RB} > 0, \forall j \tag{4.38}$$

$$b_{i,j} \in \{0,1\}, \forall i, j \tag{4.39}$$

where $\mathbf{B}$ is the $N_U \times N_{BS}$ UE-BS association matrix that contains all $b_{i,j}$ as elements; $(\mathbf{x}, \mathbf{y})$ are the $N_{nBS} \times 1$ location vectors of new small cells, (4.35) limits the deployment area of new small cells to the feasible deployment area $|H|$ as presented in [59] while avoiding coverage overlap between any two small cells; (4.36) guarantees that the throughput of each UE is above the threshold γ_{th}; (4.37) ensures that each UE is associated with one BS; (4.38) guarantees that the number of RBs each UE can be allocated is not beyond the total number of available RBs; and (4.39) is the binary constraint on the user association indicators.

4.2.4 Solving the Optimisation Problem

Since the joint optimisation of the number (N_{nBS}) and locations $(\mathbf{x}, \mathbf{y})$ of new small cells together with the user association ($\mathbf{B}$) in (4.34) is a mixed integer programming problem, which is NP-hard, the global optimal solution is difficult to obtain. Therefore, we devise a reduced-complexity iterative algorithm (in Algorithm 3) to solve the optimisation problem in (4.34) based on the branch and bound (B&B) method, where the binary constraint in (4.39) is relaxed to $0 \le b_{i,j} \le 1, \forall i, j$, forming a relaxed matrix $\mathbf{B}_r$. The algorithm iteratively optimises the number of new small cells, starting with the initial value $\sigma = 1$. In each iteration, the optimisation problem in (4.34) is solved with the number of new small cells from the previous iteration (or the initial value); if no feasible solution is obtained, then the number of new small cells is increased by 1; otherwise, the iteration terminates and the optimal solution is returned, where $(\mathbf{x}, \mathbf{y})_\sigma$ denote the location vectors of new small cells with the size of $\sigma \times 1$.

Algorithm 3 Iterative Algorithm

Initialisation: $\sigma = 1; b^r_{i,j} \leftarrow 0, \forall i,j; (\mathbf{x},\mathbf{y})_\sigma, \sigma^* \leftarrow \varnothing$; MAXVAL$\leftarrow 0$

Main Algorithm:

function MAIN ALGORITHM(Loop until $\sigma^* \neq \varnothing$)

3: $g_\sigma(\mathbf{B}^*, (\mathbf{x},\mathbf{y})^*) \leftarrow$ B&B_SEARCH($\mathbf{B}_r, (\mathbf{x},\mathbf{y})_\sigma$, MAXVAL)

return $\mathbf{B}^*, (\mathbf{x},\mathbf{y})^*, \sigma^*$

end function

B&B Algorithm:

6: **function** B&B_SEARCH($\mathbf{B}_r$,$(\mathbf{x},\mathbf{y})_\sigma$, MAXVAL)

$|\mathcal{H}| \leftarrow |\mathcal{H}|_{(\mathbf{x},\mathbf{y})_\sigma}$

$(g_\sigma(\mathbf{B}_r, (\mathbf{x},\mathbf{y})_\sigma)) \leftarrow$ Solve (4.34) for $\mathbf{B}_r, (\mathbf{x},\mathbf{y})_\sigma$

9: **if** (4.34) $> \gamma_{th}$, $\mathbf{B}_r \in \mathbb{Z}^+$ **then**

if $g_\sigma(\mathbf{B}_r, (\mathbf{x},\mathbf{y})_\sigma)$ >MAXVAL **then**

MAXVAL$\leftarrow g_\sigma(\mathbf{B}_r, (\mathbf{x},\mathbf{y})_\sigma)$

12: $\mathbf{B}^* \leftarrow \mathbf{B}_r$

$(\mathbf{x},\mathbf{y})^* \leftarrow (\mathbf{x},\mathbf{y})_\sigma$

$\sigma^* \leftarrow \sigma$

15: **else if** $g_\sigma(\mathbf{B}_r, (\mathbf{x},\mathbf{y})_\sigma) \leq$MAXVAL **then**

return $\sigma = \sigma + 1$

end if

18: **end if**

return MAXVAL, $\mathbf{B}^*, (\mathbf{x},\mathbf{y})^*, \sigma^*$

if (4.34) $\geq \gamma_{th}$, $\mathbf{B}_r \notin \mathbb{Z}^+$ **then**

21: **for all** $b^r_{i,j} \notin \mathbb{Z}^+$ **do**

B&B_SEARCH($(b^r_{i,j} = 0) \rightarrow \mathbf{B}_r, (\mathbf{x},\mathbf{y})_\sigma$, MAXVAL)

B&B_SEARCH($(b^r_{i,j} = 1) \rightarrow \mathbf{B}_r, (\mathbf{x},\mathbf{y})_\sigma$, MAXVAL)

24: **end for**

else if (4.34) $< \gamma_{th}$ **then**

return $\sigma = \sigma + 1$

27: **end if**

end function

4.2.5 Simulation Results

The performance of the proposed iterative algorithm of deploying new small cells on top of the existing HetNet is evaluated through simulations in comparison with the solution obtained by numerically solving (4.34) using the method of exhaustion for optimizing user associations and using the generalized reduced gradient (GRG) method [62] for optimizing new small cells' locations; as well as the random deployment of new small cells following a homogeneous spatial Poisson point process (SPPP). In the simulation, two existing small cells are distributed in the coverage area of the macrocell

following a homogeneous SPPP. 40 non-HS UEs are uniformly distributed in the coverage area of the macrocell, and the N_{hs} HS UEs are uniformly distributed in a circular area with a radius of 40m centred at a randomly selected point in the coverage area of the macrocell. The transmit power of the macro BS and a small-cell BS is 46dBm and 23dBm, respectively. In this work, the parameters are set as $W = 180$ kHz, $N_{RB} = 50$, $N_0 = -174$ dBm/Hz, $\gamma_{th} = 4$ Mbps and $\alpha = 4$, the coverage radius of a macrocell is 300m and that of each small cell is 20m.

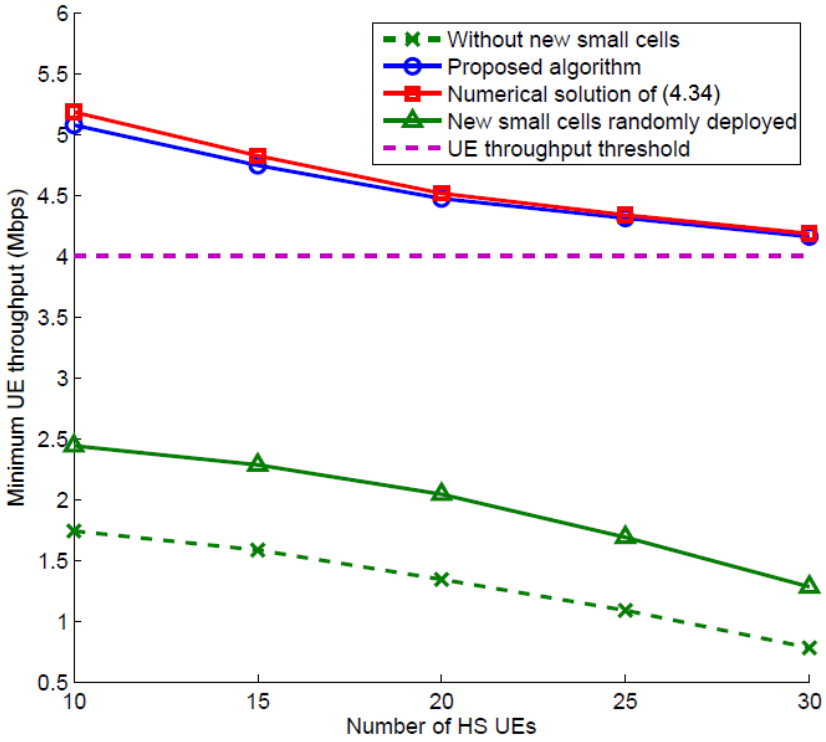


Fig. 4.6 The minimum UE throughput vs. the number of HS UEs.

Fig. 4.6 plots the minimum UE throughput versus the number of HS UEs. It shows that without deploying new small cells, the minimum UE throughput falls below the threshold, because the existing HetNet cannot fulfil the extra traffic

demand from HS UEs. While with new small cells deployed following the proposed algorithm or by numerically solving the joint optimisation in (4.34), the minimum UE throughput can be kept above the threshold even for a large number of HS UEs. If the same number of new small cells (same as our proposed algorithm) are deployed randomly over the macrocell coverage area, the minimum UE throughput is increased as compared to the case without new small cells, but still falls below the threshold. The solution obtained by numerically solving (4.34) offers a slightly higher minimum UE throughput than the proposed iterative algorithm, at the cost of a higher computational complexity. The proposed algorithm needs 2×10^4 to 5×10^4 loops to get the optimal solution, while numerically solving (4.34) needs 3.75×10^{11} to 8×10^{12} loops. The performance of the iterative algorithm gets closer to that of the joint optimisation for more HS UEs.

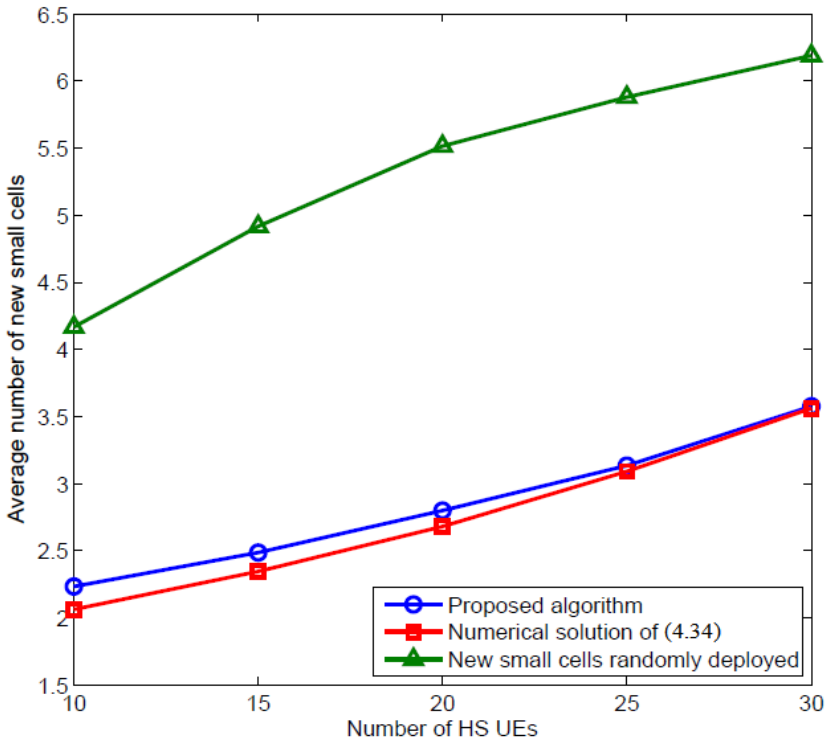


Fig. 4.7 The average number of new small cells vs. the number of HS UEs.

Fig. 4.7 shows the average number of new small cells required for all UEs to meet the UE throughput threshold versus the number of HS UEs. We can see that the proposed iterative algorithm requires slightly more new small cells than the numerical optimisation solution, with the gap between them reducing as the number of HS UEs increases. The random deployment requires a lot more additional small cells than both the proposed algorithm and the numerical optimisation solution.

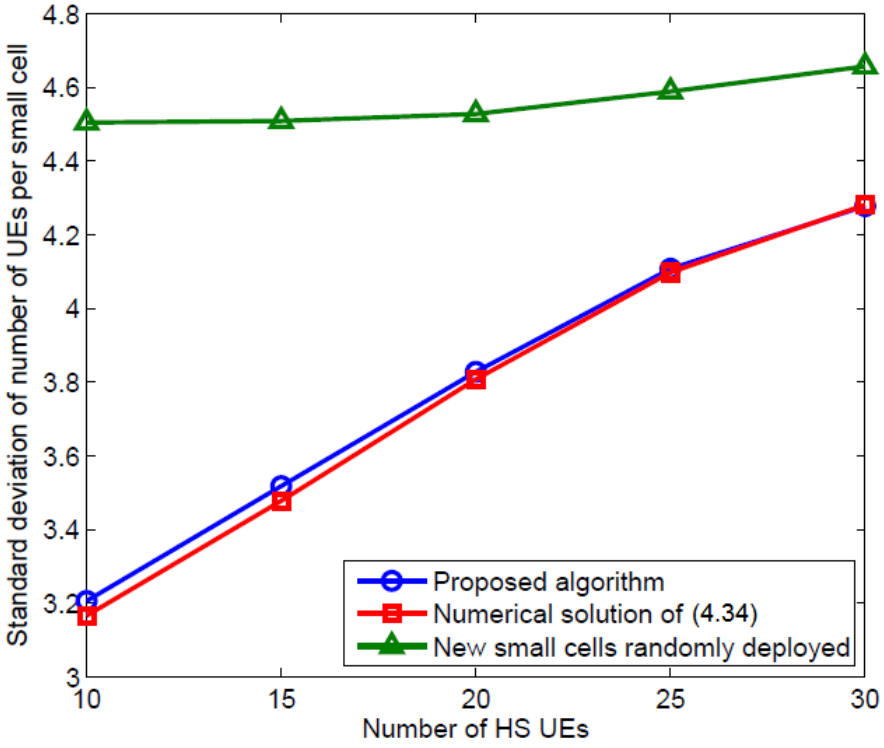


Fig. 4.8 The SD of number of small cell UEs vs. the number of HS UEs.

Fig.4.8 plots the standard deviation (SD) of the number of UEs per small cell under the same setting as Fig. 4.7. We can see that the proposed iterative algorithm achieves a SD of UEs per small cell very close to that of the numerical optimisation solution, which is much lower than that of the random deployment of new small cells. This shows that the proposed algorithm

achieves a more balanced load distribution using less new small cells than the random deployment, due to the optimised locations of new small cells.

4.2.6 Conclusion

In this section, an iterative algorithm is proposed to optimise the number and locations of new small cells to be deployed on top of an existing HetNet and the user associations of all cells, in order to fulfil the extra traffic demands of HS UEs. The simulation results have shown that the iterative algorithm offers a much higher minimum UE throughput and requires less new small cells for all UEs to meet the UE throughput threshold than the random deployment of new small cells. Moreover, the iterative deployment algorithm achieves a more balanced load distribution using less new small cells than the random deployment.

Chapter 5

Small Cell Deployment Based on Stochastic Geometry Analysis

5.1 Introduction

In this chapter, HetNet is analysed before and after an unexpected HS occurs and additional small cells that would be required to mitigate the HS effect under a spatial bivariate PPP model. The spatial distributions of different network nodes (e.g., existing small cells, additional small cells to be deployed, HS UEs, and non-HS UEs) have been modelled into five possible events. By minimizing the difference between the numbers of macrocell UEs after and before the HS occurs based on the analytical results, the optimal numbers of additional small cells required for the HS and non-HS areas are obtained. After, an algorithm is proposed to maximize the average UE throughput in a HetNet affected by unexpected HS by jointly optimizing the locations of additional small cells on top of the existing HetNet and the user associations of all cells.

The relationship between HS UE intensity and additional small cell intensities in HS and non-HS areas is analysed. The simulation results show that our proposed algorithm can effectively maintain the average UE throughput requirement with excellent fairness among all UEs even for a very high intensity of HS UEs.

5.2 System Model

A two-tier HetNet that contains macrocells and small cells is considered as the scenario in this work. In each macrocell coverage area, the location of the macro BS is constant, while the small cell BSs and UEs are considered as random points [63]. Based on the spatial bivariate PPP [52], five events are defined for the problem of deploying additional small cells on top of an existing HetNet as presented in Table 5.1. Event A represents non-HS UEs (that are not jointly distributed with event C, which will be explained later), with the spatial intensity α. Event B denotes the HS UEs (that are not jointly distributed with event E, which will be explained later) in the HS area with the spatial intensity β. Since the HS area overlaps with the macrocell coverage, non-HS UEs may also occur in the HS area. Event C represents the existing small cells. It is assumed that each existing small cell serves at least one non-HS UE and that each existing small cell BS is jointly distributed with a non-HS UE. Hence, event C is considered as an existing small cell and a non-HS UE jointly occurring at a random point y. Event D denotes additional small cells that will be deployed in non-HS areas, with intensity μ. Event E denotes additional small cells deployed in the HS area with intensity ρ, each serving at least one HS UE. Thus, each event E node is jointly distributed with an HS UE.

TABLE 5.1: Events.

Event	Type	Intensity	Given
A	non-HS UEs	α	Yes
B	HS UEs	β	Yes
C	Existing small cells	ν	Yes
D	Additional small cells outside HS	μ	No
E	Additional small cells in HS	ρ	No

After a significant HS occurs and additional small cells are deployed, the overall spatial intensity of small cells in non-HS areas is $\nu+\mu$ and that of small cells in the HS area is $\nu+\rho$ [64].

For event C, denote $f_1(\|\mathrm{x}_u-\mathrm{y}\|)$ and $g_1(\|\mathrm{x}_s-\mathrm{y}\|)$ as the probability density functions (PDFs) of a UE at x_u and an existing small cell at x_s when event C occurs at y, respectively. Zero-mean isotopic Gaussian distributions with standard deviation σ_{f_1} and σ_{g_1} are chosen for $f_1(\mathrm{X})$ and $g_1(\mathrm{X})$, respectively [52].

Before the HS occurs, the probability of event C occurring at any point in the macrocell coverage area is calculated as [52]:

$$\begin{aligned}\mathbf{Pr}\{\mathrm{C}\} &= h_1(\mathrm{x}_u,\mathrm{x}_s)\int_{\mathrm{y}\in R_M} f_1(\|\mathrm{x}_u-\mathrm{y}\|)g_1(\|\mathrm{x}_s-\mathrm{y}\|)d\mathrm{y} \\ &= h_1(\mathrm{x}_u-\mathrm{x}_s,0)=h_1(\mathrm{X}_1,0)=\frac{1}{2\pi\sigma_1^2}e^{\frac{-\|\mathrm{X}_1\|^2}{2\sigma_1^2}}\end{aligned} \tag{5.1}$$

where R_M is the macrocell coverage area, $\mathrm{X}_1=\mathrm{x}_u-\mathrm{x}_s$, and $\sigma_1^2=\sigma_{f_1}^2+\sigma_{g_1}^2$, which is a correlation factor that is determined by the distance of a non-HS UE to its jointly distributed existing small cell BS for event C.

Similarly, for event D, denote $f_2(\| \mathrm{x}_{hsu} - \mathrm{y} \|)$ and $g_2(\| \mathrm{x}_{ns} - \mathrm{y} \|)$ as the PDFs for a HS UE occurring at x_{hsu} and an additional small cell BS deployed at x_{ns} in the HS area, respectively. $f_2(\mathrm{X})$ and $g_2(\mathrm{X})$ are zero-mean isotopic Gaussian distributions with standard deviation σ_{f_2} and σ_{g_2}, respectively. The probability of a joint event between a HS UE at x_{hsu} and an additional small cell BS at x_{ns} occurring in the HS area is given by [52]:

$$\begin{aligned} h_2(\mathrm{x}_{hsu}, \mathrm{x}_{ns}) &= \int_{\mathrm{y} \in R_H} f_2(\| \mathrm{x}_{hsu} - \mathrm{y} \|) g_2(\| \mathrm{x}_{ns} - \mathrm{y} \|) d\mathrm{y} \\ &= h_2(\mathrm{X}_2, 0) = \frac{1}{2\pi\sigma_2^2} e^{\frac{-\|\mathrm{X}_2\|^2}{2\sigma_2^2}} \end{aligned} \tag{5.2}$$

where R_H is the HS area, $\mathrm{X}_2 = \mathrm{x}_{hsu} - \mathrm{x}_{ns}$, and $\sigma_2^2 = \sigma_{f_2}^2 + \sigma_{g_2}^2$.

5.2.1 The Network without HS

By using the given intensities of event A and event C and $h_1(\mathrm{x}_u, \mathrm{x}_s)$ in (5.1), we can calculate the expected number of non-HS UEs in the coverage of an existing small cell BS at $\mathrm{x}_s, \mathrm{x}_s \in R_M$, as follows:

$$\chi_{AC}(r_s) = \int_{\mathrm{x}_u \in C_s} (\alpha + h_1(\mathrm{x}_u, \mathrm{x}_s)) d\mathrm{x}_u = \alpha\pi r_s^2 + (1 - e^{\frac{-r_s^2}{2\sigma_1^2}}) \tag{5.3}$$

where C_s is the coverage area of a small cell, and r_s is the coverage radius of a small cell.

The expected number of existing small cells in the macrocell coverage area is given by:

$$\chi_C(r_m) = \int_{\mathrm{x}_s \in R_M} \nu d\mathrm{x}_s = \nu\pi r_m^2 \tag{5.4}$$

where r_m is macrocell coverage radius.

The expected total number of non-HS UEs in the macrocell coverage area before HS occurs is given by:

$$\begin{aligned} E\{N_U\} &= \int_{\mathrm{x}_u \in R_M} \alpha d\mathrm{x}_u + \int_{\mathrm{x}_u \in R_M} \nu h_1(\mathrm{x}_u, \mathrm{x}_s) d\mathrm{x}_u \\ &= \alpha\pi r_m^2 + (1 - e^{\frac{-r_m^2}{2\sigma_1^2}})\nu\pi r_m^2 \end{aligned} \tag{5.5}$$

5.2.2 The Network with HS

After a HS with UE intensity β occurs in the coverage of a macrocell, by using $h_2(\mathrm{x}_{hsu}, \mathrm{x}_{ns})$ in (5.2), the expected number of UEs (including HS UEs and non-HS UEs) in an additional small cell coverage which is in the HS area can be written as:

$$\begin{aligned} \chi_{ABE}(r_s) &= \int_{\mathrm{x}_{hsu} \in C_s} (\beta + h_2(\mathrm{x}_{hsu}, \mathrm{x}_{ns})) d\mathrm{x}_{hsu} + \int_{\mathrm{x}_u \in C_s} \alpha d\mathrm{x}_u \\ &= (\alpha + \beta)\pi r_s^2 + (1 - e^{\frac{-r_s^2}{2\sigma_2^2}}) \end{aligned} \tag{5.6}$$

where σ_2 is defined in (5.2).

The expected number of UEs (including HS UEs and non-HS UEs) in an existing small cell coverage that overlaps with the HS area is given by:

$$\chi_{AB}(r_s) = \chi_{AC}(r_s) + \int_{\mathrm{x}_{hsu} \in C_s} \beta d\mathrm{x}_{hsu} = (\alpha + \beta)\pi r_s^2 + (1 - e^{\frac{-r_s^2}{2\sigma_1^2}}) \tag{5.7}$$

where C_h is the area of the small cell coverage in the HS area. For simplicity, it is assumed that the existing small cell is completely in the HS area in (5.7).

The expected number of non-HS UEs served by an additional small cells in non-HS areas is given by:

$$\chi_{AD}(r_s) = \int_{x_u \in C_s} \alpha dx_u = \alpha \pi r_s^2 \tag{5.8}$$

The expected number of existing small cells in non-HS areas is:

$$\chi'_C = \int_{x_s \in R_{nH}} \nu dx_s = \nu\pi(r_m^2 - r_{hs}^2) \tag{5.9}$$

where R_{nH} is the non-HS area in the macrocell coverage.

The expected number of additional small cells in the HS area is given by:

$$\chi_E(r_{hs}) = \int_{x_{ns} \in R_{H'}} \nu dx_{ns} = \rho(\pi r_{hs}^2 - \nu\pi^2 r_{hs}^2 r_s^2) \tag{5.10}$$

where r_{hs} is the radius of the HS area, which is assumed to be a disc area, $R_{H'}$ is the HS area excluding the coverage area of existing small cells in the HS. Since the HS area overlaps with the existing network, the coverage areas of existing small cells located in the HS need to be removed from the calculation of additional small cells in the HS area.

The expected number of existing small cells in the HS area is given by $\chi_C(r_{hs})$ following (5.5).

The expected number of additional small cells in non-HS areas is:

$$\chi_D = \int_{x_{ns} \in R_{nH}} \mu dx_{ns} = \mu\pi(r_m^2 - r_{hs}^2) \tag{5.11}$$

The expected total number of UEs in the macrocell coverage area after the HS occurs is given by:

$$E\{N'_U\} = \alpha\pi r_m^2 + \nu\pi r_m^2(1-e^{\frac{-r_m^2}{2\sigma_1^2}}) + \beta\pi r_{hs}^2 + \rho(\pi r_{hs}^2 - \nu\pi^2 r_{hs}^2 r_s^2)(1-e^{\frac{-r_{hs}^2}{2\sigma_2^2}}) \quad (5.12)$$

5.3 Intensity of Additional Small Cells

It is assumed that all the UEs in the HetNet are satisfied with their service before a significant HS occurs. After a significant HS occurs, without deploying additional small cells, the HS UEs that cannot be accommodated by existing small cells will be served by the macrocell. This extra burden on the macro BS may reduce the QoS of macrocell UEs. In order to avoid overloading the macrocell, the object is keeping the number of macrocell UEs after the HS occurs as close as possible to that before the HS occurs.

The expected number of macrocell UEs before the HS occurs is given by:

$$E\{N_{MU}\} = E\{N_U\} - \chi_{AC}(r_s) \cdot \chi_C(r_m) \quad (5.13)$$

After the HS occurs and additional small cells are deployed, the expected number of UEs served by the existing small cells in non-HS areas can be calculated as:

$$E\{N_{nhs}\} = \chi_{AC}(r_s) \cdot \chi'_C \quad (5.14)$$

The expected number of UEs served by the existing small cells in the HS area can be calculated as:

$$E\{N'_{hs}\} = \chi_{AB}(r_s) \cdot \chi_C(r_{hs}) \quad (5.15)$$

The expected number of UEs served by additional small cells in the HS area can be calculated as:

$$E\{N_{hs}\} = \chi_{ABE}(r_s) \cdot \chi_E(r_{hs}) \tag{5.16}$$

The expected number of UEs served by additional small cells in non-HS areas can be calculated as:

$$E\{N'_{nhs}\} = \chi_{AD}(r_s) \cdot \chi_D \tag{5.17}$$

Hence, the expected number of macrocell UEs after the HS occurs can be calculated as:

$$E\{N'_{MU}\} = E\{N'_U\} - (E\{N_{nhs}\} + E\{N_{hs}\} + E\{N'_{hs}\} + E\{N'_{nhs}\}) \tag{5.18}$$

The intensities of additional small cells in the HS and non-HS areas is optimised through minimizing the difference between the numbers of macrocell UEs after and before the HS occurs as follows:

$$\arg\min_{\rho,\mu} | E\{N'_{MU}\} - E\{N_{MU}\} | \tag{5.19}$$

$$\text{s.t. } \rho, \mu \geq 0. \tag{5.20}$$

By using the results in (5.3)-(5.18), the objective function in (5.19) can be rewritten as:

$$\begin{aligned} &\arg\min_{\rho,\mu} | a \cdot \rho + b \cdot \mu + c | \\ &a = [(\alpha + \beta)\pi r_s^2 + (1 - e^{\frac{-r_s^2}{2\sigma_2^2}})](\pi r_{hs}^2 - v\pi^2 r_{hs}^2 r_s^2) \\ &b = \alpha\pi r_s^2 (\pi r_m^2 - \pi r_{hs}^2) \\ &c = (\beta\pi r_s^2 + e^{\frac{-r_s^2}{2\sigma_1^2}} - e^{\frac{-r_s^2}{2\sigma_2^2}})v\pi r_{hs}^2 - \beta\pi r_{hs}^2 \end{aligned} \tag{5.21}$$

TABLE 5.2: System Settings for stochastic geometry analysis.

Parameter	Value
α	$50\text{UE}/km^2$
ν	$10\text{BS}/km^2$
σ_1^2	0.1
σ_2^2	0.1
r_m	500m
r_s	20m
r_{hs}	50m

The optimisation problem in (5.19) can be readily solved using numerical methods. Here numerical solutions of the optimal ρ and μ obtained under the system parameters in Table II is presented. The centre of the HS area is randomly generated in the macrocell coverage area while ensuring the whole HS area is within the macrocell coverage area with a minimum distance from the macrocell BS of 100m.

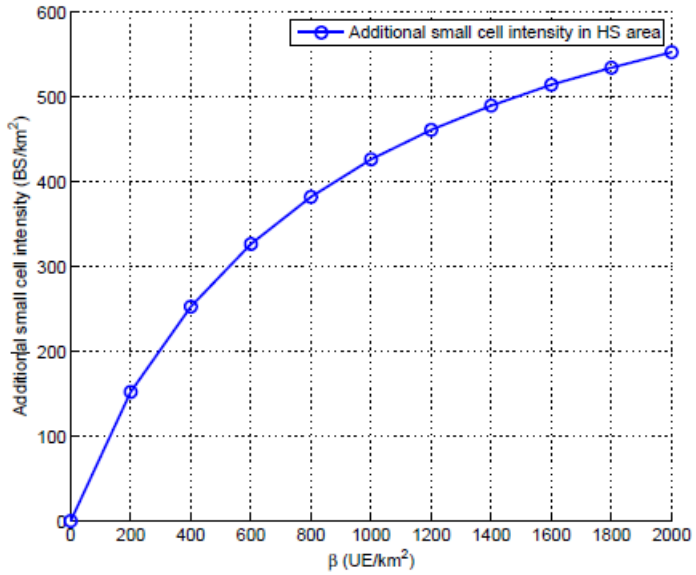


Fig. 5.1 The intensity of additional small cells in the HS area (ρ) versus the intensity of HS UEs.

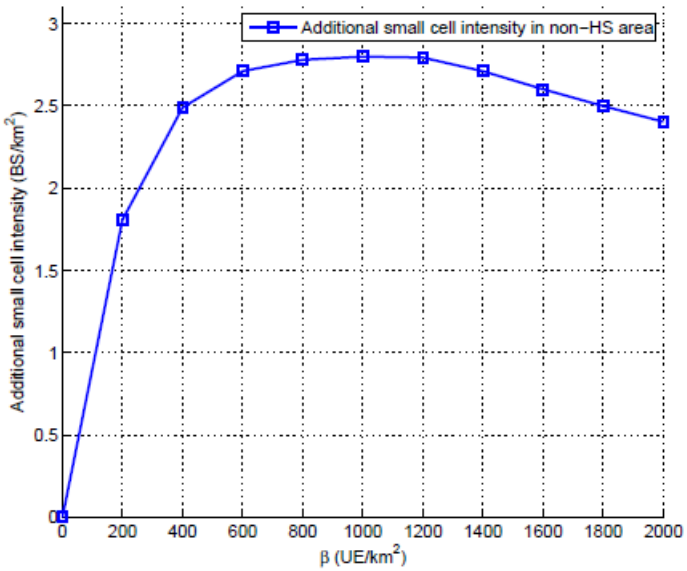


Fig. 5.2 The intensity of additional small cells in non-HS areas (μ) versus the intensity of HS UEs.

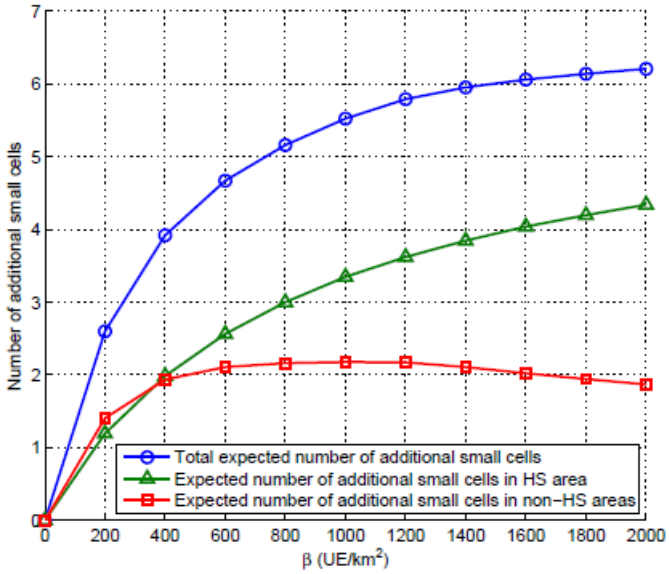


Fig. 5.3 An example of the total number of new small cells under the setting in Table 5.2.

The intensity of additional small cells in the HS area versus the intensity of HS UEs is presented in Fig. 5.1. It can be seen that the intensity of additional small cells in the HS area increases with the intensity of HS UEs at a decreasing rate. Fig. 5.2 shows the intensity of additional small cells in non-HS areas versus the intensity of HS UEs. Comparing Fig. 5.2 with Fig. 5.1, the intensity of additional small cells in the HS area is much higher than that in non-HS areas; the rapid increase of the intensity for additional small cells in non-HS areas when $\beta < 1000\text{UE}/km^2$ is because the intensity of additional small cells in the HS area is not enough to serve all the HS UEs, thus requiring additional small cells in non-HS area to share the extra burden caused by HS UEs to the cellular network. For large values of β, the intensity of additional small cells in non-HS areas decreases with β, because the additional small cells in the HS area are now able to serve the majority of HS UEs. Fig. 5.3 shows the expected numbers of additional small cells in the HS and non-HS areas under the setting in Table 5.2.

5.4 Deployment of Additional Small Cells

Based on the analytical results in Section 5.3, the objective is maximizing the average UE throughput among all UEs by jointly optimizing the locations of new small cells and user associations of all cells for a HetNet affected by a HS not expected in the original network planning.

The DL of a two-tier HetNet consisting of one central macrocell and $\grave{}\chi_C(r_m)$p small cells randomly distributed in the macrocell coverage area is considered in this work. The number of existing BSs is $N_{eBS} = 1 + \grave{}\chi_C(r_m)$p. Each cell has access to the total of N_{RB} RBs. Denoting the number of additional small cells to be deployed in the HS area as N_{ns_1}, and the new small cells to be deployed in non-HS area as N_{ns_2}, the total number of BSs is given by $N_{BS} = N_{eBS} + N_{ns_1} + N_{ns_2}$. Denote the set of BSs as $N_{BS} = \{1, 2, ..., N_{BS}\}$.

According to the system model in Section 5.2, the expected number of HS UEs is given by:

$$\chi_{BE}(r_{hs}) = \beta\pi r_{hs}^2 + (1 - e^{\frac{-r_{hs}^2}{2\sigma_2^2}})\nu\pi r_{hs}^2 \tag{5.22}$$

Let $N_{nhs} = \grave{}E\{N_U\}$p denote the number of non-HS UEs and $N_{hs} = \grave{}\chi_{BE}(r_{hs})$p denote the number of HS UEs in the macrocell coverage area. The set of all UEs is given by $N_U = \{1, 2, ..., N_U\}$, where $N_U = N_{nhs} + N_{hs}$.

The throughput of the i th UE ($i = \{1, 2, ..., N_U\}$) is given by:

$$\gamma_i = \sum_{j=1}^{N_{BS}} N_j^{RB} \cdot W \cdot \log_2(1 + \frac{P_j \cdot g_{i,j} \cdot u_{i,j}}{I_{i,j} + N_0}) \tag{5.23}$$

where $u_{i,j}$ = 1 if the i th UE is served by the j th BS, $u_{i,j} = 0$ otherwise; W is the bandwidth of a RB; without considering power control, P_j is the DL transmit power of the j th BS in a RB, and $P_j = P_{M(S)}$ if it is a macro (small-cell) BS. Assuming that the channel on each RB sees independent and identical Rayleigh fading, the channel power gain of the link between the i th UE and the j th BS in a RB is expressed as (4.2). N_j^{RB} is the number of RBs per UE in cell j and is given by (4.32). As illustrated and proven by Ye *et al.* in [65], equal resource allocation among the UEs in all cells is optimal for the logarithmic utility, it is assumed that all the available RBs are allocated in each cell following the round robin algorithm with full bandwidth allocation [60], so that each cell is fully loaded and there will be inter-cell interference in each RB; N_0 is the additive white Gaussian noise (AWGN) power; and $I_{i,j}$ is the interference power in an RB received by UE i from BSs other than BS j, i.e.,

$$I_{i,j} = \sum_{j'=1, j' \neq j}^{N_{BS}} P_{j'} \cdot g_{i,j'}, \forall j \in N_{BS} \tag{5.24}$$

The additional small cell deployment optimisation problem is formulated as,

$$\arg\max_{\mathbf{U},(\mathbf{x}_1,\mathbf{y}_1),(\mathbf{x}_2,\mathbf{y}_2)} \frac{\sum_{i=1}^{N_U} \gamma_i}{N_U} \tag{5.25}$$

$$\text{s.t. } (\mathbf{x}_1, \mathbf{y}_1) \in |H_{hs}| \tag{5.26}$$

$$(\mathbf{x}_2, \mathbf{y}_2) \in |H_{nhs}| \tag{5.27}$$

$$\sum_{j=1}^{N_{BS}} u_{i,j} = 1, \forall i \in N_U \tag{5.28}$$

$$\gamma_i \geq \gamma_{th}, \forall i \in N_U \tag{5.29}$$

$$N_{RB} \geq N_j^{RB} > 0, \forall j \in N_{BS} \tag{5.30}$$

$$u_{i,j} \in \{0,1\}, \forall i \in N_U, \forall j \in N_{BS} \tag{5.31}$$

where $\mathbf{U}$ is the $N_U \times N_{BS}$ UE-BS association matrix that contains all $u_{i,j}$ as elements; $(\mathbf{x}_1, \mathbf{y}_1)$ are the $N_{ns_1} \times 1$ location vectors of additional small cells in the HS area; $(\mathbf{x}_1, \mathbf{y}_2)$ are the $N_{ns_2} \times 1$ location vectors of additional small cells in non-HS areas; $|H_{hs}|$ and $|H_{nhs}|$ are the feasible deployment areas [59] for additional small cells in the HS and non-HS areas, respectively; and γ_{th} is the minimum UE throughput threshold.

Note that the joint optimisation of the locations of additional small cells and the user associations of all cells in (5.25) is a mixed integer programming problem, which is NP-hard and requires a high computational complexity to find the global optimal solution. Therefore, a practical algorithm in Algorithm 4 is proposed to solve the optimisation problem in (5.25).

Algorithm 4

Initialization:

$u^r_{i,j} \leftarrow 0, \forall i,j$; $\mathbf{X}_1, \mathbf{X}_2 \leftarrow \varnothing$; MAX$\leftarrow$ 0; $N_{ns_1} \leftarrow \lfloor \chi_E(r_{hs}) \rfloor$; $N_{ns_2} \leftarrow \lfloor \chi_D \rfloor$

Main Function:

$|\mathcal{H}_{hs}| \leftarrow |\mathcal{H}|_{\mathbf{X}_1}$
$|\mathcal{H}_{nhs}| \leftarrow |\mathcal{H}|_{\mathbf{X}_2}$
$(g, \mathbf{U}_r, \mathbf{X}_1^*, \mathbf{X}_2^*) \leftarrow$ GRG solve (5.25) for $\mathbf{X}_1, \mathbf{X}_2$
if (5.25) is feasible, $\mathbf{U}_r \in \mathbb{Z}^+$ **then**
    if $g >$MAX **then**
        MAX$\leftarrow g(\mathbf{U}_r, (\mathbf{X}_1, \mathbf{X}_2))$
        $\mathbf{U}^* \leftarrow \mathbf{U}_r$
        $\mathbf{X}_1^* \leftarrow \mathbf{X}_1$
        $\mathbf{X}_2^* \leftarrow \mathbf{X}_2$
        if $N_{ns_1} = \lfloor \chi_E(r_{hs}) \rfloor$ **then**
            $N_{ns_1} \leftarrow \lceil \chi_E(r_{hs}) \rceil$
            return
        else if $N_{ns_1} = \lceil \chi_E(r_{hs}) \rceil, N_{ns_2} = \lfloor \chi_D \rfloor$ **then**
            $N_{ns_2} \leftarrow \lceil \chi_D \rceil$
            return
        else if $N_{ns_1} = \lceil \chi_E(r_{hs}) \rceil, N_{ns_2} = \lceil \chi_D \rceil$ **then**
            return
        end if
    else if $g \leq$MAX **then**
        if $N_{ns_1} = \lceil \chi_E(r_{hs}) \rceil$ **then**
            $N_{ns_1} \leftarrow \lfloor \chi_E(r_{hs}) \rfloor$
            $N_{ns_2} \leftarrow \lceil \chi_D \rceil$
            return
        else if $N_{ns_1} = \lfloor \chi_E(r_{hs}) \rfloor, N_{ns_2} = \lceil \chi_D \rceil$ **then**
            $N_{ns_2} \leftarrow \lfloor \chi_D \rfloor$
            return
        end if
        return
    end if
    return MAX, $\mathbf{U}^*, \mathbf{X}_1^*, \mathbf{X}_2^*$
else if (5.25) is feasible , $\mathbf{U}_r \notin \mathbb{Z}^+$ **then**
    for all $u^r_{i,j} \notin \mathbb{Z}^+$ **do**
        $g(u^r_{i,j} = 0) \rightarrow \mathbf{U}_r, \mathbf{X}_1, \mathbf{X}_2$, MAX
        $g(u^r_{i,j} = 1) \rightarrow \mathbf{U}_r, \mathbf{X}_1, \mathbf{X}_2$, MAX
    end for
else if (5.25) is infeasible, **then**
    return
end if

In Algorithm 4, the binary constraint of user association indicators is relaxed to $0 \le u_{i,j}^{r} \le 1$ [61]. All user association indicators are initialized as 0. Denote $\mathbf{X}_1$ and $\mathbf{X}_2$ as the sets of additional small cells in the HS and non-HS areas, respectively. The two sets are both initialized as empty. The sizes N_{ns_1} and N_{ns_2} of vectors $(\mathbf{x}_1, \mathbf{y}_1)$ and $(\mathbf{x}_2, \mathbf{y}_2)$ respectively are initialized based on the analytical results in Section 5.3. The optimised numbers of additional small cells to be deployed will be refined in the algorithm. The additional small cells' locations are optimised by the generalized reduced gradient (GRG) method [62].

5.5 Simulation Results

In this section, the simulation results are presented to evaluate the performance of the proposed algorithm for deploying additional small cells on top of an existing HetNet affected by HS. In the simulation, in addition to Table II the transmit power of the macro BS and a small-cell BS are 46dBm and 23dBm, respectively, the bandwidth of each RB $W = 180$ kHz, the minimum UE throughput threshold $\gamma_{th} = 3$ Mbps, the number of RBs in each cell $N_{RB} = 50$, and $N_0 = -174$ dBm/Hz.

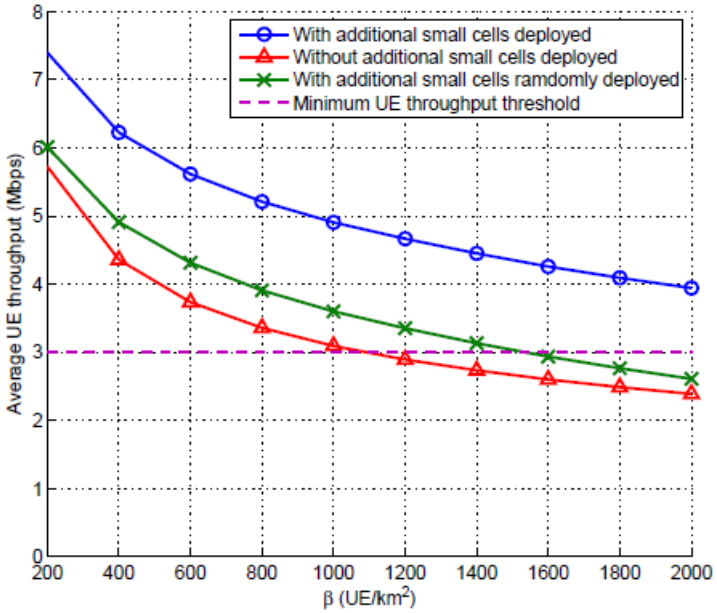


Fig. 5.4 Average UE throughput versus the intensity of HS UEs.

Fig. 5.4 shows the average UE throughput with or without additional small cells deployed following our proposed algorithm and with the same number of additional small cells randomly deployed versus the spatial intensity of HS UEs. We can see that by deploying additional small cells following our proposed algorithm, the average UE throughput is much higher than without deploying any additional small cells or randomly deploying the same number of additional small cells. The average UE throughput can be kept above the threshold even for a very high intensity of HS UEs. If without deploying additional small cells or randomly deploying the additional small cells, the average UE throughput drops below the threshold as the intensity of HS UEs increases, indicating that some of the UEs will have unsatisfactory QoS.

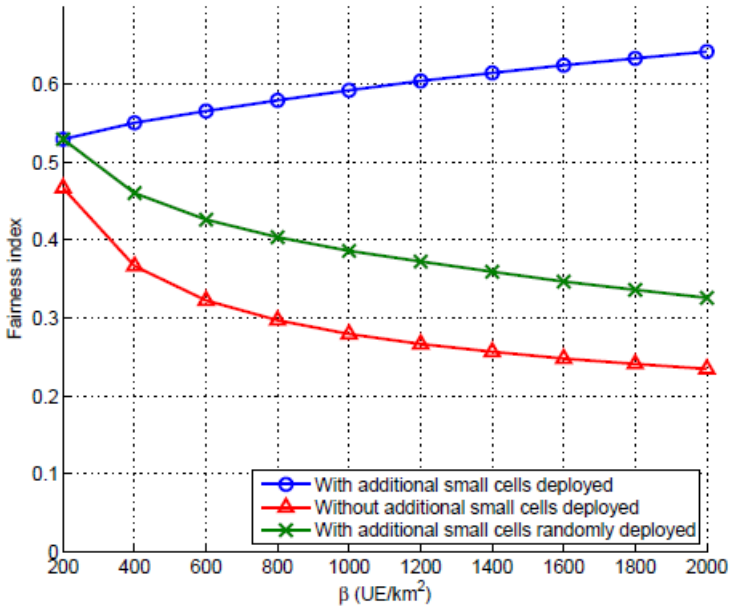


Fig. 5.5 Fairness index versus the intensity of HS UEs.

In order to evaluate the load balancing performance of our algorithm, the fairness index of all UEs is calculated as [66],

$$F = \frac{(\sum_{i=1}^{N_U} \gamma_i)^2}{N_U \cdot \sum_{i=1}^{N_U} \gamma_i^2} \tag{5.32}$$

The fairness index versus the intensity of HS UEs is shown in Fig. 5. Since the number and locations of additional small cells and user associations are optimised, the fairness index achieved by our proposed algorithm is much larger than those without deploying additional small cells or randomly deploying the additional small cells, and even increases with the spatial intensity of HS UEs. The fairness indexes of without deploying additional small cells or randomly deploying additional small cells decreases with the spatial intensity of HS UEs, indicating that not only the QoS of UEs but also the fairness among UEs (load balancing performance) is significantly degraded due to the HS.

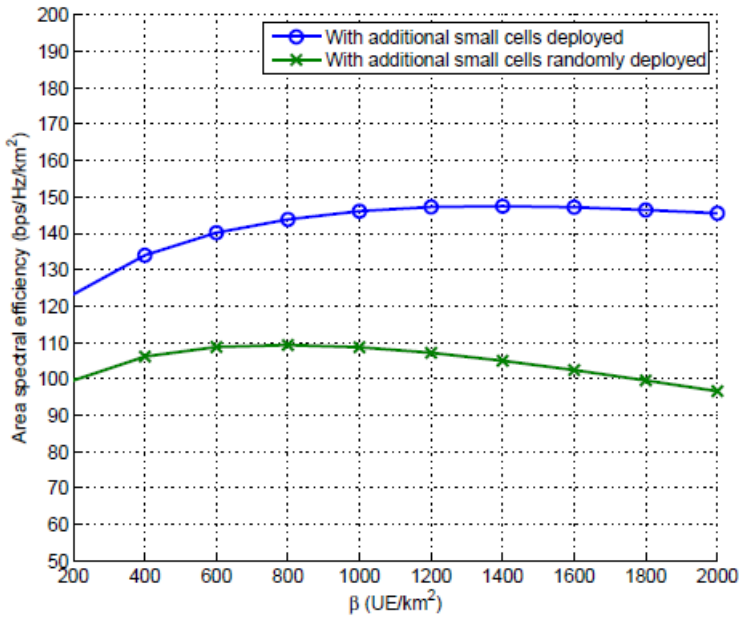


Fig. 5.6 Area spectral efficiency versus the intensity of HS UEs.

Fig. 5.6 shows the Area Spectral Efficiency (ASE) [67] with additional small cells deployed following our proposed algorithm and with the same number of additional small cells randomly deployed versus the spatial intensity of HS UEs. We can see that by deploying additional small cells following our proposed algorithm, the ASE is much higher than randomly deploying the same number of additional small cells, and the gain in ASE increases with the intensity of HS UEs. This is because the numbers of additional small cells deployed in HS and non-HS areas are respectively optimised for given intensity of HS UEs, as shown in Fig. 5.3.

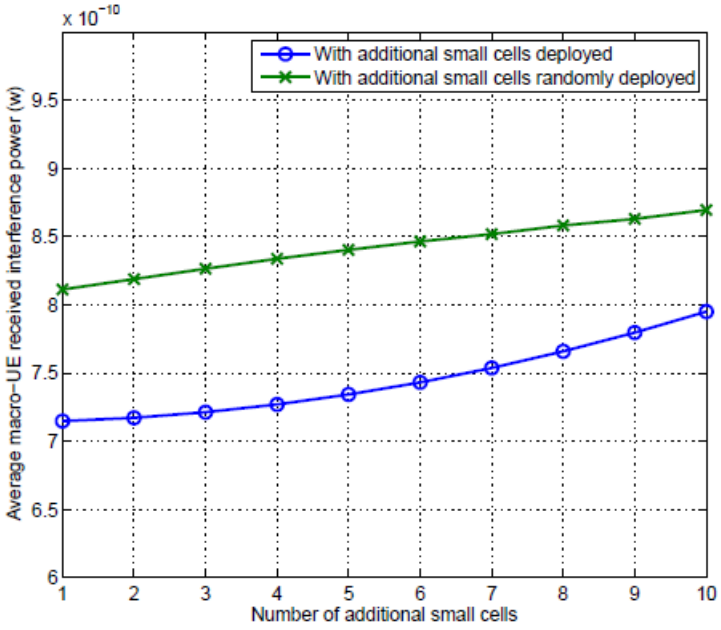


Fig. 5.7 Macro-UE received interference power versus the number of additional small cells.

The average interference power received by a macro UE caused by the additional small cells is simulated to evaluate the performance. The simulation result is presented in Fig. 5.7. We can see that the average macro-UE received interference power caused by additional small cells increases with the number of additional small cells. However, compared with randomly deploying the same number of additional small cells, our proposed algorithm results in much less interference for any given number of additional small cells.

5.6 Conclusion

In this work, the relationship between the intensities of HS UEs (not expected in the original network planning) and additional small cells required to fulfill the extra traffic demand of the HS UEs based on a spatial bivariate PPP model of the HetNet is analysed. Based on the analytical results, the optimal numbers of

additional small cells for the HS and non-HS areas is obtained through an algorithm that proposed in this work to maximize the average UE throughput (after unexpected HS occurs) by jointly optimizing the locations of additional small cells and the user associations of all cells. Simulation results show that the proposed algorithm for optimizing the deployment of additional small cells on top of an existing HetNet can guarantee the QoS requirement of all UEs even for a very high density of HS UEs while achieving excellent load balancing performance for the HetNet.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

To sum up, the overall research work mentioned above can be classified generally into three parts. The first part includes analysing and statistically modelling of the DL throughput per cell distributions over time and over different cells of an existing 3G network based on real network throughput data. The second part work contains the studies of small cell deployment for an existing HetNet in order to solve the extra data requirement from unexpected HSs. The third part includes the work of analysing the relationship between the intensities of HS UEs and additional small cells required through spatial bivariate PPP model.

In the first part of research work, a characteristic that the throughput per cell during office hours of a weekday and a weekend has been found. Hence, the

features of data requirement for an existing 3G network can be acknowledged for later research. Moreover, statistical models of cellular throughput for weekdays and weekends are also constructed, which shows that the instantaneous throughput per cell follows exponential distribution. Accordingly, by giving an exponential PDF parameter λ, the exponential distribution PDFs of instantaneous throughputs per cell at different sampling time instants and the time-varying function can be provided. In order to verify the accuracy of the statistical models in this work, evaluation of the corresponding R-square and RMSE values are provided with the comparison of the model reproduced throughputs with real network measurements.

In the second part of the research work, a study of deployment strategy for additional mobile small cells on top of an existing HetNet has been undertaken. The strategy is proposed for satisfying the extra traffic demand of recurring HSs that have not been considered in the original network planning. The main problem is considered as an optimisation of maximizing the minimum UE throughput jointly over the number and locations of mobile small cells, user associations of all cells, and resource allocation in each cell. By formulating and simplifying the optimisation problem, a joint optimisation of mobile small cell deployment and user association is obtained in this work. A relatively simple algorithm to solve the simplified optimisation problem by giving a fixed feasible number of additional small cell BSs is provided as FNDA. After, FDNA is extended into DOENA algorithm to jointly optimise the number and locations of mobile small cells together with user association. Besides these two algorithms, an iterative algorithm is also studied to optimise the number and locations of new small cells to be deployed on top of an existing HetNet and the user associations of all cells. It has shown that DOENA offers higher minimum UE throughput while requiring less mobile small cells to be deployed, and higher average UE throughput when the number of HS UEs is large in simulation compare to existing small cell deployment strategies. Moreover, for the iterative algorithm, the simulation results have shown that a much higher minimum UE throughput can be obtained with less new small cells required for

all UEs to meet the UE throughput threshold compare to the random deployment of new small cells. And a more balanced load distribution with the iterative deployment algorithm can be achieved by using less new small cells compare to the random deployment.

In the third part of the work, the relationship between the intensities of unexpected HS UEs and additional small cells required is analysed based on a spatial bivariate PPP HetNet model. According to the analytical results, an algorithm is proposed in this work to maximize the average UE throughput by jointly optimizing the locations of additional small cells and the user associations of all cells. The optimal numbers of additional small cells for the HS and non-HS areas is also obtained through this algorithm. By simulation, it shows that the optimised deployment of additional small cells on top of an existing HetNet can guarantee the QoS requirement of all UEs even for a very high density of HS UEs with an excellent load balancing performance.

6.2 Future Work

For the research work in Chapter 3, the exponential distribution PDFs of instantaneous throughputs per cell at different sampling time instants and the time-varying function based on exponential PDF parameter λ has been obtained. As a result, 3G network modelling and simulating can be implemented by using this function because it is close to real network. On the other hand, the locations for each BS in the data set can be found since each BS has a unique series number. Another research of user data traffic based on the features of the living area can be implemented, for instance, the different data traffic of weekday and weekend day for office area and residential area could be found. Moreover, study of mobile network deployment and design could also be undertaken since the user traffic data set is from a real network.

The proposed small cell deployment strategies in Chapter 4 are theoretically used to solve the problem of extra data requirements caused by unexpected HS in an existing HetNet. However, based on the user data traffic statistical results from operators, the strategies can be also used into the whole HetNet design and planning. Moreover, the proposed algorithms are implemented to solving simplified optimisation problems due to the complexities, hence, the future work can focus on extending the proposed algorithms into an algorithm that can be used for small cell deployment, user allocation and resource allocation by extending the resource allocation into a more complicated model. The main difficulties of the work will be the time scales of these three problems being different. Another aspect of the future work based on this research could be focus on energy saving HetNets. According to the simulation results, the proposed small cell deployment algorithms requires less additional small cells to fulfil the extra data requirements. Therefore, be considering the sleeping mode of small cells, the algorithm can be extended to optimise an existing HetNet in order to reduce the power consumption.

The research work in Chapter 5 provides an analysing method of HetNets by using spatial bivariate PPP model. For the whole HetNet, BSs and UEs can be categorized into different events according to their characteristics. As a result, the proposed analysing method can be adjusted to be applied into ultra-dense HetNets for 5G. On the other hand, the theoretical analysing for unexpected HSs in existing HetNets can be also implemented into HetNet design and planning by using different approaches to solve the problem. In the work, optimisation is applied to solve the problem, however, game theory can be also adjusted to solve the problem based on the analysing results.

Bibliography

[1] Tania Teixeira, “Meet marty cooper-the inventor of the mobile phone,” BBC News April 2010.

[2] Cisco Visual Networking Index, “Global mobile data traffic forecast update, 2015-2020, feb. 2016.

[3] 3GPP TR 36.912 V2.0.0, “3GPP; technical specification group radio access network; feasibility study for further advancements for eutra (release 9),” August 2009.

[4] A. Khandekar, N. Bhushan, Ji Tingfang and V. Vanghi, “LTE-advanced: Heterogeneous networks,” in Wireless Conference, 2010, pp. 978–982.

[5] K. Jansson, “First in the world with 4G,” URL: http://www.teliasonerahistory.com/pioneering-the-future/pioneering-the-future/firstin-the-world-with-4g/, 2015.

[6] P. Xia, V. Chandrasekhar and J.G. Andrews, “Open vs. closed access femtocells in the uplink,” IEEE Transactions on Wireless Communications, vol. 9, no. 12, pp. 3798–3809, 2010.

[7] D. López-Pérez, A. Valcarce, G. D. Roche, E. Liu and J. Zhang,” Access methods to wimax femtocells: A downlink system-level case study,” IEEE Singapore International Conference on Communication Systems. 2008, pp. 1657–1662.

[8] H. Y. Hsieh, S. E. Wei and C. P. Chien, “Optimizing small cell deployment in arbitrary wireless networks with minimum service rate constraints,” IEEE Transactions on Mobile Computing, vol. 13, no. 8, pp. 1801–1815, Aug 2014.

[9] M.H. Qutqut, H. Abou-zeid, H.S. Hassanein, A.M. Rashwan and F.M. Al-Turjman, “Dynamic small cell placement strategies for LTE heterogeneous networks,” in IEEE Symposium on Computers and Communication, June 2014, pp. 1–6.

[10] H. Yin and S. Alamouti, “OFDMA: A broadband wireless access technology,” IEEE Sarnoff Symposium, 2006, pp. 1–4.

[11] A.C. Stocker, “Small-cell mobile phone systems,” IEEE Transactions on Vehicular Technology, vol. 33, no. 4, pp. 269–275, 1984.

[12] J.G. Andrews, H. Claussen, M. Dohler, S. Rangan and M.C. Reed, “Femtocells: Past, present, and future,” IEEE Journal on Selected Areas in Communications, vol. 30, no. 3, pp. 497–508, 2012.

[13] 3GPP, “TR 25.820: 3G home nodeB study item technical report,” http://www.3gpp.org/.

[14] V. Chandrasekhar and J.G. Andrews, “Uplink capacity and interference avoidance for two-tier femtocell networks,” IEEE Transactions on Wireless Communications, vol. 8, no. 7, pp. 3498–3509, 2009.

[15] H. Claussen and F. Pivit, “Femtocell coverage optimisation using switched multi-element antennas,” in IEEE International Conference on Communications, 2009, pp. 1–6.

[16] R.M. Radaydeh and M. S. Alouini, “Switched-based interference reduction scheme for open-access overlaid cellular networks,” IEEE Transactions on Wireless Communications, vol. 11, no. 6, pp. 2160–2172, 2012.

[17] Evolved Universal Terrestrial Radio Access, “Further advancements for e-utra physical layer aspects (release 9),” 3GPP TS, vol. 36, pp. V9, 2010.

[18] 3GPP TR 36.921 v11.0.0, “E-UTRA FDD Home eNode B Radio Frequency requirements analysis (Release 11),” October 2012.

[19] J.Y. Wu, Xiaoli Chu and D. Lopez-Perez, “Downlink outage probability of co-channel femtocells in hierarchical 3-sector macrocells,” IEEE Communications Letters, vol. 16, no. 5, pp. 698–701, 2012.

[20] B. W. Silverman, Density estimation for statistics and data analysis, Chapman & Hall/CRC, vol. 26, 1986.

[21] A. Zielesny, From Curve Fitting to Machine Learning, Springer, vol. 18, 2011.

[22] J. R. Taylor, An introduction to error analysis: the study of uncertainties in physical measurements, University science books, 1997.

[23] R. J. Hyndman and A. B. Koehler, “Another look at measures of forecast accuracy,” International Journal of Forecasting, vol. 22, no. 4, pp. 679–688, 2006.

[24] G. McLachlan and D. Peel, Finite mixture models, Wiley-Interscience, 2004.

[25] D. Willkomm, S. Machiraju, J. Bolot and A. Wolisz, “Primary user behavior in cellular networks and implications for dynamic spectrum access,” IEEE Communications Magazine, vol. 47, no. 3, pp. 88–95, 2009.

[26] D. Willkomm, S. Machiraju, J. Bolot and A. Wolisz, “Primary users in cellular networks: A large-scale measurement study,” 3rd IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks, pp. 1–11, 2008.

[27] J. S. Evans and D. Everitt, “On the teletraffic capacity of CDMA cellular networks,” IEEE Transactions on Vehicular Technology, vol. 48, no. 1, pp. 153–165, 1999.

[28] M. Laner, P. Svoboda, S. Schwarz and M. Rupp, “Users in cells: a data traffic analysis,” IEEE Wireless Communications and Networking Conference, pp. 3063–3068, 2012.

[29] U. Paul, A. P. Subramanian, M. M. Buddhikot and S. R. Das, “Understanding traffic dynamics in cellular data networks,” Proceedings IEEE INFOCOM, pp. 882–890, 2011.

[30] U. Paul, A. P. Subramanian, M. M. Buddhikot and S. R. Das, “Understanding spatial relationships in resource usage in cellular data networks,” Conference on Computer Communications Workshops, 2012, pp. 244–249.

[31] M. Z. Shafiq, L. Ji, A. X. Liu, J. Pang and J. Wang, “Characterizing geospatial dynamics of application usage in a 3g cellular data network,” INFOCOM 2012, 2012, pp. 1341–1349.

[32] A. Dong, X. Luo and Q. Du, “A small cell deployment strategy towards amorphous coverage in the cellular network,” International Wireless Communications and Mobile Computing Conference, 2015, pp. 745–750.

[33] G. J. Yu and K. Y. Yeh, “A k-means based small cell deployment algorithm for wireless access networks,” IEEE International Conference on Networking and Network Applications, 2016, pp. 393–398.

[34] W. Chen, H. Li, Z. Li, Z. Xiao and D. Wang, “Optimisation of small cell deployment in heterogeneous wireless networks,” IEEE Information and Telecommunication Systems, 2016, pp. 1–5.

[35] Y. Jung, H. Kim, S. Lee, D. Hong and J. Lim, “Deployment of small cells with biased density in heterogeneous networks,” IEEE Asia-Pacific Conference on Communications, 2016, pp. 541–544.

[36] S. F. Chou, Y. J. Yu and A. C. Pang, “Mobile small cell deployment for service time maximization over next generation cellular networks,” IEEE Transactions on Vehicular Technology, no. 99, pp. 1–1, 2016.

[37] G. Qiao, S. Leng, K. Zhang and K. Yang, “Joint deployment and mobility management of energy harvesting small cells in heterogeneous networks,” IEEE Access, no. 99, pp. 1–1, 2016.

[38] D. O. Abonyi and J. M. Rigelsford, “Localisation system for network planning in 2-tier heterogeneous networks,” Progress in Electromagnetic Research Symposium, 2016, pp. 494–494.

[39] D. K. Shin, W. Choi and T. Yu, “Statistically controlled opportunistic resource block sharing for femto cell networks,” Journal of Communications and Networks, vol. 15, no. 5, pp. 469–475, Oct 2013.

[40] D. Fooladivanda and C. Rosenberg, “Joint resource allocation and user association for heterogeneous wireless cellular networks,” IEEE Transactions on Wireless Communications, vol. 12, no. 1, pp. 248–257, January 2013.

[41] W. Zhao, S. Wang, C. Wang and X. Wu, “Cell planning for heterogeneous networks: An approximation algorithm,” IEEE INFOCOM 2014, April 2014, pp. 1087–1095.

[42] R. Razavi and H. Claussen, “Urban small cell deployments: Impact on the network energy consumption,” IEEE Wireless Communications and Networking Conference Workshops, April 2012, pp. 47–52.

[43] I. Siomina and D. Yuan, “Optimisation approaches for planning small cell locations in load-coupled heterogeneous LTE networks,” IEEE International Symposium on Personal Indoor and Mobile Radio Communications, Sep 2013, pp. 2904–2908.

[44] Y. Park, J. Heo, H. Kim, H. Wang, S. Choi, T. Yu and D. Hong, “Effective small cell deployment with interference and traffic consideration,” IEEE VTC Fall, Sept 2014, pp. 1–5.

[45] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse and M. Franceschetti, “Stochastic geometry and random graphs for the analysis and design of wireless networks,” IEEE Journal on Selected Areas in Communications, vol. 27, no. 7, pp. 1029–1046, 2009.

[46] K. Huang and J. G. Andrews, “An analytical framework for multicell cooperation via stochastic geometry and large deviations,” IEEE Transactions on Information Theory, vol. 59, no. 4, pp. 2501–2516, 2013.

[47] A. K. Gupta, X. Zhang and J. G. Andrews, “Sinr and throughput scaling in ultradense urban cellular networks,” IEEE Wireless Communications Letters, vol. 4, no. 6, pp. 605–608, 2015.

[48] J.G. Andrews, F. Baccelli and R.K. Ganti, “A tractable approach to coverage and rate in cellular networks,” IEEE Transactions on Communications, vol. 59, no. 11, pp. 3122–3134, 2011.

[49] T. Bai and R. W. Heath, “Coverage and rate analysis for millimeter-wave cellular networks,” IEEE Transactions on Wireless Communications, vol. 14, no. 2, pp. 1100–1114, 2015.

[50] E. Mugume, D. K. C. So and E. Alsusa, “Energy efficient deployment of dense heterogeneous cellular networks,” IEEE Global Communications Conference, 2015, pp. 1–6.

[51] J. Peng, P. Hong and K. Xue, “Energy-aware cellular deployment strategy under coverage performance constraints,” IEEE Transactions on Wireless Communications, vol. 14, no. 1, pp. 69–80, 2015.

[52] A. Babaei and B. Jabbari, “Distance distribution of bivariate Poisson network nodes,” IEEE Communications Letters, vol. 14, no. 9, pp. 848–850, September 2010.

[53] C. C. Lin, K. H. S. Liu and S. L. Su, “Coverage performance of heterogeneous cellular networks with dependent cell deployment,” IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, 2016, pp. 1–4.

[54] C. T. Peng, L. C. Wang and C. H. Liu, “Optimal base station deployment for small cell networks with energy-efficient power control,” IEEE International Conference on Communications, 2015, pp. 1863–1868.

[55] J. Zhang, W. Wang, X. Zhang, Y. Huang, Z. Su and Z. Liu, “Base stations from current mobile cellular networks: Measurement, spatial modelling and analysis,” WCNCW, 2013, pp. 1–5.

[56] J. Zhang and H. Tian, “An isotropic SPPP model for femtocells networks with outage probability constraints,” Chinese Journal of Electronics, vol. 24, no. 4, pp. 824–831, 2015.

[57] M. Tahani and M. Sabaei, “A distributed data-centric storage method for hot spot mitigation in wireless sensor networks,” International Symposium on Telecommunications, 2010, pp. 401–408.

[58] W. Zuo, H. Xia and C. Feng, “A novel coordinated multi-point transmission in dense small cell deployment,” in PIMRC 2015, 2015, pp. 1872–1877.

[59] J. Wu, X. Chu, D. Lopez-Perez and H. Wang, “Femtocell exclusion regions in hierarchical 3-sector macrocells for co-channel deployments,” in IEEE International Conference on Communications in China, Aug 2012, pp. 541–545.

[60] 3GPP TR 36.814 V9.0.0, “Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects (Release 9),” March 2010.

[61] J. M. Ortega and W. C. Rheinboldt, Iterative solution of nonlinear equations in several variables, vol. 30, Siam, 1970.

[62] L.S. Lasdon, A.D. Waren, A. Jain and M. Ratner, “Design and testing of a generalized reduced gradient code for nonlinear programming,” ACM Transactions on Mathematical Software (TOMS), vol. 4, no. 1, pp. 34–50, 1978.

[63] T.C. Brown, B.W. Silverman and R.K. Milne, “A class of two-type point processes,” Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, vol. 58, no. 3, pp. 299–308, 1981.

[64] J.E. Paloheimo, “A spatial bivariate Poisson distribution,” Biometrika, vol. 59, no. 2, pp. 489–492, 1972.

[65] Q. Ye, B. Rong, Y. Chen, M. Al-Shalash, C. Caramanis and J.G. Andrews, “User association for load balancing in heterogeneous cellular networks,” IEEE Transactions on Wireless Communications, vol. 12, no. 6, pp. 2706–2716, 2013.

[66] J. Raj, D. Chiu and W. Hawe, “A quantitative measure of fairness and discrimination for resource allocation in shared computer systems,” CoRR, vol. cs.NI/9809099, 1998.

[67] M.S. Alouini and A.J. Goldsmith, “Area spectral efficiency of cellular mobile radio systems,” IEEE Transactions on Vehicular Technology, vol. 48, no. 4, pp. 1047–1066, Jul 1999.