UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

DYNAMIC USER-CENTRIC CAPACITY MAXIMIZATION BY OPTIMIZING ANTENNA PARAMETERS

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE

By

MD SALIK PARWEZ
Norman, Oklahoma
2016

DYNAMIC USER-CENTRIC CAPACITY MAXIMIZATION BY OPTIMIZING ANTENNA PARAMETERS

A THESIS APPROVED FOR THE
SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

BY

Dr. Ali Imran, Chair

Dr. Pramode Verma

Dr. Kam Wai Clifford Chan

Dedicated to Amma and Abba

Acknowledgements

First of all, I would like to thank almighty GOD, the Supreme, for strengthening and supporting me through all phases and through every means.

I would like to express my sincere gratitude to my thesis committee for invaluable support and guidance throughout the research work that helped me to complete this research successfully.

I am extremely grateful to my advisor, Dr. Imran, for his untiring support and research guidance that helped me to accomplish this research work.

To Dr. Verma, I will always be grateful for his wonderful support especially in my difficult time. Without his support, it would have been difficult to carry out this task till this end. Thank you again for being there.

To Dr. Chan, thank you very much for the important questions that you put forth and valuable comments in my thesis. Indeed, your feedbacks were helpful in giving my thesis a good shape.

Last but not the least, I am thankful to Amal and Hasan for their company on different occasions.

Table of Contents

List of Tables

List of Figures

Abstract

Cellular network has become the primary means of voice as well as data communication. With sophisticated but affordable end user devices e.g. smartphones, tablet PCs etc. and ubiquity of mobile connectivity, users are able to access a range of multimedia services requiring low to high data rate and with desired quality of experience everywhere and all the times. However, mobile network operators (MNOs) always have limited bandwidth resources as compared to users' demand, as bandwidth is the most expensive resource in the network. Thus MNOs always seek new tools and technologies to optimally utilize the available bandwidth to accommodate maximum number of users and provide high quality of services, maximizing the revenue in return. Especially, in the case of ultra-dense heterogeneous deployment of small cells equipped with *massive*-MIMO antenna configuration operating over mmWave spectrum in 5G, automated solution for dynamic spectrum optimization with respect to rapidly changing users and network requirement will be of critical importance.

This thesis presents a novel scheme for spectral efficiency (SE) optimization through clustering of users. By clustering users with respect to their geographical concentration we propose a solution for dynamic steering of antenna beam by dynamically adjusting antenna azimuth and tilt angles with respect to the most focal point in every cell that would maximize overall SE in the system. The proposed framework thus introduces the notion of elastic cells that can be potential component of 5G networks. The proposed scheme decomposes large-scale system-wide optimization problem into small-scale local sub-problems and thus provides a low complexity

solution for dynamic system wide optimization. Every sub-problem involves clustering of users to determine focal point of the cell for given user distribution in time and space, and determining new values of azimuth and tilt that would optimize the overall system SE performance. To this end, we proposed three user clustering algorithms to transform a given user distribution into the focal points that can be used in optimization process: the first is based on received signal to interference ratio (SIR) at the user; the second is based on received signal level (RSL) at the user; the third and final one is based on relative distances of users from the base stations. We also formulate and solve an optimization problem to determine optimal radii of clusters. The performances of proposed algorithms and framework are evaluated through system level simulations. Performance comparison against benchmark where no elastic cell deployed, shows that a gain in spectral efficiency of up to 26% is achievable depending upon user distribution in each cell.

Chapter 1: Introduction

1.1 Background & Scope

Mobile communication network is under rapid adaptation as a common medium for almost every type of communication in our daily life. It will proliferate mobile connectivity, connecting 5.5 billion users (70% of the global population), and 12 billion devices by 2020, out of them 26.4% will be machine to machine (M2M) modules enabling internet of things (IoT) [1]. In addition, the penetration of social media has contributed from another dimension in this revolution of communication, making social media the second largest contributor to the global data traffic (15% of the global traffic volume) [2]. The data rate (bit per second) requirement of these tremendously high number of devices vary from few bps to tens of Mbps that further vary spatio-temporally. Moreover, diverse applications and technologies are emerging consistently that put forth additional set of requirements in terms of data rate, signaling, delay etc. However, the diversity of users and their spatio-temporally varying requirements mandate the future networks to be not only heterogeneous and dense but also highly elastic. High network node density further increases the complexity to manage them. Hence, manual optimization of the network becomes highly challenging [3]. Self-Organizing Networks (SON) has emerged as a technique to replace the manual handling by embedding intelligence and elasticity into the network [4]. SON enables the network to adapt to the changing environment by adjusting the network parameters autonomously. SON not only makes network highly efficient but also yields significant reduction in the network operational expenses (OPEX).

1.2 Thesis Statement, Approach and Contribution

Spectrum, which is regarded as one of the scarcest resources, must be efficiently utilized to meet these demands alongside the innovation and invention of new technologies and architectures. On one hand, there are a number of schemes being researched including, amongst others, *massive*- multiple input multiple output (MIMO), base station (BS) densification, mmWave networks, control and data plane separated architecture (CDSA), information centric networks (ICN), device to device communication (D2D) etc. that target the 5G and beyond networks to improve overall network efficiency. This thesis, on the other hand, proposes to improve network spectral efficiency by optimizing the existing network parameters such as antenna azimuth and tilt angles, within the available resources. Based on this idea, this thesis attempts the following:

Thesis Statement: *It is possible to achieve significant gain in system-wide spectral efficiency by steering the antenna beam towards the densest region of the cell without compromising sparse regions.*

Proposed Approach: In this thesis, we propose a framework for self-optimization of system-wide spectral efficiency (SE) through dynamic clustering of users and steering the beam adaptively towards the clustered users. Spectral efficiency is measured as bits/s/Hz which is measured as bits/s/Hz, i.e. data rate that can be achieved over every Hz of frequency. To determine the highly dense regions of users within a cell, we propose and investigate three clustering algorithms, which when implemented, determine focal points in each cell. The clustering algorithms are based upon three different KPIs: signal to interference ratio (SIR), received signal level (RSL) and relative distance of the user from the BS. The optimal radius of the cluster further depends upon factors such as cell size, user distribution etc. We considered uniform user distribution with cluster radius equivalent to small cell radius of 150 meters. Once the clusters and their focal points are found, SE optimization algorithm is utilized to calculate new optimal azimuth and tilt values of antennae in each cell. Such online dynamic beam steering in real network could potentially be exploited using e.g., multi element antenna systems such as MIMO or massive MIMO which are being considered for emerging networks. The kind of beam steering proposed in this thesis is much simpler and easier to implement as it does not require 1) tracking of individual users, 2) estimation of angle of arrival 3) estimation of channel. Instead the proposed solution requires simple antenna adjustments to change its azimuth and/or tilt by a few degrees as we will explain later.

Contribution: From above discussion, the contribution of this thesis can be summarized into following points.

- We proposed self-optimization framework for system-wide spectral efficiency by changing antenna azimuth and tilt
- We proposed and compared three clustering algorithms based upon SIR, RSL and relative distance of users from BS
- We proposed algorithms for calculation of optimal radii of the clusters

The significance of this work lies in the fact that joint optimization of both azimuth and tilt effectively paves the way for newly conceived cell-less deployment architecture—an architecture where cells won't have rigid foot prints. In such architecture active cells' shapes, sizes and numbers will vary with user distribution and demand [3]. Such elastic cell-less architecture is one of the key features being envisioned for 5G. Proposed framework can be implemented in such elastic cell-less architecture by harnessing the beam steering capabilities of multi antenna element systems which are also a key component of 5G landscape.

1.3 Thesis Outline

The rest of the thesis is organized in four chapters. Chapter 2 discusses the relevant works in the literature focused on spectral efficiency optimization. In Chapter 3, we discuss and detail the proposed framework under which we discuss problem formulation of the considered system model. We detail the proposed clustering algorithms and the notion of triplet for optimization in the same chapter. Some of the major and widely considered emerging techniques and evolving approaches for achieving significantly higher spectral efficiency in 5G are discussed in Chapter 4. There we mainly discuss big data analytics and caching at the edge approaches for optimal spectrum utilization in 5G. In Chapter 5, we finally conclude the thesis with a guide to future work.

Chapter 2: Relevant Work

There are various techniques and approaches undertaken in literature to enhance and optimize spectral efficiency for given bandwidth resource. Some of the major examples techniques are antenna parameters adjustments e.g. transmit power, antenna azimuth and tilt angles, antenna height; interference mitigation technique; inter-cell interference coordination techniques; cooperative MIMO techniques etc. A number of them mainly focus on enhancement through technique and algorithms in physical layer (PHY), whereas only few focus on network level system-wide optimization. Although the later one is difficult to achieve, it usually results in significant gain. Without going into each technique's details, we would like to discuss the relevant works on spectral efficiency optimization using antenna parameters.

Antenna parameters has been widely considered for spectral efficiency optimization. In [5], user's average throughput has been maximized using BS-coordinated tilting. The proposed user-centric framework adjusts the antenna tilt of each of BS facing a user jointly to maximum sum throughput of a user. However, it focuses on a single user and it does not incorporate antenna azimuth angle into optimization process. Authors in [6] propose to adaptively adjust antenna tilt and pilot power to meet varying traffic load in a WCDMA system. However, they also consider only the tilt and do not consider azimuth optimization. A similar solution is also proposed in [7] to study capacity and coverage optimization (CCO) use case of SON. However, solution proposed in [7] optimizes the throughput of a single hotspot, it neither considers system-wide optimization nor does it address the dynamically changing user density throughout the cell. In [8] – [11], switched beam tilting has been proposed, in which each BS utilizes one of the many pre-determined fixed tilts to maximize the users' throughput in certain region within a cell. In [11], a framework has been proposed for dividing the cell into co-centric region and apply switched beam tilting. However, solution in [11] is studied in context of an isolated cell, and does not take into account interference from neighboring cells.

In a recent work [12], authors proposed SON enabled system-wide SE optimization solutions for network with hotspots and relay stations. However, work in [12] considers only tilt angle as the optimization parameter. Whereas in this article, we propose self-optimization of both azimuth and tilt angle for dynamically changing user density. Second distinction of this work from [12] is that we present and compare three different user clustering algorithms to determine best representative point in a cell that can be used in the joint azimuth and tilt optimization processes, whereas the work in [12] does not take into account clustering of users. Furthermore, in this work we also present a method to determine optimal radius to cluster the users into groups. These contributions allow the optimization framework to be more user centric and elastic than that presented in [12].

In all previous works, either antenna tile or azimuth angle has been considered for optimization purposes. In addition, previous works either focus on single user, or single cell without taking the interference from neighboring cells into account. Also the previous works are not concerned with user distribution. To the best of authors' knowledge, this is first framework that considers both antenna azimuth as well as tilt simultaneously considering clustering of users for system-wide optimization. This framework achieves its goal without sacrificing other resources and at without increase in signaling load but with significant decrease in computational complexity as compared to optimization with fixed azimuth and tilt.

Chapter 3: User-centric Capacity Optimization using Antenna Azimuth and Tilt

In this chapter, we discuss dynamic user-centric capacity maximization using self-optimization framework. The framework relies on the concept of triplet that assumes that for any BS, the immediate neighbors are the most interfering BSs and thus interference from other BSs can be discarded while optimizing capacity. The framework divides the whole system into small sub-systems of triplet for which capacity can be optimized individually. Not only limited to that, we propose clustering algorithms based upon SIR, RSL and distance parameters, to cluster users and optimize antenna parameter for each triplet with respect to focal point (centroid) of the clusters in each cell. Each centroid effectively represents the group of users clustered together, thus antenna parameters are optimized with respect to only the focal points as against all the users in the case of without clustering. This largely reduces the size of the problem and eases the optimization process. To calculate the optimal radius of cluster, we derive relation based upon SIR received at the user and density of user in the cell.

3.1 System Model

We consider downlink transmission in hexagonal multicellular network in which each BS has three sectors, each covering a span of 120 degrees, as shown in Fig.1(a). Figure 1(b) shows the antenna sectors and antenna azimuth and tilt angle of individual tri-sectored BS. Users are uniformly distributed in each of the cell in the network. It is assumed that all user equipment (UEs) are outfitted with omnidirectional antenna with 0 dB gain. We use spectral efficiency (SE) in b/s/Hz as the optimization metric and we define it as the long term average bandwidth normalized throughput on a link given by $\log_2(1 + SIR)$, where SIR stands for Signal to Interference Ratio.

Let N denotes the set of points corresponding to transmission antenna location of all sectors and K denotes the set of points representing the location of users in the system. The geometrical SIR perceived at a user location k being served by n^{th} sector can be given as:

$$\eta_k^n = \frac{P^n G_k^n \alpha (d_k^n)^{-\beta} \, \sigma_k^n}{\sum_{\forall m \in N \setminus n} (P^m G_k^m \alpha (d_k^m)^{-\beta} \, \sigma)}; m,\ n\ \in N,\ k \in K \qquad (1)$$

where P^n indicates transmission power of the n^{th}cell, d_k^n is the distance between transmitting antenna location n and UE location k. α and β are pathloss coefficient and exponents, respectively. σ_k^n represents shadowing experienced by users at location k while receiving signal from n^{th} transmitting antenna. G_k^n represents antenna gain perceived at k user location from n^{th}antenna.

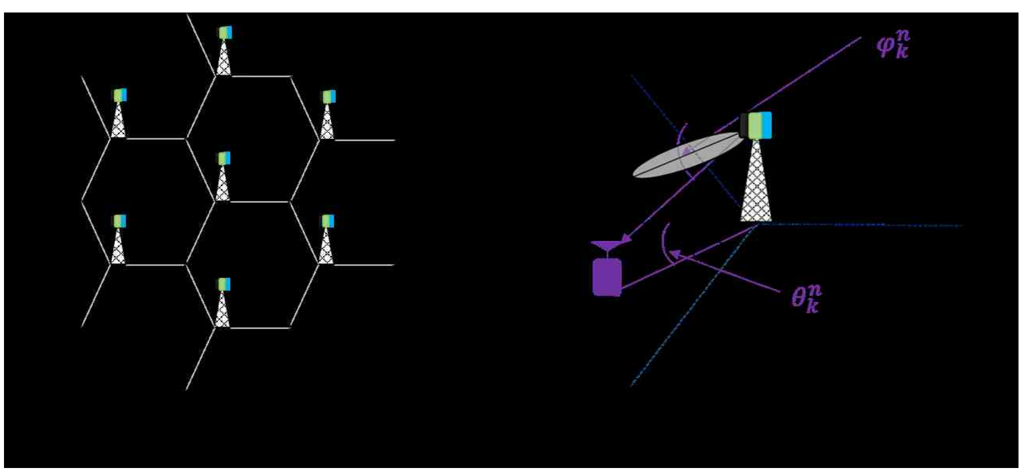

Figure 1: Multi-cell network with antenna parameters at each BS

As proposed by 3GPP [13], the three dimensional antenna pattern can be given as

$$G_k^n = 10^{0.1\left(\begin{array}{c}\lambda_v\left(G_{max}-\min\left(12\left(\frac{\varphi_k^n-\varphi_{tilt}^n}{B_v}\right)^2, A_{max}\right)\right)+\\ \lambda_h\left(G_{max}-\min\left(12\left(\frac{\theta_k^n-\theta_a^n}{B_h}\right)^2, A_{max}\right)\right)\end{array}\right)} \qquad (2)$$

and with simplification introduced in [13], the above expression is reduced to

$$G_k^n = 10^{-1.2\left(\lambda_v\left(\frac{\varphi_k^n-\varphi_{tilt}^n}{B_v}\right)^2+\lambda_h\left(\frac{\theta_k^n-\theta_a^n}{B_h}\right)^2\right)} \qquad (3)$$

where φ_k^n is the vertical angle at the n^{th} BS in degrees from reference axis to the k^{th} UE. φ_{tilt}^n is the tilt angle of the n^{th} cell as shown in Fig.1. Also, θ_a^n represents azimuth angle orientation with respect to horizontal reference axis and θ_k^n is the angular distance of the k^{th} user from horizontal reference axis. For simplicity, we use substitution in (1) as follows:

$$\delta_k^n = \sigma_k^n\alpha(d_k^n)^{-\beta}\ ,\ \delta_k^m = \sigma_k^m\alpha(d_k^m)^{-\beta}\ \text{ and }\ \mu = \frac{-1.2\lambda_v}{{B_v}^2}$$

Using the above substitution and the gain from (3) into (1), we obtain the SIR at the UE which is represented as

$$\eta_k^n = \frac{\delta_k^n 10^{-1.2\left(\lambda_v\left(\frac{\varphi_k^n-\varphi_{tilt}^n}{B_v}\right)^2+\lambda_h\left(\frac{\theta_k^n-\theta_a^n}{B_h}\right)^{2^2}\right)}}{\sum_{\forall m\in N\backslash n}\left(\delta_k^m 10^{-1.2\left(\lambda_v\left(\frac{\varphi_k^m-\varphi_{tilt}^m}{B_v}\right)^2+\lambda_h\left(\frac{\theta_k^m-\theta_a^m}{B_h}\right)^2\right)}\right)} \qquad (4)$$

where, $n, m\ \in N$; the indexes n, m and N represents the serving BS, interfering BS and the set of all the BSs respectively.

3.2 Self-Optimization Framework

The self-optimization framework includes problem formulation and distributed solution to the problem to achieve self-optimization of system throughput. Distributed solution is achieved by dividing the whole system into individual triplets and representing a number of users by a single focal point determined through clustering based upon SIR, RSL and distance. The self-optimization framework determined optimal radius of cluster through relation derived at the end of this section.

3.2.1 Problem Formulation

As assumed, set K represents the location of all the users in the system. So, utilizing Shannon's capacity relation, the bandwidth normalized system throughput maximization can be expressed as (5) below.

$$\max_{\varphi_{tilt}^{N},\theta_{a}^{N}} \hat{\xi} = \max_{\varphi_{tilt}^{N},\theta_{a}^{N}} \sum_{\forall k \in K} log_2\left(1 + \eta_k^n\left((\varphi_{tilt}^{N},\theta_{a}^{N})\right)\right) \qquad (5)$$

It can be easily noticed that η_k^n is a function of system wide azimuth and tilt angle and is calculated using (4). Since K represents user and is usually a large number, (5) results in a large-scale non-linear problem and becomes NP-Hard to solve.

To overcome the difficulty of solving a large scale optimization problem that further require real-time locations of all users in the system, we exploit the concept of determining a single focal point in each cell. The key attribute of such point is that the single point can effectively represent all the users in that cell during the optimization process. The validity of this approach was demonstrated in Lemma 1 in [12] in context of tilt optimization. In this paper, we extend that approach for joint azimuth and tilt approach and propose, in next section, three alternative algorithms to heuristically compute such single point for each cell.

Let S denotes the set of all such focal points in all the cells. Then the objective function in (5) can be optimized only with respect to those focal points and thus can be approximated as

$$\max_{\varphi_{tilt}^{N}, \theta_{a}^{N}} \sum_{\forall s \in S} log_2 \left(1 + \eta_s^n \left((\varphi_{tilt}^{N}, \theta_{a}^{N}) \right)\right) \qquad (6)$$

Thus, compared to problem (5) which optimizes the antenna azimuth and tilt with respect to every user 'K' reduces to optimization problem (6) which optimizes with respect to only few focal points 'S' in the system, reducing the computational complexity.

3.2.2 Achieving Distributed Solution through Concept of Triplet

Equation (6), though has far less number of variables than (5), its solution still requires global coordination among all cells. A distributed solution can be enabled by exploiting the fact that in low power, small cell, high frequency band deployments being envisioned for 5G, interference will not propagate far beyond immediate neighbors. Thus the problem in (6) can be further approximated as

$$\max_{\varphi_{tilt}^{N}, \theta_{a}^{N}} \hat{\xi} = \max_{\varphi_{tilt}^{N}, \theta_{a}^{N}} \sum_{\forall s \in S} log_2 \frac{1}{|N|} \left(1 + \hat{\eta}_s^n \left((\varphi_{tilt}^{N}, \theta_{a}^{N}) \right)\right) \qquad (7)$$

where $\hat{\eta}_s^n$ represents the approximate SIR at the optimal point that considers observations from only its two immediate neighbors as shown in Fig. 2. Lemma 1, corollary 3 in [12], actually proves that for large β, $\hat{\xi}$ approaches the true value ξ. One particular scenario where beta is expected to be significantly large is mmWave based deployment [14], which is being considered as an integral part of 5G landscape.

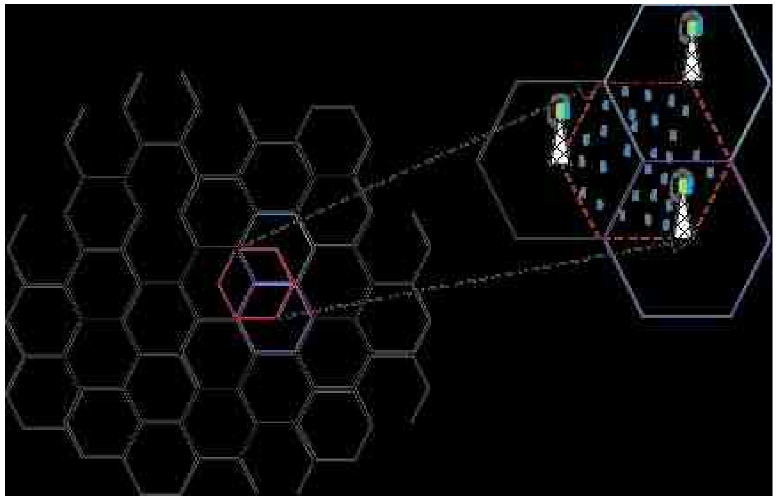

Figure 2: Illustration of concept of triplet

Thus, according to the propositions 1 and 2 in [12], (7) can be expressed as

$$\hat{\xi}_{N,max} = \frac{1}{|N|}\sum_{\forall n \in N}\left\{\max_{\varphi^{N}_{tilt},\theta^{N}_{a}} \frac{1}{|T_n|}\sum_{\forall s \in S_n} log_2\left(1+\hat{\eta}^{n}_{s}\left((\varphi^{N}_{tilt},\theta^{N}_{a})\right)\right)\right\} \quad (8)$$

where, S_n is the set of focal points in n^{th} triplet and $|S_n| = |T_n| = T_n = 3, \forall n = N$. Thus large scale centralized problem represented by (6) is now reduced to small scale distributed optimization represented by (8), for each triplet to be solved independently.

3.2.3 Throughput Optimization of User Clusters in the Triplet

From the analysis presented above our problem in (5) has now been reduced to optimization of azimuth and tilt angle with respect to single focal points (to be determined) in individual triplets of cells. This optimization problem to be solved for every triplet now can be expressed as

$$\max \hat{\xi}\ (\varphi_{tilt}^{1}, \phi_{a}^{1}, \varphi_{tilt}^{2}, \theta_{a}^{2}, \varphi_{tilt}^{3}, \theta_{a}^{3}) \qquad (9)$$

$$\text{s.t.} \qquad 0 < \varphi_{tilt}^{1}, \varphi_{tilt}^{2}, \varphi_{tilt}^{3} < 90 \qquad (10)$$

$$0 < \theta_{a}^{1} < 120 \qquad (11)$$

$$121 < \theta_{a}^{2} < 240 \qquad (12)$$

$$241 < \theta_{a}^{3} < 360 \qquad (13)$$

$$\hat{\xi} = log_2\left(1 + \frac{\delta_1^1 10^{-1.2\left(\lambda_v\left(\frac{\varphi_1^1-\varphi_{tilt}^1}{B_v}\right)^2 + \lambda_h\left(\frac{\theta_1^1-\theta_a^1}{B_h}\right)^2\right)}}{\left(\delta_1^2 10^{-1.2\left(\lambda_v\left(\frac{\varphi_1^2-\varphi_{tilt}^2}{B_v}\right)^2 + \lambda_h\left(\frac{\theta_1^2-\theta_a^2}{B_h}\right)^2\right)}\right) + \left(\delta_1^3 10^{-1.2\left(\lambda_v\left(\frac{\varphi_1^3-\varphi_{tilt}^3}{B_v}\right)^2 + \lambda_h\left(\frac{\theta_1^3-\theta_a^3}{B_h}\right)^2\right)}\right)}\right) +$$

$$log_2\left(1 + \frac{\delta_2^2 10^{-1.2\left(\lambda_v\left(\frac{\varphi_2^2-\varphi_{tilt}^2}{B_v}\right)^2 + \lambda_h\left(\frac{\theta_2^2-\theta_a^2}{B_h}\right)^2\right)}}{\left(\delta_2^1 10^{-1.2\left(\lambda_v\left(\frac{\varphi_2^1-\varphi_{tilt}^1}{B_v}\right)^2 + \lambda_h\left(\frac{\theta_2^1-\theta_a^1}{B_h}\right)^2\right)}\right) + \left(\delta_2^3 10^{-1.2\left(\lambda_v\left(\frac{\varphi_2^3-\varphi_{tilt}^3}{B_v}\right)^2 + \lambda_h\left(\frac{\theta_2^3-\theta_a^3}{B_h}\right)^2\right)}\right)}\right) + \ log_2\left(1 + \right.$$

$$\left. \frac{\delta_3^3 10^{-1.2\left(\lambda_v\left(\frac{\varphi_3^3-\varphi_{tilt}^3}{B_v}\right)^2 + \lambda_h\left(\frac{\theta_3^3-\theta_a^3}{B_h}\right)^2\right)}}{\left(\delta_3^1 10^{-1.2\left(\lambda_v\left(\frac{\varphi_3^1-\varphi_{tilt}^1}{B_v}\right)^2 + \lambda_h\left(\frac{\theta_3^1-\theta_a^1}{B_h}\right)^2\right)}\right) + \left(\delta_3^2 10^{-1.2\left(\lambda_v\left(\frac{\varphi_3^2-\varphi_{tilt}^2}{B_v}\right)^2 + \lambda_h\left(\frac{\theta_3^2-\theta_a^2}{B_h}\right)^2\right)}\right)}\right) \qquad (14)$$

Equation (14) presents the full form of objective function expressed in (9). Although the tilt angle varies from 0 to 90, in practice, the optimal value of tilt generally varies from 0 to 20 degree unless all users in the cell are concentrated near the base of base station. On the other hand, in tri-sector system the optimal value of azimuth can be safely assumed to lie within $\pm$ 15 degree of the nominal azimuth value unless user distribution is extremely skewed towards one edge of the sector, in which case it will be better to serve those users with that sector toward which user distribution is skewed. Our repeated computer simulations show that capping the range of azimuth adaptation is also necessary to limit inter-sector interference. These observations shorten the search space to $20\times20\times20\times15\times15\times15 = 27\times10^6$. This search space can be explored by any state of the art heuristic search algorithm that promises a global solution despite of non-convexity of the objective function. Noting that solution space is fairly small and well defined, we apply Simulated Annealing (SA) to explore the optimal azimuth and tilt and to make sure a global optimal within the curtailed search space is guaranteed.

3.2.4 Proposed Clustering Algorithms for Determining Cell Loci

We investigate three different clustering algorithms to determine focal point (highly user dense region) in each cell which then can be used for azimuth and tilt optimization in real-time fashion. The key common idea behind the three clustering algorithms is that for every user in each cell, the number of users within a pre-specified radius will be determined. The users will be grouped into a cluster if they lie within that radius and if they meet the selected KPIs which are SIR, RSL and MS-BSs distance. The underlying KPI used to group users differentiates the three clustering algorithms. Users are first grouped in clusters based on the selected KPI (SIR/RSL/distance) and clusters are formed with respect to every user present in the cell. The cluster which has the highest number of users is then chosen as the desired cluster. The mean (focal) point of the chosen cluster, for each cell in a triplet, is then determined. Those focal points are then used as the representative point with respect to which the optimization of antenna azimuth and tilt is performed in each triplet. In this study we assume cluster radius of 150 meter, but optimal radii of clusters will vary with cell radii, user distribution and propagation conditions.

The clustering algorithms based on different KPIs are described as follows:

***a) Clustering based on SIR*:** Since users in each cell are served by their serving BS, other BSs in the triplet work as interferer. *In this case,* we first find the mean SIR of every cell in the triplet, which acts as threshold SIR (SIR_{th}) for the users to be grouped into a cluster in addition to the condition that they should lie within the radius. In other words, to form the cluster around every user, the cluster user*s* should be within the radius and the*ir* received SIR should be greater than or equal to SIR_{th} of the cell under evaluation. Similar procedure is followed in others cells in the triplet, which will have their own $SIR_{th.}$ The pseudocode for the algorithm is given in Fig. 3. Note that in emerging future cellular networks, with advent of location based services, accurate locations of individual users within a cell is known by the network. Given that location information is available at each base station, the SIR can be estimated using (4). Alternatively, third generation partnership project (3GPP) Channel Quality indicator (CQI) reported by users can also be exploited to estimate real time SIR. Thus, the clustering algorithms proposed here are implementable in an online fashion with no additional signaling overhead.

Figure 3: Algorithm for SIR based clustering

b) ***Clustering based on RSL*:** This method is similar to the one with SIR with the difference that SIR is replaced with RSL, i.e. users are clustered based upon threshold RSL_{th} and number of users within given radius of cluster. The pseudocode for this algorithm is given in Fig. 4 below. In emerging cellular networks, minimization of drive test (MDT), recently standardized by third 3GPP, contains RSL reports [15]. Thus standardization of MDT allows online implementation of the algorithm in Fig. 4 without additional signaling overhead.

Figure 4: Algorithm for RSL based clustering

c) ***Clustering based on distance from the BS*****:** This method also considers the highest number of user within the pre-specified cluster radius with a center being the base station. However, the second criteria in this scheme is that the distance of the user from its serving BS is assumed to be lesser than the distance to the interfering BSs. The cluster formed using this method will ensure that it is nearer to the serving BS. This algorithm will be more useful to realize under cell-less architecture where any user can connect to any BS depending upon signal availability. The pseudocode for this algorithm is given in Fig. 5.

Figure 5: Algorithm for distance based clustering

The choice of algorithm depends upon the motive behind clustering. For example, SIR based clustering will ensure that most of the users will be served with higher throughput most of the time. Similarly, RSL based clustering can be chosen where throughput fairness is the major objective, while distance based clustering can be used for scenarios where there are no high rise buildings and towers that can work as interferers. Since their implementation, performance and complexity vary, there lies trade-off to be considered while selecting one scheme over the other. We also noticed that all three algorithms require the user density to be the highest but the location of the cluster for same distribution of user will differ based on the selected KPI.

3.2.5 Determining Optimal Radius

As discussed earlier, the optimal threshold radii used in clustering algorithms will vary with a number of factors, e.g. cell size, user distribution etc. Hence it is essential to determine optimal radius for a given scenario. Optimal radius calculation becomes more important while realizing cell-less architecture in emerging 5G cellular network, in that optimal radius would play key role in determining width of the beam directed toward each user. Furthermore, in emerging user-centric small cell network (SCN), a number of small cells will be closely spaced and will cooperate to serve group of users overcoming the interference. Moreover, in emerging D2D communication, the clustering of users analogically applies to determine optimal radius within which D2D communication can be effectively realized.

To maximizing spectral efficiency in our case, we consider clustering based up on SIR and user density. Thus the spectral efficiency is a function of SIR experienced at the user and total number of user (denoted by *N*) in the cell.

$$SE\ (\xi) = f(SIR\ (azimuth, tilt), N) \qquad (15)$$

In our case, we considered the radius of cluster to be constant for the assumed user distribution. However, considering emerging SCN and D2D scenarios, we assume that it can take values between $50 < r < 150$.

If the radius of the hexagon is denoted by R, then the optimal radius can be determined by solving

$$\hat{r} = \operatorname{argmax}_{r\in[50,150]} \left(\left(\pi r^2 \times \frac{2N}{3\sqrt{3}R^2}\right) \times SIR_n\right) \qquad (16)$$

where $\hat{r}$ denotes the optimal radius of the cluster. It is important here to note that the cluster of range of radius $50 < r < 150$ is formed around every user and the radius satisfying (16) is chosen to be optimal. The SIR_n indicates the SIR at each user $(n = 1, .., N)$. The above optimization problem is a differentiable and constrained problem; it can be easily solved through any nonlinear optimization algorithm.

3.3 Performance Evaluation

In this section, we evaluate the system performance and present numerical results. The parameters used for evaluation were 3GPP recommended simulation parameters for LTE systems, given in Table 1.

Table 1: Simulation Parameters

Parameters	**Values**
System Topology	19 BS with 3 sectors per BS
BS Transmission Power	46 dBm
BS Inter-site Distance	500 meter
BS and UE height	32 meter, 1.5 meter
UE Antenna Gain	0 dB
Vertical Beamwidth	70^0
Horizontal Beamwidth	10^0
Vertical Gain Weight, λ_v	0.5
Horizontal Gain Weight, λ_h	0.5
Maximum Gain, G_{max}	18 dB
Maximum Attenuation, A_{max}	25 dB
Shadowing	8 dB
Frequency	2 GHz
Path Loss Model	3GPP Urban Macro

To obtain the optimum azimuth and tilt, we applied heuristic search method 'Simulated Annealing' [16]. The detail of this method is skipped as it is out of scope. We also implemented sequential quadratic programming (SQP) technique to achieve optimal value of the antenna azimuth and tilt. Optimal azimuth and tilt were found using a brute force algorithm and the obtained results were compared to those achieved using the SQP algorithm. The SQP results were confirmed and validated. We compared the bandwidth normalized throughput obtained by the proposed self-organized azimuth and tilt optimization framework, referred to as SAT hereforth, with those obtained using fixed azimuth and tilt optimization which is referred to as fixed azimuth and tilt optimization (FAT) in [12].

3.3.1 Performance Comparison of FAT and SAT Framework

Figure 6 depicts the commutative density function (CDF) of SE achieved at the focal points using the proposed SAT framework that uses SIR based clustering, with that that obtained using FAT. The antenna tilt angle was fixed at 5^o and 12^o with fixed typical azimuth of 0^o, 120^o, 240^o for cell 1, 2 and 3, respectively in the triplets for FAT scenario. We observe that at lower tilt of 5^o, the performance is poor because at smaller tilt the beam is pointing towards the edge of the cell and hence exposed to higher interferences from other BSs. For uniform user distribution and for standard BS height and user end (UE) height, the fixed optimal tilt is centered around 12^o[12]. At the fixed tilt of $12^o$and at regular values of azimuth, the performance improves as compared to that at 5^o as the interference from other BSs decreases. However, if the tilt goes on increasing, the performance degrades and limits the coverage at the cell edge. In Fig. 6, we observe that at fixed tilt 5^o with regular azimuth, 60% of the user experiences SE of up to 6 bps/Hz, whereas at tilt of 12^o and regular azimuth, 60% of the user get up to 8 bps/Hz while with FAT framework outperforms with 10.6 bps/Hz. Also, maximum SE of up to 7 bps/Hz and 10 bps/Hz is achievable with SAT framework, i.e. at tilt of 5^o and 12^o respectively and with regular azimuth, whereas with FAT framework the achievable SE goes up to 12.8 bps/Hz. Thus using the SAT framework, SE gain of 0.8 to 2.8 bps/Hz is achievable as compared to that obtained using FAT framework. The SAT framework adjusts its azimuth and tilt automatically based on user density, thus optimizing the throughput at UE. It is clearly confirmed that the SAT technique outperforms the fixed scheme FAT.

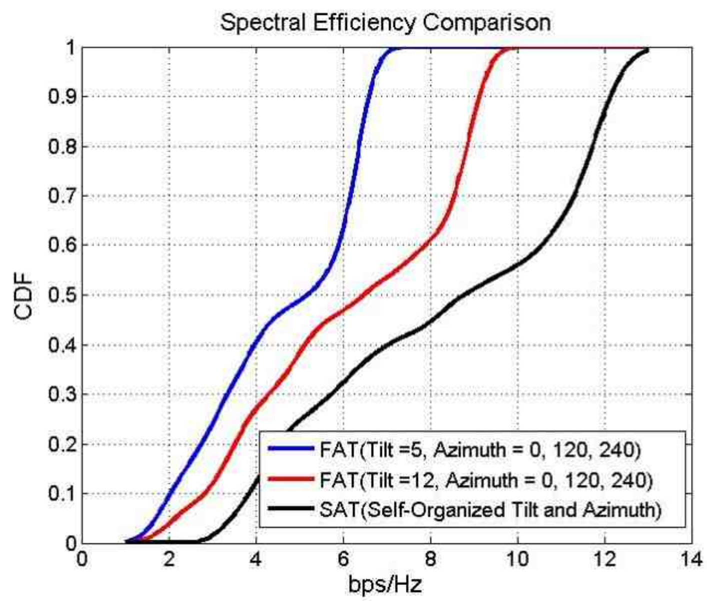

Figure 6: CDF Representation of achievable SE using FAT and SAT framework

In Fig. 7, we observe that the proposed SAT framework achieves per user average throughput of 6.5 bps/Hz, 6.2 bps/Hz and 5.5 bps/Hz using SIR, RSL and MS-BS distance based clustering respectively, whereas that obtained using FAT framework with fixed tilt of 12^o and regular azimuth values of $0^o, 120^o, 240^o$ was 4.5 bps/Hz. Thus SAT framework outperforms FAT framework by per user average throughput of at least 1 bps/Hz.

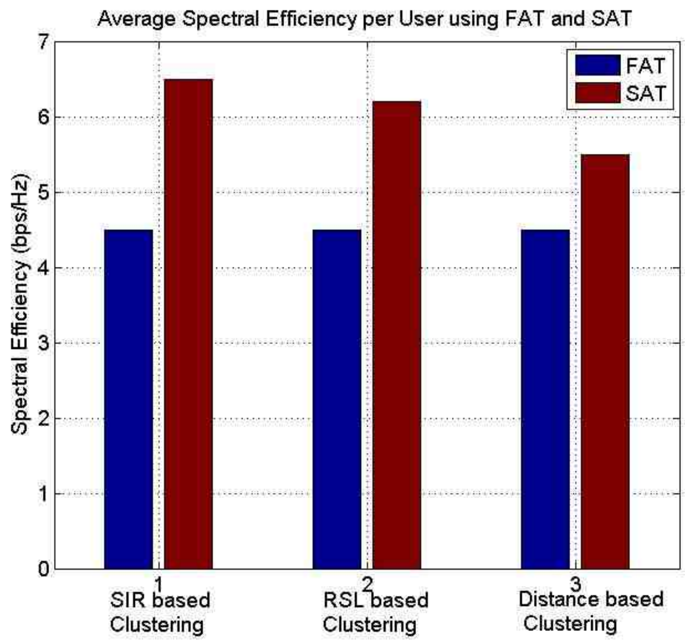

Figure 7:Average spectral efficiency per user obtained using FAT and SAT

3.3.1 Performance Comparison of Clustering Schemes

Figure 8 presents the effect of different clustering techniques on SE performance, as the location of optimal focal points will vary for the investigated clustering schemes. The performance has been observed with fixed tilt of 12^o and at regular azimuth of 0^o 120^o 240^o. It can be noticed that SE gain of 2 bps/Hz, 1.5 bps/Hz and 1.1 bps/Hz is achievable using clustering based on SIR, RSL and MS-BS distance respectively and adjusting antennae' tilt and azimuth applying SAT framework. Thus, it is possible to achieve 20% to 26% gain in overall spectral efficiency, the highest being for the clustering based on SIR.

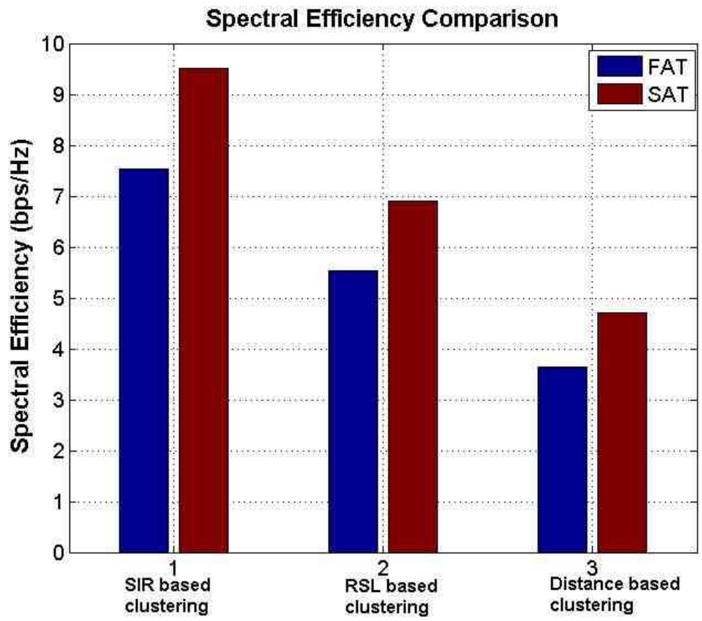

Figure 8: Comparison of bandwidth normalized throughput obtained using FAT and SAT

The SIR based clustering performed the best as it considers not only the signal strength of the serving BS but also the interferences from other BSs. Other clustering algorithms although considers signal strength as well as the relative distance to the BS, they don't incorporate interference directly. Hence their performances are not as good as that of SIR based SAT. However, it is worth noting that the RSL and distance- based SAT are implementable with lower computational complexity, thereby offering a trade-off between performance and complexity.

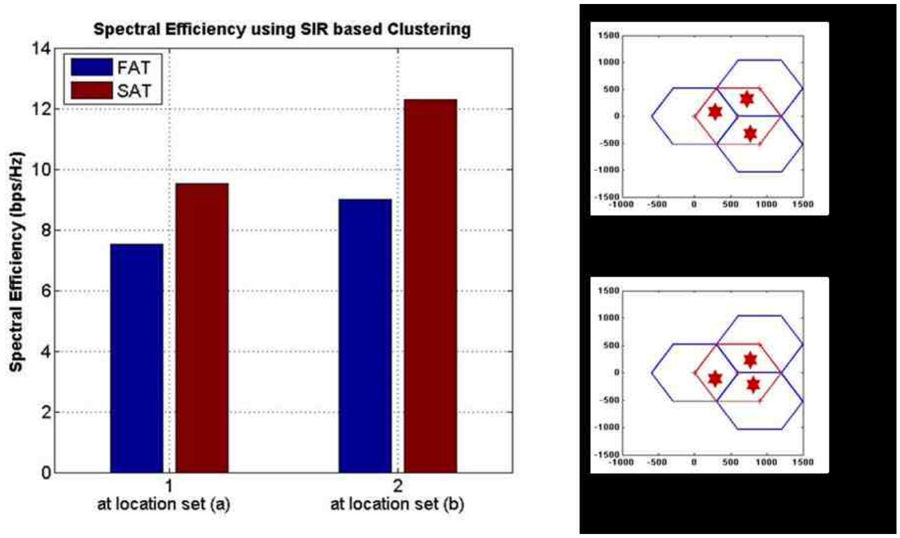

Figure 9:Spectral Efficiency gain using FAT and SAT for different focal points based on SIR clustering

Moreover, for each algorithm the gain in SE varies depending upon the location of the focal points determined through the clustering process. Figure 8 shows the variation in SE gain due to different focal point locations determined using SIR based clustering. Focal point location will change with different distribution of users thus it is another major aspect to be considered while considering clustering. As we can see the SE achieved at location in Fig. 9(b) is higher than that in Fig. 9(a). Not to mention, other clustering algorithms discussed in earlier section will result in different focal points and hence different SE gain.

Chapter 4: Emerging Technologies and Approaches for Spectral Efficiency Optimization

Spectrum is the scarcest resources in the network mainly because of its lack of availability and extra high cost. Therefore, technologies and approaches are continuously evolving to optimally utilize available spectrum in order to maximize the revenue. In this chapter, we cover major technologies and approaches widely acknowledged as 5G candidates for meeting exponentially increasing data rate demand in emerging cellular networks.

4.1 Emerging Technologies for Spectral Efficiency Maximization

Some of the widely foreseen technologies for spectral efficiency enhancement in 5G are ultra-dense BS deployment, increased spectrum, large scale antenna system, D2D communication, interference management, cooperative transmission etc. However, we limit our discussion to only first three of these techniques and few important approaches.

4.1.1 Base Station Densification

Network capacity can be significantly enhanced by shrinking the cell size, each cell covering an area of as small as 200 meters. Large number of small cells allows reuse of spectrum across a given geographic area, reducing the number of users competing for resources at BSs [17]. According to [18], the SIR is preserved even if the cell size shrinks indefinitely until every BS serves a single user. 5G cellular network expects deployment of thousands of small cell nodes per unit area and per unit Hz. This will result into data rate many times more bits/s/Hz for every unit of area.

4.1.2 Increased Spectrum (mmWave)

The current wireless systems' operations are restricted to a short range of frequencies ranging from several hundred MHz to few GHz, also known as *beachfront spectrum* [17]. This spectrum is nearly fully occupied, especially when the demand is peak. If we observe the advancement in technologies in different layers to enhance the spectral efficiency, we find that all techniques and technologies has already reached Shannon limit and falls far short as compared to exponentially rising data rate demand. Thus, along with innovation and invention in technologies, much more bandwidth is needed to meet the future application requirement [19] [20]. A large portion of spectrum exists in mmWave (millimeter Wave) range between 30 – 300 GHz, with reasonable spectrum in 20 – 30 GHz range. Researches are already considering frequencies within mmWave range to test suitability for 5G [21]. Apart from bandwidth proportional data rate enhancement (according to Shannon relation) using mmWave, BS densification will simply multiply the data rate for every unit of bandwidth resource. Thus the combination of more nodes per unit area per Hz and more Hz (from mmWave) will give rise to spectral efficiency by even higher bits/s/Hz per unit area than just BS densification.

4.1.3 Large Scale Antenna System (Massive MIMO)

With a vision of putting a number of antennae much larger than number of active users per time-frequency resources and assuming possibility of accurate channel estimation for tens of users per resources under reasonable time-frequency selectivity, authors in [22] came up with initial idea of putting hundreds of antennae at BS. This idea was known as large scale antenna system, also popularly recognized as *massive*-MIMO. In other words, massive MIMO, with the help of large number of antennae at the BS simultaneously serves tens of users in the same time-frequency resource resulting in increase in capacity by more than 10 times. An experiment performed in [23] finds that for an antenna array consisting of 6400 omnidirectional antennas at the BS transmitting over 20 MHz bandwidth in PCS band serving 1000 fixed users distributed randomly, provides 95% of the terminals a throughput of 21.2 Mbps/terminal. This offers a total downlink throughput of 20 Gbps resulting in sum-spectral efficiency of 1000 bps/Hz and this is sufficient to provide 20 Mpbs broadband service to each of a thousand homes [23]. Thus Massive MIMO is considered as game-changing technologies for boosting spectral as well as energy efficiency in 5G.

4.2 Emerging Approaches for Spectral Efficiency Maximization

Apart from technologies, numerous frameworks and approaches to be implemented at different level of network are currently under investigation for possible inclusion in 5G. Examples are centralized (cloud) RAN (CRAN), software defined network (SDN), network function virtualization (NFV), self-organizing network (SON), caching at the edge nodes, big data empowered network system design etc. All these technologies play essential role in availing spectral efficiency gain in one way or other. In this section, we shorten our discussion to only following three topics.

4.2.1 Self-Organizing Network

Self-Organizing Network (SON) describes abilities of a network to autonomously configure when a network is in deployment stage and autonomously optimize and heal the network after the network enters operation stage. SON was addressed in 3GPP Release 8 with the aim to automate the network so that network can function in every situation without or with least human intervention. Although SON was not paid much attention until Release 10, researchers have recognized it as an indispensable need for enabling 5G. With the rise in number of nodes, antenna and terminals, it would be extremely difficult to perform network configuration, optimization and healing with the help of offline tools. Moreover, in current heterogeneous network scenario, user demands are changing dynamically. To respond promptly and precisely to those demands, emerging network needs to be embedded with intelligence to predict users and network demand in proactive manner, so that network resources such as bandwidth can be allocated in advance and optimally utilized in self-organizing manner [3]. Additionally, 5G network should be able to proactively identify probability of fault occurrence in the network, so that network can prepare necessary action plan prior to occurrence of any problem enabling self-healing ability of the network. Recently, with the help of big data analytics (explained in the section followed) and intelligent machine learning tools, intelligent functionalities can be enabled in the network making a completely self-organized network.

SON features can be utilized in the network for any functions. For spectral efficiency enhancement, SON will intelligently utilize the available spectrum without wastage. For example, using SON, a small cell detects that there are only few users in its coverage, thus SON engine will allocate only few resource blocks for that small cell and the rest for other cells which have higher number of users. Nevertheless, this is a simple example, SON actually can be applied for various nested functionalities to enable coordination among conflicting parameters and functionalities to ensure stable operation of the network.

4.2.1 Big Data Analytics Driven Spectral Efficiency Optimization

Although data analytics has existed since decades back, big data analytics has emerged as a revolutionary approach for improving every walks of life. Big data analytics is identified through its widely accepted features: volume, velocity, variety. These unconventional 3V features (volume, velocity, variety) of current data generation give birth to big data and thus its management and analysis require big data analytics scheme. Big data analytics is an umbrella term, that incorporates methods and technologies, hardware and software for collecting, managing and analyzing large scale structured and unstructured data in real-time. Big data analytics works on entire data collected from every sources in the system as opposed to only sample data in conventional data analytics schemes. Thus big data analysis is expected to create data-driven opportunities as opposed to hypothesis-based solutions in conventional data analytics scheme.

In cellular networks, data are generated at very large scale (volume) with expected size of 24.3 Exabyte (EB) per month [1], with fast input/output to/from the network (velocity) and from various sources within and outside the network (variety). The various data sources in cellular (LTE-Advanced) network are UE at user side, eNodeB at RAN layer and nodes such as MME, PGW, SGW, PCRF at CN layer. Data collected from these nodes can be utilized in a number of ways for variety of purposes. From spectral efficiency optimization point of view, the network first needs to understand spatio-temporally varying user's demand. With the help of advanced data analytics scheme and intelligent learning tools, data collected from social networks such as Facebook, YouTube, Instagram, WhatsApp etc. can be analyzed to understand user's contextual information that includes user's temporal location, mobility pattern, network usage behavior, user's social connection and ties, user's choice of contents and dynamicity of choice and so on.

Having known above information, network can predict user's future context information and accordingly allocate necessary resources on real-time basis. This makes the network intelligent to optimally utilize resources at right place in right time and in right quantity. For example, with the help of big data, network can predict user's choice of content types at a particular time and place, and thus will allocate the bandwidth resource accordingly instead of consistently assigning certain bandwidth for that user. So, big data analytics assists in enabling on-demand services and resource utilization, optimizing bandwidth utilization.

From above discussion, it can be inferred that, big data analytics is actually making the network intelligent which partially overlaps with the objective of self-organizing network. In fact, big data analytics is the enabling approach that helps to visualize the end-to-end picture of present, past as well as future of user and network. Big data analytics digs deeper down and explore wider in the ocean of data to extract actionable insights. Machine learning tools are then applied on the extracted information in order for the network to make intelligent decisions. Thus big data analytics is the core concept behind network intelligence.

4.2.2 Caching at the Edge

Caching has emerged as one of the most disruptive technologies for enabling 5G [5]. It basically brings most popular contents to the network nodes closest possible to the user. This facilitates the users to be served with minimum delay without requiring backhaul usage, which in turn reduces backhaul bandwidth requirement. However, an efficient caching depends upon various factors such as cacheability of contents, cache location, caching strategy and last but not the least cache size. In other words, *what to cache*; *where to cache*; and *how to cache* are the important questions behind making caching approach effective and efficient. For caching to be beneficial, content-level characterization of cellular traffic is essential for strategizing caching effectively. A study on 3G traffic reveals that more than 95% of the downlink traffic flow uses transmission control protocol (TCP) [24]. Additionally, since hypertext transfer protocol (HTTP) has become workhorse for various applications, from data to video, HTTP dominates the cellular traffic with more than 80% of the average downstream traffic [25]. Depending upon caching techniques, up to 42% bandwidth can be saved by caching redundant TCP flows, and up to 37% during peak usage in night time [24].

Similarly, the experiment in [25] shows that up to 70% of overall user requests were found to be cacheable, 49% of them could be served from the cache.

Thus caching has become an integral part and an inevitable means for enabling 5G. With the advent of big data analytics in conjunction with machine learning tools, it has become possible to predict, track network conditions as well as build user demand profile to act proactively. By understanding user's spatio-temporal network usage pattern and social ties, necessary resource can be allocated to the relevant node (including UE) for caching the forecasted contents. It is not surprising to note that device to device (D2D) communication further enables to cache and disseminate contents among the group of users with similar preference. Proactivity in caching would expedite the delivery of content to the desired location that would reduce not only the response delay and backhaul load but also optimize the network resource utilization.

In cellular networks, caching basically stores popular contents demanded by the user at different layers (layer 0: **UE**, layer 1: radio access network (**RAN**), layer 2: Core Network (**CN**)) of the network, thus reduces the need to access each content from the main server every time a user requests. This avoids a user request to traverse the whole network (UE – RAN – CN – Internet – CN – RAN – UE) to receive a requested content. Since it reduces the use of backhaul link but requires storage memory at different layers to store popular contents, it is basically replacement of backhaul **but not** in same proportion. Because the cost of storage memory is much cheaper than cost of bandwidth, caching of contents is always preferred over utilization of backhaul. In addition to backhaul bandwidth reduction, UE receives the content from the nearest cache thus minimized service delay and improved QoE.

Chapter 5: Conclusion and Future Work

Bandwidth resource optimization has consistently been challenging task especially for mobile network operators (MNOs). MNOs always keep an eye on the optimization techniques especially on those that can be availed at minimum cost and that can be implemented with minimal human intervention.

In this thesis, we presented a novel framework for dynamic optimization of spectral efficiency with respect to varying user density utilizing antenna azimuth and tilt angles. To achieve system-wide spectral efficiency, we presented two important concepts as foundation for the proposed framework. First, we discussed and analytically developed the self-organizing mechanism to simultaneously configure two key antenna parameters (azimuth and tilt) in automated fashion. We decomposed the large-scale computationally taxing task into small-scale tasks by introducing the concept of triplet. Second, we determined single focal point in each cell of a triplet that effectively represent all the users in that cell. We evaluated three different clustering algorithms used to calculate the focal points that can best represent a given user distribution. These algorithms offer different levels of trade-offs in implementation complexity and performance gain. The SIR based clustering algorithm offers the highest gain. We also proposed a method to determine optimal radius of a user cluster that further enables dynamic user centric optimization, which is essential to manage the continuously changing user spatial distribution. We compared the SE gain against the conventional setting of fixed azimuth and tilt angles (FAT). Obtained results confirm a 20% to 26 % system-wide gain using the proposed scheme based upon clustering technique.

Moreover, the average throughput per user obtained was the best in case of SIR based clustering with 6.5 bps/Hz followed by RSL and MS-BS clustering with values 6.2 bps/Hz and 5.5 bps/Hz respectively. This indicates gain in average spectral efficiency per user of 1-3 bps/Hz as compared to conventional FAT scheme with per user spectral efficiency of 4.5 bps/Hz. This performance was achieved considering uniform user distribution, and different user distribution will result in different gain. Thus SAT scheme results in spectral efficiency gain of at least 1 bps/Hz/user and overall spectral efficiency gain of 20% to 26%.

5.1 Future Works

The proposed framework is expandable towards cell-less deployment architecture in next generation network 5G, where cells are expected to adapt their sizes and shapes in user centric fashion by harnessing the flexibility inherent in multi element antenna systems. The proposed solution is also applicable to mmWave based systems as the interference in such system is mainly inter-sector interference and interference from neighboring cells can be neglected to implement distributed, low complexity, online cell footprint adaptation to make the best of spatio-temporally changing user distributions. In the following, we discuss the possible future work in the proposed system to realize higher gain in spectral efficiency.

5.1.1: Consideration of Non-Uniform User Distribution

The performance of the proposed framework was realized using uniform user distribution. Consideration of user distribution that can represent (closer to) realistic scenario will result in understanding and realizing real-network performance. For example, hotspots like stadium, railway stations etc. are common cases to be considered while optimizing the network. Moreover, under a single antenna coverage (i.e. within a cell), more than one such hotspots can appear dynamically. The proposed framework can be extended in such direction by taking into account different non-uniform user distribution.

5.1.2: Consideration of User's Mobility

In the proposed system, all users were considered to be stationary, whereas the users can exhibit different nomadicity, i.e. users can be moving with less-frequency/high frequency, in a real-network scenario. In such cases, how to adjust antenna's azimuth and tilt to meet varying movement dynamics of the user such that overall spectral efficiency is optimized without compromising user's quality of experience. Towards this, exploitation of user's movement information and proactive allocation of proper resources for user's next location can be a possible solution. Apart from user's movement, various other information can be exploited with emerging big data analytics tools that can be further utilized to track different behavior of users to serve them with optimized QoS. A glimpse of this is described in the sub-section followed.

5.1.3: Exploitation of User's Contextual Information

The framework discussed can be expanded and can be proved to be highly effective by addressing the notion of '*Caching at the Edge*'. Recently, with huge penetration of social media, user's contextual information can be easily predicted with the help of highly efficient big data analytics and intelligent machine learning tools. Users' context information includes their whereabouts, network usage behavior, social connection and ties, their choice of contents, dynamicity of popularity of contents etc. Having known user related above information, popular contents can be proactively cached at the nodes closest to the user. This will reduce the need for the contents to be fetched from origin server every time the contents are requested. Additionally, it will decrease backhaul bandwidth requirement which instead can be effectively and efficiently utilized during peak service hours. Since caching enables desired (popular) contents available at the edge nodes, backhaul load can be significantly reduced and users can be served with minimum delay. Thus introduction of caching in the discussed framework will eventually decrease the total cost of operation (TCO) for MNOs and improve QoS of the users.

In addition to the user's choice of contents, user's mobility and network usage behavior can be predicted that will further enable the network to strategize network resource allocation in right time and in right place. Big data analytics of user's context information will unearth a number of key information that can potentially be exploited in a variety of ways which actually cannot be visualized before their analysis. Similar will be the case with big data analysis performed on the network parameters. Thus exploitation of big data analytics technique on user as well as network information will open numerous unimagined directions for network optimization.

Bibliography

[1] Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update 2015-2020", *White Paper, [Online]*, http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.pdf

[2] Ericsson, "Ericsson Mobility Report, November2015" White Paper, [Online], http://www.ericsson.com/res/docs/2015/mobility-report/ericsson-mobility-report-nov-2015.pdf.

[3] Imran, A., Zoha, A. and Abu-Dayya, A., 2014. Challenges in 5G: how to empower SON with big data for enabling 5G. *IEEE Network*, *28*(6), pp.27-33.

[4] Aliu, O.G., Imran, A., Imran, M.A. and Evans, B., 2013. A survey of self organisation in future cellular networks. *IEEE Communications Surveys & Tutorials*, *15*(1), pp.336-361.

[5] Boccardi, F., Heath, R.W., Lozano, A., Marzetta, T.L. and Popovski, P., 2014. Five disruptive technology directions for 5G. *IEEE Communications Magazine*, *52*(2), pp.74-80.

[6] Shen, X., 2015. Device-to-device communication in 5G cellular networks. *Network, IEEE*, *29*(2), pp.2-3.

[7] Tehrani, M.N., Uysal, M. and Yanikomeroglu, H., 2014. Device-to-device communication in 5G cellular networks: challenges, solutions, and future directions. *Communications Magazine, IEEE*, *52*(5), pp.86-92.

[8] Mohamed, A., Onireti, O., Imran, M.A., Imran, A. and Tafazolli, R., 2015. Control-data separation architecture for cellular radio access networks: A survey and outlook. *IEEE Communications Surveys & Tutorials*, *18*(1), pp.446-465.

[9] Saur, S. and Halbauer, H., 2011, May. Exploring the vertical dimension of dynamic beam steering. In *Multi-Carrier Systems & Solutions (MC-SS), 2011 8th International Workshop on* (pp. 1-5). IEEE.

[10] Halbauer, H., Saur, S., Koppenborg, J. and Hoek, C., 2012, April. Interference avoidance with dynamic vertical beamsteering in real deployments. In *Wireless Communications and Networking Conference Workshops (WCNCW), 2012 IEEE* (pp. 294-299). IEEE.

[11] Seifi, N., Coldrey, M. and Svensson, T., 2012, December. Throughput optimization in MU-MIMO systems via exploiting BS antenna tilt. In *2012 IEEE Globecom Workshops* (pp. 653-657). IEEE.

[12] Imran, A., Imran, M.A., Abu-Dayya, A. and Tafazolli, R., 2014. Self organization of tilts in relay enhanced networks: a distributed solution. *IEEE Transactions on Wireless Communications*, *13*(2), pp.764-779.

[13] Viering, I., Dottling, M. and Lobinger, A., 2009, June. A mathematical perspective of self-optimizing wireless networks. In *2009 IEEE International Conference on Communications* (pp. 1-6). IEEE.

[14] MacCartney, G.R., Zhang, J., Nie, S. and Rappaport, T.S., 2013, December. Path loss models for 5G millimeter wave propagation channels in urban microcells. In *2013 IEEE Global Communications Conference (GLOBECOM)*(pp. 3948-3953). IEEE.

[15] Zoha, A., Saeed, A., Imran, A., Imran, M.A. and Abu-Dayya, A., 2014, September. A SON solution for sleeping cell detection using low-dimensional embedding of MDT measurements. In *2014 IEEE 25th Annual International Symposium on Personal, Indoor, and Mobile Radio Communication (PIMRC)*(pp. 1626-1630). IEEE.

[16] Peyvandi, H., Imran, M.A. and Tafazolli, R., On Performance Optimization in Self-Organizing Networks using Enhanced Adaptive Simulated Annealing with Similarity Measure.

[17] Andrews, J.G., Buzzi, S., Choi, W., Hanly, S.V., Lozano, A., Soong, A.C. and Zhang, J.C., 2014. What will 5G be?. *IEEE Journal on Selected Areas in Communications*, *32*(6), pp.1065-1082.

[18] Ramasamy, D., Ganti, R. and Madhow, U., 2013, July. On the capacity of picocellular networks. In *Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on* (pp. 241-245). IEEE.

[19] Connecting America: The National Broadband Plan (Chapter 5) 2010. [Online]. Available: http://download.broadband.gov/plan/national-broadband-plan.pdf

[20] "Estimated spectrum bandwidth requirements for the future development of IMT-2000 and IMT-Advanced," ITU-R, Geneva, Switzerland, Rep. M.2078, 2006.

[21] Rangan, S., Rappaport, T.S. and Erkip, E., 2014. Millimeter-wave cellular wireless networks: Potentials and challenges. *Proceedings of the IEEE*,*102*(3), pp.366-385.

[22] Marzetta, T.L., 2010. Noncooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Transactions on Wireless Communications*, *9*(11), pp.3590-3600.

[23] Larsson, E.G., Edfors, O., Tufvesson, F. and Marzetta, T.L., 2014. Massive MIMO for next generation wireless systems. *IEEE Communications Magazine*, *52*(2), pp.186-195.

[24] Woo, S., Jeong, E., Park, S., Lee, J., Ihm, S. and Park, K., 2013, June. Comparison of caching strategies in modern cellular backhaul networks. In*Proceeding of the 11th annual international conference on Mobile systems, applications, and services* (pp. 319-332). ACM.

[25] Erman, J., Gerber, A., Hajiaghayi, M.T., Pei, D., Sen, S. and Spatscheck, O., 2011. To Cache or not to Cache: The 3G case. *Internet Computing, IEEE*, *15*(2), pp.27-34

Appendix A

Publications

Parwez, M.S., Farooq, H., Imran, A., Refai, H., "Spectral Efficiency Self-Optimization Through Dynamic User Clustering and Beam Steering," *IEEE Globecom*, San Diego, USA, Dec. 2015

Farooq, H., Parwez, M.S., Imran, A., "Continuous Time Markov Chain Based Reliability Analysis for Future Cellular Networks," *IEEE Globecom*, San Diego, USA, Dec. 2015

Parwez, M.S., Imran, A., Verma, P., "Caching Approaches for Enabling 5G Cellular Networks" (Under Review for IEEE Survey & Tutorial)

Parwez, M.S., Taufique, A., Imran, A., "Big Data Analytics for Emerging Cellular Networks: Challenges and Opportunities from 5G Perspective" (Under Review for IEEE Proceedings)

Parwez, M.S., Farooq, H., Zoha, A., Imran, A., Abu-Dayya, A., "CDR Based Anomaly Detection in Emerging Wireless Cellular Network," (Under Review for IEEE Globecom 2017)