# Network Design and Power Allocation in 5G Networks

Priyanka Zade, Dr. Achala Deshmukh
Student Department of Electronics and Telecommunication,
Professor Department of Electronics and Telecommunication,
Sinhgad College of Engineering, Pune, Maharashtra
India

***Abstract:*** The future demand device- to-device (D2D) cellular services is satisfied by the heterogeneous 5G communication. In cellular network on mobile Internet related services has increased the need for higher bandwidth by the demand. However multi-tier 5G in power allocation among secondary transmitter that maintain quality of services (QoS). In this paper, we can used a non-cooperative methodology with by focused on the power allocation problem in the multi-tier 5G network. Therefore, we propose energy efficiency consideration in the 5G system for power allocation on online learning scheme, where power allocation approach can be used transmitter hypothesis for without information exchange. The concurrence of achieve in system energy efficiency by proved the proposed scheme and the numerical results determine the capacity and fast convergence with QoS guarantee.

***Index Terms* – online learning, 5G networks, power allocation, energy efficiency.**

## I. INTRODUCTION

Digital mobile communication access market in high number of mobile user, service provider and data services. The existing standard (e.g. LTE-A, HSPA and Wi-Fi) having a global unified platform that provide seamless connectivity among the vision of 5G network. Data service has to satisfy the future demand by the 5G is a promising technology as it is provide to end to end latency of 2 to 5 milliseconds wth high data rate up to 10 Gbps [1]. In mobile data user will not able to new applications for increase the future usage in current wireless system. The heterogeneous wireless system is multi-tier network is one of the 5G structure with conceptualize transmission power and aberrant smart device [2]. Consider this two tier: primary and secondary tier. The primary tier is high power microcells and secondary tier be composed to Device to Device (D2D) communication, picocells,and femtocells. The secondary base station (BSs) and secondary transmitter (STs) involves D2D transmitters are denoted as pico and femto BSs [3]. They apply the available resources in an decide mode affected on micro tier delay below threshold. However, the power consumption which increases the interference to the primary tier and they tend to increase performance of transmission power to maximize. The secondary users the non- cooperative several interference to access, increase the QoS of the micro users power consumption.[4]

Efficient resource management and a key feature of 5G networks is essential for power control and interference allocation among heterogeneous networks to be optimize the usage of network resources nodes, and it is optimize to necessary the usage of network resources. This requires learning how to move toward additional solutions to a wider base and improve the model to better appeal. In this paper provide Integer Linear programming (ILP) for D2D communication has been based on a two-tier 5G heterogeneous network. Finally, power efficiency on how Fractional knapsack using greedy algorithm used in experiments and allowing of future accumulation problems [5].

The energy efficiency is emphasize and motivation in 5G network. This paper explain it system description and problem formulation. This will be described the power allocation structure to identify the challenging issues. Also describe the intuitive scheme for power allocation. This system will be used in future demand for easily communicate with one person other.

## II. SYSTEM DESCRIPTION

This section, problem formulate the power allocation and describes the the system model.

### 2.1. System Model

Our system provide for the two tiers: primary tier and secondary tier, the primary tier provides of the macrocell (MUEs) and the secondary tier consist of picocells and femtocells for the D2D communications as in Figure 1. All the D2D communicator and secondary tier BSs are affected to be inconsistently coverage of the macrocell. The secondary BSs is devote by N ={1,2…,N}, the affiliated users (SUEs) are devoted by X ={1,2….,X}, the set of MUEs is consider by K ={1,2….,K}, and the active D2D pairs as D ={1,2….,D}. The dth D2D team provide by the D2D transmitter and receiver, where DT ={1,2…,DT}and DR ={1,2….,DR}. The nth secondary BS is referred as Xn can associate with the assumption by each one secondary BS. problem of resource allocation introduce the power allocation and SUE. They can combination with a single BS.

A two tier 5G network can be identify the microcells and devices. The device tier is the D2D communication architecture and microcell tier is the conventional architecture. Two tier communication define Device Relaying with Operator Controlled. Device Relaying with Operator Controlled is relays message from one device to another device.

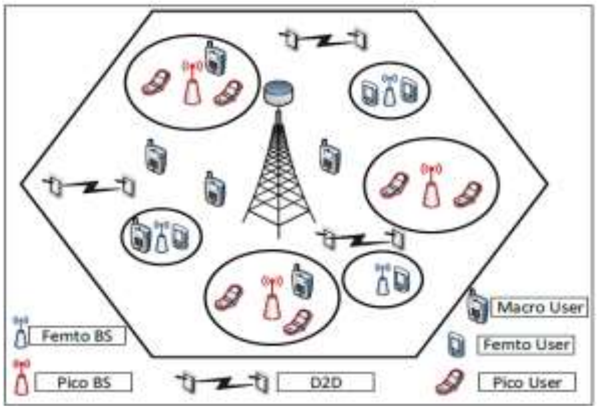


Figure 1. System Model

**2.2. A Network Design Problem Formulation**

Power allocation can be realize the non-cooperative efficient, the ratio of the efficiency in the process of power allocation energy to the power absorb by transmitters and BSs as follows,

$$EE_i = \frac{B\log_2(1+\gamma)}{P_i+P_{cc}} \quad (1)$$

where $P_{cc}$ is the power absorb in ST, B is the bandwidth and.$\gamma$ is the SINR accomplish at the secondary receiver. And index $i$ is the secondary transmitter (ST). Therefore, the 5G heterogeneous structure has the optimization problem by the non-cooperative power allocation is defined as follow,

$$\max_{P_i} EE_i \quad (2)$$

Subjected to

C.1: $\gamma_{n,x} \geq \gamma_{n,x}^* \ \forall_x \in \boldsymbol{X} \ and \ \gamma_d \geq \gamma_d^* \ \forall d_R \in \boldsymbol{D}_R$

C.2: $\gamma_k \geq \gamma_k^* \ \forall k \in \boldsymbol{K}$

C.3: $\sum_{n\in N} \upsilon_n^x = 1 \ \forall k \in \boldsymbol{X}$

C.4: $\sum_{n\in X} \upsilon_n^x = k \ \forall n \in \boldsymbol{N}$

The restraint C.1 is the assurance by the SINR of the $d$th D2D and $x$th SUE which is the thresholds $\gamma_{n,x}^*$ and $\gamma_d^*$ respectively. C.2 is the restraint to the SINR of the MUEs is above to the threshold $\gamma_k^*$ and added the SINR of the macro tier in MUE as follows,

$$\gamma_k = \frac{P_m, G_{m,k}}{\sum_{n\in N} P_n G_{n,k} + \sum_{d_T \in \boldsymbol{D}_T} P_d G_{d_{T,k}} + \sigma} \quad (3)$$

where $G_{m,k}, G_{d_{T,k}}$ and $G_{n,k}$ is the power gains of the $k$th MUE and macro BS, secondary BSs $n$ and D2D transmitters $d_T$ respectively.C.2 is promoted by the expectation of the macro BS and can be interchange its correlated MUEs SINR orientation of secondary BSs and D2D transmitters. The restrain C.3 suggests the correlation with each SUE on the only one secondary BS and C.4 accentuate can each secondary BS assist the $k$ SUEs.

## III. RESEARCH IN POWER ALLOCATION SCHEME

In this section, we authorize the power allocation in ST model by applying online learning.

**3.1.Online Learning Structure**

For different network state, can be optimize the power allocation strategy for every ST on the role of a learning servant. The online research scheme are as follows:

**Action**: the action is the ST transmission $(a_i) = (P_i)$.

**State** are the local consideration to assign their background state for the time slot $t$. The state checked ST $i$ is assign as,

$$s_i^t = (i\ , P_i) \quad (4)$$

**Reward**: the reward action assign the state/action group, to accomplish the desired efficiency as well as to return the selection of power action /state to required transmission QoS.

**Transition Function**: the state $s_i^t$ to $s_i^{t+1}$on the transition is resolved by the ST conjectural action. Any ST chose the approach $\pi_i(s_i)$ to conventional reward. The approach $\pi_i(s_i)$ is assign the a probability vector $\pi_i(s_i) = [\pi_i(s_i, 0), \ldots\ldots, \pi_i(s_i, P_i^{max})]$ where $\pi_i(s_i, a_i)$ selects action $a_i$at the state $s_i$.

For the other STs approach about entire information, total conventional reward of is defined as follows,

$$V_i(s_{i,}\pi_i, \pi_{-1}) = E\left[\sum_{t=0}^{\infty} \beta^t R_i(s_i^t, \pi_i(s_i^t), \pi_{-1}(s_i^t)), s_i^0 = s_i\right]$$
$$= E\left[R_i\left(s_i, \pi_i(s_i), \pi_{-1}(s_i)\right)\right] + \beta \sum_{s_i^t \epsilon s_i} T_{s_i, s_i^t}\left(\pi_i(s_i), \pi_{-1}(s_i)\right) V_i(s_i^t, \pi_i, \pi_{-1}) \quad (5)$$

where $T_{s_i, s_i^t}(\,.\,)$ is transition probability,

$$E\big[R_i\big(s_i,\pi_i(s_i),\pi_{-1}(s_i)\big)\big]=\sum_{(a_i,a_{-1})\in A}[R_i(s_i,a_i,a_{-1})\prod_{j\in N}\pi_j(s_j,a_j)] \quad (6)$$

where $a_{-i}$ STs represent the selected action for other state $s_i$. In the conditional research, learn the optimal power allocation for each STs $\pi_i^*$ for environment state $s_i$.

The optimal strategy satisfies the Bellman's optimality equation, that is, for ST $i$

$$V_i(s_i,\pi_{-i}^*,\pi_{-i}^*)=max_{a_{i\in A_i}}\{E\big[R_i\big(s_i,a_i,\pi_{-i}^*(s_i)\big)\big]+\beta\sum_{s_i'\in s_i}T_{s_i,s_i'}(a_i,\pi_{-i}^*(s_i))V_i(s_i',\pi_i^*,\pi_{-i}^*)\} \quad (7)$$

Where $E\big[R_i\big(s_i,a_i,\pi_{-i}^*(s_i)\big)\big]=\sum_{(a_i,a_{-1})\in A}[R_i(s_i,a_i,a_{-1})\prod_{j\in N}\pi_j(s_j,a_j)]$

Hence, evaluate Q-value of ST in conventional reward plus discounted reward when these optimal approach for all STs are as follows,

$$Q_i^*(s_i,a_i)=E\big[R_i\big(s_i,a_i,\pi_{-i}^*(s_i)\big)\big]+\beta\sum_{s_i'\in s_i}T_{s_i,s_i'(a_i,}\pi_{-i}^*(s_i)V_i(s_i',\pi_i^*,\pi_{-i}^*) \quad (8)$$

By bringing together (7) and (8),

$$Q_i^*(s_i,a_i)=E\big[R_i\big(s_i,a_i,\pi_{-i}^*(s_i)\big)\big]+\beta\sum_{s_i'\in s_i}T_{s_i,s_i'}(a_i,\pi_{-i}^*(s_i))\ max_{a_{i\in A_i}}Q_i^*(s_i,b_i) \quad (9)$$

The active online research scheme of optimal Q-value assign the recursive way $(a_{i,}s_i,s_i',\pi_i^t)$ by the two states $s_i=s_i^t$and$s_i'=s_i^{t+1}$ with the time slot t and t + 1 respectively,$a_i$ and $\pi_i^t$ are the ST of end of time slot t and the power allocation approach in time slot t respectively. the online learning in optimal Q-value can be spread is given by,

$$Q_i^{t+1}\big(s_{i,}a_i\big)=(1-\alpha^t)Q_i^t(s_i,a_i)+\alpha^t\left\{\sum_{\alpha_i\epsilon A_{-i}}[R_i(s_i,a_{i,}a_{-i}\right)\times\prod_{j\epsilon N}\pi_j^t(s_{j,}a_j)]+\beta max_{b_i\epsilon A_i}Q_i^t\left(s_i',b_i\right) \quad (10)$$

**3.2Challenging Issues In System Design**

This paper is active power allocation scheme for different tier communicator, in the online learning structure. The reward in every ST is requires other STs information about a function of joint actions. This produce a problem to the following reasons:

- Each ST not be the number of other ST in the system.
- Every ST is access its background state, received historical reward or its transmission strategy.
- Therefore, the common online learning methodology for each state/action pair to maintain the large capacity in Q-value of 5G heterogeneous system, which has the system concurrence and the curse of the required computations in extensity increases.

## IV.ESTIMATED-INTUITION IN ONLINE RESEARCH

In this section, STs power allocation approaches in 5G heterogeneous network and the power allocation problem. In such environment is considered state/ action Q-value are due to the large space of convergence. Thus, we introduce estimated – intuition in online learning scheme, which allows STs power allocation for without exchange the ST information to other strategies. A smaller set of Q-value function will be uses a brief for approximated representation. This will reduce the algorithm related computations and advance the convergence.

**4.1. Intuition Based Power Allocation**

At the same work state for the different STs will be similar to the approach in power allocation by the derived intuition. To be ready to evaluate other STs power allocation strategies $\pi_{-i}^t(s_i)=(\pi_i^t(s_1)\ldots,\pi_{i-1}^t(s_{i-1}),\pi_{i+i}^t(s_{i+1}),\ldots,\pi_{N+D}^t(s_{N+D}))$ applying non-cooperative research scheme,

$$\mu_i^t\big(s_{i,}a_{-i}\big)=\prod_{j\in N\cup\frac{D}{\{i\}}}\pi_j^t\big(s_j,a_j\big) \quad (11)$$

ST $i$ is the time $t$. Calculate the Q- value $Q_i^{t+1}(s_i,a_i)$ of next time slot $t+1$ by the other STs to the result. ST $i$ is define background state $s^{t+1}$ as well as achieves probability in ST $i$ is reward $R_i(s_i,a_i,a_{-i})$,

$$T_i=\pi_i^t(s_i,a_i)\mu_i^t(s_i,a_{-i}) \quad (12)$$

Let $\delta$ is the continue time slot. According to its strategy is used to transmission power for the available to ST $i$, the method explicit the power allocation approach in intuition factor are as follow,

$$\mu_i^t(s_i,a_{-i})=\mu_i'(s_i,a_{-i})+\omega_i\big[\pi_i^t(s_i,a_i)-\pi_i'(s_i,a_{-i})\big] \quad (13)$$

where $\mu_i^t(s_i, a_{-i})$ and $\pi_i^{'}(s_i, a_{-i})$ are the intuition based action for the mention points and probability, and $\omega$ is a linearization of positive scalar. The reference points are given to common knowledge. They calculate with other ST $i$ and it derived from its reference point $\pi_i^{'}(s_i, a_{-i})$ and $\mu_i^{'}(s_i, a_{-i})$ is the proportional to $[\pi_i^t(s_i, a_i) - \pi_i^{'}(s_i, a_{-i})]$. If the reference point are $\mu_i^{'}(s_i, a_{-i}) = \prod_{j\in N\cup D/\{i\}} \pi_j^*(s_j, a_j)$ and $\pi_i^{'}(s_i, a_i) = \pi_i^*(s_i, a_i)$, then the intuition factor is optimized to $\mu_i^{'}(s_i, a_{-i}) = \prod_{j\in N\cup D/\{i\}} \pi_j^*(s_j, a_j)$. The achieved maximum Q-value on their historical local information are refer to the based points about transmissions of the STs. reference points in time slot $t$ define the following rule on the intuition factor of STs,

$$\mu_i^t(s_i, a_{-i}) = \mu_i^{t-1}(s_i, a_{-i}) + \omega_i[\pi_i^t(s_i, a_i) - \pi_i^{t-1}(s_i, a_{-i})] \tag{14}$$

where $\mu_i^{'}(s_i, a_{-i})$ and $\pi_i^{'}(s_i, a_i)$ are set to $\mu_i^{t-1}(s_i, a_{-i})$ and $\pi_i^{t-1}(s_i, a_{-i})$ respectively. Each ST perform changes will other STs to induce the next time slot. We consider the its reference points of ST $i$ from deviation to catch of other STs as follows,

$$\mu_i^t(s_i, a_{-i}) - \mu_i^{t-1}(s_i, a_{-i}) = \omega_i[\pi_i^t(s_i, a_i) - \pi_i^{t-1}(s_i, a_{-i})] \tag{15}$$

Balancing analysis and bleeding is conditional issue in power allocation. Analysis new allocated approaches and it does not assign new ones. Well-established strategy is using exploitation process. Assign the $\varepsilon$-greedy technique for balance the selection. However, the available approach selects actions equally. In order to beaten the $\varepsilon$ -greedy drawback of approach, the Q-value are varied as a graded function of the action selection probabilities. Other levels are ranked is given the highest selection probability the best power level according to their Q-values. Determine the energy efficiency of the power allocation probability action that compete the maximization to the learning algorithm exploits Boltzmann probability distribution.

### 4.2. Power Allocation Algorithm Concurrence

Prove the proposed intuition-based online learning algorithm for power allocation concurrence. Our proof calculate ordinary differential equations (ODE). The following assumption required to the proof:

**Algorithm 1** Approximate intuition based online learning algorithm for power allocation

**Require:** $\pi_i^t(s_i, a_i), t, \omega_i > 0, \gamma_{n,x}^*, \gamma_d^*$ and $\gamma_k^*$
**Ensure:** Transmission power allocation for STs

```
initialization
Let t = 0
for each (s_i, a_i, ∈ A_i) do
    initialize power allocation strategy π_i^t(s_i, a_i);
    initialize approximated Q- value ξ_i^t ψ^T(s_i, a_i);
    initialize intuition factor μ_i^t(s_i, a_-i);
end for
evaluate the state s_i = s_i^t
while (true) do
    select action a_i according to π_i^t(s_i, a_i);
    Measured the received γ_n,x, γ_d and γ_k with feedback from the receiver and observed the states s_i' by identifying P_i and
    comparing SINR;
    if (γ_n,x ≥ γ*_n,x, γ_d ≥ γ*_d and γ_k ≥ γ*_k) then
        R_i(s_i, a_i, a_-i) is achieved ;
    else
        R_i(s_i, a_i, a_-i = 0) as the receiver could not received the data correctly
    end if
    Update ξ_i^{t+1} ψ^T(s_i, a_i) based on μ_i^t(s_i, a_-i)
    Update π_i^{t+1}(s_i, a_i)
    Update μ_i^{t+1}(s_i, a_i)
    s_i = s_i^{t+1}
    t = t+1
end while
```

**Assumption 1.** The basic actions $\psi_z(s_i, a_i)$ are independent for all $(s_i, a_i)$and $Q_i^t(s_i, a_i)$ are applied to the dot product of vector $\xi_i^t\psi^T(s_i, a_i)$.
**Assumption 2.**Each z = (1,2,….Z), $\psi_z(s_i, a_i)$ is bounded, which means $E\{\psi_z^2(s_i, a_i)\} < \infty$ and the reward function supplied $E\{R_z^2(s_i, a_i, a_{-i})\} < \infty$.
**Assumption 3.** The research rate supplied $\sum_{t=1}^{\infty} \alpha^t = \infty$ and $\sum_{t=1}^{\infty} (\alpha^t) < \infty$.

**Definition 1.** Let $\Psi = E[\psi^T(s_i, a_i)\psi(s_i, a_i)]$. For the vector $\xi$ and a network state $s^i \in S'$, we assign a $\psi(s_i, \xi) = [\psi_z(s_i, a_i)]$ for z=1 $\rightarrow Z$ where $a_i \in \{a_i i = argmax_{b_{i\in}A_i} \xi_i \psi^T(s_i, a_i)\}$ is power allocation optimization action for $s_i$.

**Proposition 1.**the assumption 1-3 and Definition 1, calculate the intuition in online learning with coverage of approximation probability (w.p)1, if

$$\psi' < \Psi, \forall \xi$$

Proof: The proof of concurrence is to finding the ODE in stable fixed points which is define on the basis of the derivation of the rule with respect to $t$ as follows,

$$\xi_i^t = E[(\sum_{a_{-i}\in A_{-i}}[\mu_i^t(s_i, a_i) - \mu_i^{t-1}(s_i, a_i)]R_i(s_i, a_i, a_{-i}) + \beta\xi^t\psi^T(s_i, a_i) - \xi^t\psi^T(s_i, a_i))\psi(s_i, a_i)] \quad (16)$$

Where $\xi_i^t = \frac{\partial \xi}{\partial t}$ as $\alpha \rightarrow 0$. From the definition of the intuition factor , we state

$$\sum_{a_{-i}\in A_{-i}}[[\mu_i^t(s_i, a_i) - \mu_i^{t-1}(s_i, a_i)]R_i(s_i, a_i, a_{-i})]] = \sum_{a_{-i}\in A_{-i}} \omega_i[\pi_i(s_i, a_i) - \pi_i'(s_i, a_i)]R_i(s_i, a_i, a_{-i}) \quad (17)$$

By substituting the value of $\pi_i(s_i, a_i)$ and when $T$ is large , we obtain,

$$e^{\xi_i^t \psi \frac{T(s_i, a_i)}{T}} = 1 + \frac{\xi_i \psi^T(s_i, a_i)}{T} + \rho\left(\frac{\xi_i \psi^T(s_i, a_i)}{T}\right)$$

Where $\rho\left(\frac{\xi_i \psi^T(s_i, a_i)}{T}\right)$ is a polynomial of order O($\left(\frac{\xi_i \psi^T(s_i, a_i)}{T}\right)$), we can find,

$$\sum_{b\in A_i} e^{\xi_i^t \psi \frac{T(s_i, a_i)}{T}} = m_i + 1 + [\frac{\xi_i \psi^T(s_i, a_i)}{T} + \rho\left(\frac{\xi_i \psi^T(s_i, a_i)}{T}\right)]$$

Where $m_i$ is the range certain number of power levels.

$$\pi_i(s_i, a_i) = \frac{1}{m_i + 1} + \frac{1}{m_i + 1} \cdot \frac{\xi_i \psi^T(s_i, a_i)}{T} + \varrho\left(\frac{\xi_i \psi^T(s_i, a_i)}{T}\right) \quad (18)$$

Where $\varrho\left(\frac{\xi_i \psi^T(s_i, a_i)}{T}\right)$ is a polynomial of order smaller than O$\left(\frac{\xi_i \psi^T(s_i, a_i)}{T}\right)$, Note that the Q-value of the polynomial is independent of the coefficient vector. Evaluate the reference approach to the optimization of historical action as follows,

$$\pi_i'(s_i, a_i) = \frac{1}{m_i + 1} + \frac{1}{m_i + 1} \cdot \frac{\xi_i \psi^T(s_i, a_i)}{T} + \varrho\left(\frac{\xi_i \psi^T(s_i, a_i)}{T}\right) \quad (19)$$

By substitution (18) and (19) in (17)

If assume a value of T, then

$$\sum_{a_{-i}\in A_{-i}}[[\mu_i^t(s_i, a_i) - \mu_i^{t-1}(s_i, a_i)]R_i(s_i, a_i, a_{-i})]] \leq \frac{1+\beta}{m_i + 1}[\xi_i \psi^T(s_i, a_i) - \xi_i \psi^T(s_i, \xi)] \quad (20)$$

W can assume that V=$\frac{1+\beta}{m_i + 1}$. Now,

$$\xi^t = E[(V\xi^t\psi^T(s_i, a_i) - \xi^t\psi^T(s_i, \xi^t)] + \beta\xi^t\psi^T(s_i', \xi^t) - \xi^t\psi^T(s_i, a_i)\psi(s_i, a_i)]$$

If $\xi_0^t = \xi_1^t - \xi_2^t$, then

$$\frac{\delta\|\xi_0^t\|^2}{\delta t} = 2(\xi_1^t - \xi_2^t)(\xi_0^t)^T$$
$$= E[(-2V\xi_1^t\psi^T(s_i', \xi_1^t) + 2\beta\xi_1^t\psi^T(s_i', \xi_1^t)\psi^T(s_i, a_i)(\xi_0^t)^T - 2V\xi_2^t\psi^T(s_i', \xi_2^t) + 2\beta\xi_2^t\psi^T(s_i', \xi_2^t)\psi^T(s_i, a_i)(\xi_0^t)^T$$
$$+(2V - 2)\xi_0^t\Psi(\xi_0^t)^T \quad (21)$$

From Definition 1, we can analyze the following two asperities,

$$\xi_1^t\psi^T(s_i', \xi_1^t) \leq \xi_1^t\psi^T(s_i', \xi_2^t) \quad (22)$$

$$\xi_2^t\psi^T(s_i', \xi_1^t) \leq \xi_2^t\psi^T(s_i', \xi_2^t) \quad (23)$$

As s result, the optimization of intuition in estimated online Q-factor,

$$Q'(s_i, a_i, \xi^*)\xi^*\psi(s_i, a_i) \quad (24)$$

## V. Results

In the simulation, calculate the complete system performance of energy efficiency (EE). The system behaves predicted action on concurrence for each scheme by showing the number of epoch as in Figure 2. The system energy efficiency is evaluated by the contact on the number of STs in Figure 3.

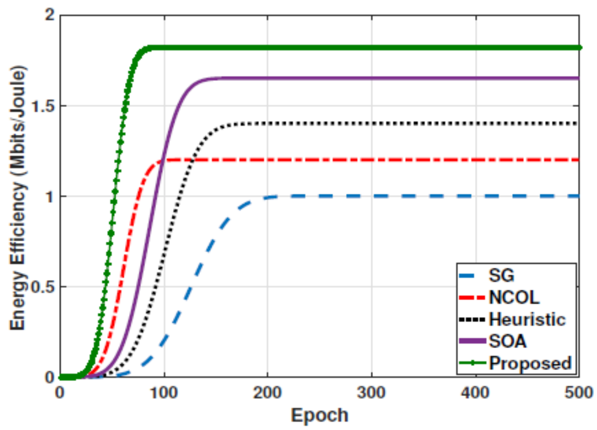


Figure 2. Average on energy efficiency on various system

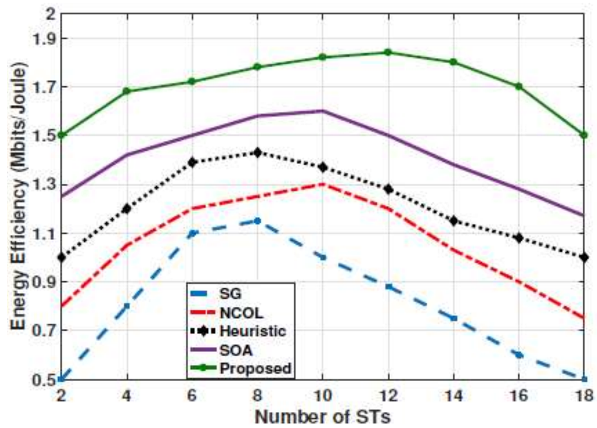


Figure 3. System energy efficiency of the contact of the STs

This simulation calculates the concurrence speed as in Figure 4, the whole 5G system presents the average spectral efficiency. Figure 5 present the macro receivers respectively the average SINR measured for all the schemes. Our scheme continue SINR above threshold for both receivers and achieved higher SINR than others. The proposed scheme calculate the QoS for both macro and secondary tier users for the satisfaction of constraints C.1 and C.2.

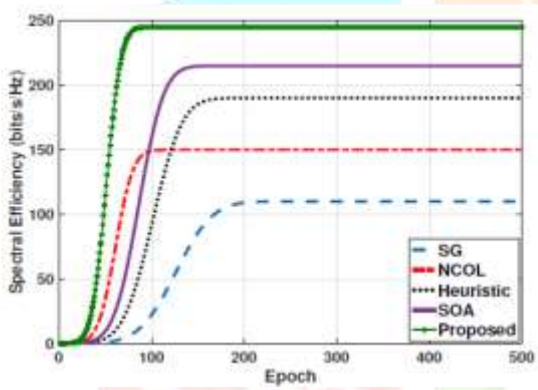


Figure 4 Average system spectral efficiency

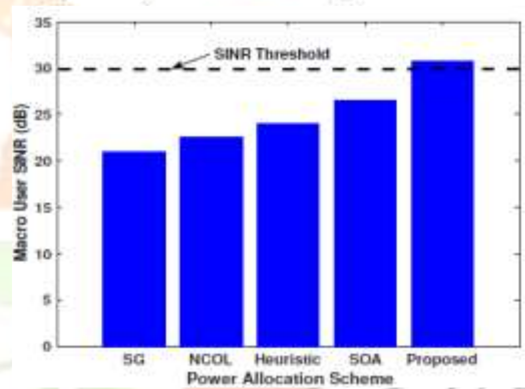


Figure 5. Average SINR for macro users

Figure 6 D2D receiver of UEs by achieved SINR and it handled the minimum SINR required by high number of UEs.

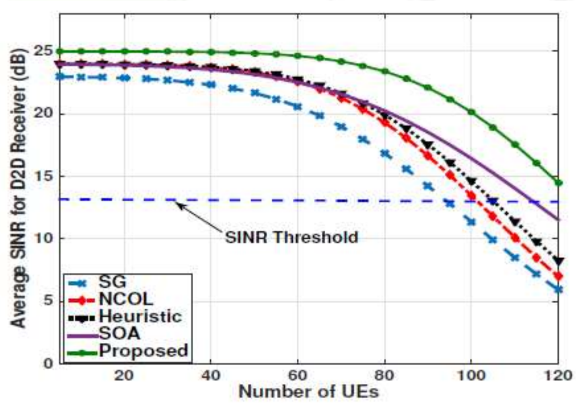


Figure 6. Average SINR for different UEs in the network

**VI. CONCLUSION**

This paper, 5G environment is handle a spectrum sharing for the downlink transmission in power allocation problem, where D2D and small cells access the spectrum. We consider the online research based scheme for maintaining QoS for both primary tier and secondary while tier reduce power consumption. Therefore, expedites the speed of convergence and reduces the state/action space for the Q-value of dissimilar mechanism. An intuition based approach selected transmission power of other STs by the online learning. In addition, the Q-value an approximation mechanism for employs the scheme , which expedites the speed of convergence and reduces the state/action space. Perform the higher energy efficiency was determined and compared to other schemes through simulation of the proposed scheme where and it conclude faster convergence.

**REFERENCES**


[1] R. Q. Hu and Y. Qian, “An energy efficient and spectrum efficient wireless heterogeneous network framework for 5g systems”, IEEE Communications Magazine, vol. 52, no. 5, pp. 94–101, May 2014.

[2] M. Agiwal, A. Roy, and N. Saxena, “Next generation 5g wireless networks: A comprehensive survey”, IEEE Communications Surveys Tutorials,

[3] Woon Hau Chin, Zhong Fan, and R. Haines, “Emerging technologies and research challenges for 5g wireless networks”, IEEE Wireless Communications,, vol. 21, no. 2, pp. 106–112, April 2014.

[4] T. O. Olwal, K. Djouani, and A. M. Kurien, “A survey of resource management toward 5g radio access networks”, IEEE Communications Surveys Tutorials, vol. 18, no. 3, pp. 1656–1686, third quarter 2016.

[5] I. AlQerm and B. Shihada, “A cooperative online learning scheme for resource allocation in 5g systems”, in 2016 IEEE International Conference on Communications (ICC), May 2016.

[6] V. Chandrasekhar, J.G. Andrews, Tarik Muharemovict, Zukang Shen, and Alan Gatherer, “Power control in two-tier femtocell networks”, IEEE Transactions on Wireless Communications,, vol. 8, no. 8, pp. 4316–4328, August 2009.

[7] M. Haddad, P. Wiecek, O. Habachi, and Y. Hayel, “A game theoretic analysis for energy efficient heterogeneous networks”, in 12th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt),, May 2014, pp. 241–246.

[8] H. Munir, S. A. Hassan, H. Pervaiz, and Q. Ni, “A game theoretical network-assisted user-centric design for resource allocation in 5g heterogeneous networks”, in IEEE 83rd Vehicular Technology Conference(VTC Spring), May 2016.

[9] S. Navaratnarajah, A. Saeed, M. Dianati, and M. A. Imran, “Energy efficiency in heterogeneous wireless access networks”, IEEE Wireless Communications, vol. 20, no. 5, pp. 37–43, October 2013.

[10] C. Niu, Y. Li, R. Q. Hu, and F. Ye, “Fast and efficient radio resource allocation in dynamic ultra-dense heterogeneous networks”, IEEE Access, vol. 5, pp. 1911–1924, 2017.