

Adaptive Rate Allocation Technique for Mixed Traffic Users in Communication Network

Neeraj Kumar^{1*}, Shashank Awasthi², Anwar Ahmad³

¹PhD student of the Department of Electronics and Telecommunication, Jamia Millia Islamia University, New Delhi.

²Department of Computer Science & Engineering, G. L. Bajaj Institute of Technology & Management, Greater Noida, G.B. Nagar, Uttar Pradesh, India

³Professor of the Department of Electronics and Telecommunication, Jamia Millia Islamia University, New Delhi.

*Corresponding author E-mail: Neeraj.mohiwal@gmail.com

Abstract

In this paper, a network utility maximization based rate allocation technique is proposed for mixed traffic users. Each user is running a network application type of either elastic traffic or inelastic traffic. A utility function is defined for each elastic and inelastic traffic behaving application. Network utility maximization approach maximizes utilities in given range of maximum throughput rate constraint. Proposed technique adaptively allocates rate to users as increasing in range of maximum rate constraint. It proposes dynamic change in parameters of utility functions with respect to change in maximum rate constraint. Results of the proposed rate allocation technique is compared with existing techniques, which shows improved throughput rate for users at any value of maximum rate constraint.

Keywords: Utility function, proportional fairness, price, demand, adaptive rate, allocation, throughput, logarithmic, sigmoidal.

1. Introduction

RAPID development has been seen in communication technology in last decade. Started from analog communication (1G) technology to complete packet switching (4G-LTE) technology is a major technology shift in a communication field. Normal voice communication service to demand of QoS based high data rate services drives this technology shift. Radio resource scheduling is always remained one of the key areas in each telecommunication network generation to satisfy higher throughput demand of users. Therefore, number of research works has been done in this area. A resource scheduling technique should be design such that it improves overall communication performance i.e. data rate, fairness, delay, packet loss rate etc. For resource scheduling, different key parameters are utilized such as channel condition, packet queue length, head of the line delay, number of carriers, past throughput etc. Some of the key parameters are time variable (for example channel condition, queue length/queue aware, delay, past throughput) and some of the parameters are non-time variable (for example number of carriers.). An adaptive resource scheduling technique is beneficial as it adapts resource allocation in accordance to the changes in key parameters and to have reliable communication performance. Literature survey papers [1]-[3] suggest different resource scheduling techniques, such as round robin, maximum throughput, proportional fairness, delay based scheduling etc. Survey paper [3] discloses scheduling techniques for uplink communication in LTE. These techniques are broadly classified in to channel unaware strategies and channel aware strategies, and QoS unaware strategies and QoS aware strategies. Each strategy provides a priority scheduling metric and accordingly resources are scheduled. Priority scheduling metric prioritizes users as per values of key scheduling parameters. Research paper [4] shows priority scheduling metrics

for round robin and proportional fair scheduling techniques and then compares its performance results. It suggests that proportional fairness technique provides high fairness to users as it allocates resources based on past throughput and current channel condition. Proportional fair technique is a technique based on trade-off between the fairness and data rate is suggested in [5]. Exponential proportional fairness technique is disclosed in [6], which utilizes features of exponential function for proportional fairness scheduling. Most of these techniques do not consider traffic type of network applications running at user ends for resource scheduling. Research papers [7] -[8] suggest resource allocation techniques for mixed traffic i.e. real time traffic and non-real time traffic. Research paper [7] suggest dynamic packet scheduling in which services are categorized in different classes and then priority based scheduling algorithm is utilized for resource allocation. Research paper [8] discusses Hebbian learning process and K-mean clustering algorithm for scheduling after categorizing services in different classes. Different research papers [9]-[12] suggest resource allocation technique based on utility function maximization. Research paper [9] suggests a utility-based resource scheduling technique for different types of traffic. It suggests a heuristic algorithm for resource scheduling for hybrid traffic users in which real time traffic users get required QoS, whereas non real time traffic users has to deal with tradeoff between data rate and fairness. Similarly, research papers [11]-[12] proposed utility proportional fairness based rate allocation method for mixed traffic users. Different utility functions (i.e. logarithmic and sigmoid function) are utilized for elastic traffic users and inelastic traffic users. It allocates rates to all users such that no users have zero rate allocation. Research papers [11]-[12] discuss that user bids for resources in iterative manner. These bids are iteratively changes as per changing in price per unit bandwidth, which is changed in accordance with utility maximization algorithm. Techniques disclosed in research papers [9] -[12]

provide utility proportional fairness in limited range of maximum rate i.e. less number of resources. As range of maximum rate is increased, than rate is dis-proportionally allocated to users. It means that it allocates higher rate to elastic traffic corresponding users while marginally increasing rates allocation for inelastic traffic corresponding users at higher range of maximum rate. Such allocation reduces overall communication performance. In telecommunication network, it is desirable that inelastic traffic (real time) performing users should get higher priority than that of elastic traffic (non-real time) performing users. In this paper, a rate allocation technique is proposed for mixed traffic users at any range of maximum throughput rate. Proposed technique adaptively prioritizes inelastic traffic (real time) running users than elastic traffic (non-real time) running users as range of maximum throughput rate is increasing. It provides optimum throughput to all users based on utility proportional fairness concept as suggested in [11] or [12] and it guarantees at least minimum throughput for all users. Proposed technique utilizes logarithmic utility function for elastic traffic and sigmoidal-like utility function for inelastic traffic. Current section is focused on introduction and related works. Next, section 2 is focused on different types of utility function for elastic traffic and inelastic traffic users. Section 3 is focused on scheduling design formulation and solution based on proposed scheduling technique. In section 4, simulations results of proposed scheduling technique and existing scheduling techniques are discussed. At last, conclusion is made for proposed scheduling technique.

2. Utility Function

In field of the economics, a utility is defined as a measurement of preferences over a set of goods and services. Its value indicates a degree of user satisfaction from the goods and services [13]. Similarly, in communication network, utility indicates a degree of satisfaction of user in terms of received rate and experienced QoS. These users can be performing of any type of communication services i.e. elastic traffic services and inelastic traffic services. As name suggest, elastic traffic are delay tolerant services and adaptable with variable throughput rate whereas inelastic traffic are rigid with respect to delay and throughput. Inelastic traffic services perform poor when it experience lower QoS. Practical examples for the elastic traffic are non-real time services such as email communication, messaging, web browsing etc. and for the inelastic traffic are real time services, such as voice calling, video calling, streaming etc. These traffic types can be modeled in different utility functions. Research paper [16] shows different utility curves for elastic traffic and for inelastic traffic. Research papers [9] and [10] reveal different utility functions for elastic traffic and inelastic traffic. Research papers [11] and [12] also disclose utility functions for elastic traffic and inelastic traffic. It shows logarithmic function for inelastic traffic and sigmoidal like elastic traffic. To satisfy users in the network, network utility maximization (NUM) approach is suggested in research paper [17]. Research paper [17] has a major contribution in all devised techniques related to network utility maximization (NUM) approach. Proposed technique also utilizes network utility maximization (NUM) approach for throughput rate allocation to users. In this paper, utility functions disclose in [9] - [12] are utilized for throughput rate allocation.

- 1) Utility functions for elastic traffic are:
 - For HTTP application type

$$U_n(x_n) = \frac{\log\left(\frac{x_n}{x_{min}}\right)}{\log\left(\frac{x_{max}}{x_{min}}\right)} \quad (1)$$

- For FTP application type

$$U_n(x_n) = \frac{\log\left(\frac{\tau_n x_n + 1}{\tau_n x_{max} + 1}\right)}{\log\left(\frac{x_{max}}{x_{min}}\right)} \quad (2)$$

- 2) Utility functions for inelastic traffic are:
 - For streaming application type

$$U_n(x_n) = \frac{1}{1 + e^{-\alpha_n(x_n - \beta_n)}} \quad (3)$$

- For voice/video communication application type

$$U_n(x_n) = \gamma_n \left(\frac{1}{1 + e^{-\alpha_n(x_n - \beta_n)}} - \delta_n \right) \quad (4)$$

Where $\gamma = \frac{1 + e^{\alpha_n \beta_n}}{e^{\alpha_n \beta_n}}$ and $\delta = \frac{1}{1 + e^{\alpha_n \beta_n}}$

All these utility functions have following features.

- $U_n = 0$ and $U_n(x_n)$ is an increasing function of x_n
- $U_n(x_n)$ is twice continuously differentiable in x_n

In fig.1, plot of logarithmic utility function (parameter τ) and sigmoidal like utility function (parameters α, β) are displayed with respect to increasing in x_n .

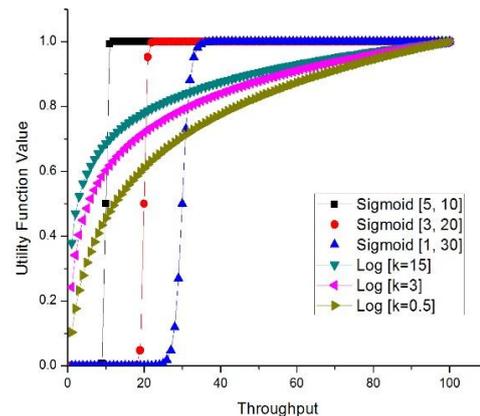


Fig. 1: Plots of logarithmic utility function and sigmoidal-like utility function at increasing in x_n

3. Scheduling Design Formulation

This section discusses rate allocation design formulation and optimization solution based on maximizing utilization functions of users. Proposed technique provides optimum throughput rate (i.e. the QoS satisfactory rate) to inelastic traffic performing users while providing minimum throughput rate to elastic traffic performing users.

3.1. Network Utility Maximization Problem

Parameters of network utility maximization problem are defined as below: let a set $\phi = \{n: n = 1, 2, 3, \dots, N\}$ of users are situated in a coverage area of a base station B_s . Each user running either elastic traffic behaving network applications or inelastic traffic behaving network applications. Each user in network of B_s receives a throughput x_{mn} over a set $\Omega = \{m: m = 1, 2, 3, \dots, M\}$ of carriers. Total throughput is received by a user is x_n over set of carriers allocated to it, which is calculated as per equ.5.

$$x_n = \sum_m x_{mn} \tag{5}$$

Network applications running at each user n can be represented by a utility function $U_n(x_n)$. Utility function $U_n(x_n)$ of a user is shown in equ.1 to equ.4 as per type of network application i.e. elastic traffic or inelastic traffic is currently running at user end.

Objective: The objective of the proposed scheduling technique is to increase overall sum of utilities of users, which is defined by equ.6, subject to maximum throughput T_{max} constraints. The network utility maximization problem is given as below:

$$\max_x \prod_{n=1}^{n=N} (U_n(x_n)) \tag{6}$$

Subject to

$$\sum_{n=1}^{n=N} x_n \leq T_{max} \tag{7}$$

Global optimal solution: Current problem shown in equ.6 can be rewritten as equ.8 by converting multiplication of utility function in to summation of natural logarithm of utility function.

$$\max_x \sum_{n=1}^{n=N} \log(U_n(x_n)) \tag{8}$$

subject to

$$\sum_{n=1}^{n=N} x_n \leq T_{max} \tag{9}$$

For global optimal solution, a Lagrange multiplier (Λ) based optimization solution is utilized, where Λ is greater than 0. The penalty function based on Lagrange multiplier (Λ) is defined for the equ.8 as in equ.10:

$$L(x, \Lambda) = \sum_{n=1}^{n=N} \log(U_n(x_n)) - \Lambda \left(\sum_{n=1}^{n=N} x_n + z - T_{max} \right) \tag{10}$$

$$L(x, \Lambda) = \sum_{n=1}^{n=N} (\log(U_n(x_n)) - \Lambda x_n) + \Lambda(T_{max} - z) \tag{11}$$

$$L(x, \Lambda) = \sum_{n=1}^{n=N} L_n(x_n, \Lambda) + \Lambda(T_{max} - z) \tag{12}$$

Here, z is a slack variable and Λ is a Lagrange multiplier. Let Λ represents price per unit bandwidth and price demand for n^{th} user is given by Θ_n . It can be represented as $\Theta_n = \Lambda x_n$ and $\sum_n(\Theta_n) = \Lambda \sum_n(x_n)$. In equ. 10, x_n is separable in $\sum_{n=1}^{n=N} (\log(U_n(x_n)) - \Lambda x_n)$. So, we can write $\max_x(n=1n=N(\log U_n x_n - \Lambda x_n)) = \sum_{n=1}^{n=N} \max_x(\log(U_n(x_n)) - \Lambda x_n)$. The primal problem defined in equ.10 can be solved by duality based optimization solution. Duality problem objective function can be written as equ.13:

$$D_\Lambda = \max_x L(x, \Lambda) \tag{13}$$

$$D_\Lambda = \sum_{n=1}^{n=N} \max_{x_n} (\log(U_n(x_n)) - \Lambda x_n) + \Lambda(T_{max} - z) \tag{14}$$

$$D_\Lambda = \sum_{n=1}^{n=N} \max_{x_n} (L_n(x_n, \Lambda)) + \Lambda(T_{max} - z) \tag{15}$$

The dual problem is given by

$$\min_\Lambda D_\Lambda \tag{16}$$

subject to $\Lambda \geq 0$ So, we have

$$\frac{\partial D_\Lambda}{\partial \Lambda} = T_{max} - \Lambda \sum_{n=1}^{n=N} x_n - z \tag{17}$$

$$\frac{\partial D_\Lambda}{\partial \Lambda} = 0 \tag{18}$$

$$T_{max} - \Lambda \sum_{n=1}^{n=N} x_n - z = 0 \tag{19}$$

Substituting $\sum_n(\Theta_n) = \Lambda \sum_n(x_n)$, we have

$$\Lambda = \frac{\sum_{n=1}^{n=N} \Theta_n}{T_{max} - z} \tag{21}$$

Hence the global optimization problem solution is divided in to two simpler optimization problems based on the duality. First is maximizing of $(\log(U_n(x_n)) - \Lambda x_n)$ and second is determining $\Lambda = \frac{\sum_{n=1}^{n=N} \Theta_n}{T_{max} - z}$ subject to $\Lambda > 0$ and $x_n \geq 0$.

Maximization of $(\log(U_n(x_n)) - \Lambda x_n)$ is depended on Λ , Θ_n and T_{max} . T_{max} is a constant during rate allocation, however, Λ , and Θ_n are iteratively updated to maximize it. Further, Maximization of $(\log(U_n(x_n)) - \Lambda x_n)$ is also depended on parameters of utility function. Here, to improve the rate allocation for inelastic traffic performing users, parameters α and β of utility functions given in equ.3 or equ.4 are utilized. Research paper [9] suggest that any change in value of parameter α changes the slope of utility function around β . As slope is increasing, user's demand for the rate is also increasing. Effect of the variability of parameter α on utility function is shown in fig.2.

Here, it is proposed that as value of α parameter decides slope of the utility function and thereby decides rate demand, its value should be changed in inverse proportionally with respect to the value of T_{max} . Thus, increasing in range of maximum throughput rate i.e. T_{max} will cause decreasing in parameter α as shown in equ.22 and thereby increasing in slope. So higher rate will be allocated to inelastic traffic running users than that of elastic traffic running users.

$$\alpha \propto \frac{1}{T_{max}}$$

subject to $\alpha > 1$, when $T_{max} < 100$ and $\alpha < 1$, when $T_{max} > 100$.

Proposed Scheduling Algorithm: Proposed scheduling algorithm is a distributed scheduling algorithm which will be run in iterative manner. Each user transmits its price demand to base station in each iteration and base station allocates rates to users based on received price demand. Proposed scheduling algorithm provides proportional fairness for all users in the network of base station while it satisfying QoS and minimum throughput rate for users. Proposed scheduling algorithm has two parts (a&b), one at user side and another at base station side.

- *Part a of proposed scheduling algorithm at user side:* In each iteration, each active user n transmits its price demand $\Theta_{n,k}$ for rate allocation x_n to base station B_s when user n does not receive stop command from base station. The price demand $\Theta_{n,k}$ is calculated based on $\Theta_{n,k} = \Lambda_k x_{n,k}$, where value of Λ_k is provided by base station B_s is updated in each iteration. When the user receives stop command from base station B_s , then user allocated rate $x_{n,k}$ will be calculated as per current

price demand $\Theta_{n,k}$ and Λ_k . The k is an index for number of iterations.

- *Part b of proposed scheduling algorithm at base station side:* Base station B_s receives price demand $\Theta_{n,k}$ of all users in the network in each iteration. It compares current price demand $\Theta_{n,k}$ of users to previous price demand $\Theta_{n,k-1}$ of users. If difference is less than a predetermined threshold ϵ , then scheduler at base station B_s allocates rates according to the current price demand $\Theta_{n,k}$. However, if difference is greater than determined the predetermined threshold ϵ , then scheduler at base station B_s calculates Λ_k based on received price demands from all users i.e. $\sum_n \Theta_{n,k}$ and maximum throughput T_{max} . The updated value of Λ_k is provided to all users in the network in each iteration. The k is an index for number of iterations.

4. Simulation Results and Analysis

Proposed rate allocation technique is simulated in MATLAB environment. Here, six users are taken in the network for current simulation. Out of six users, three users are running elastic traffic behaving network applications and three users are running inelastic traffic behaving network applications. These network application types are characterized by different utility functions defined in equ.1 to equ.4. Table I shows different values of parameters utilized for utility functions. Here, proposed technique is named as adaptive rate allocation (in short ARA) technique providing utility proportional fairness scheduling technique and scheduling technique of [12] of Ahmed et. al. is named as utility proportional fairness (in short PUF) scheduling technique.

4.1. Analyzing Performance Results:

Performance of proposed ARA technique at $T_{max} = 100$: Proposed ARA technique shows that it provides approx same throughput for respective users as provided by existing PUF technique. Because, it is considered that value of parameter α is above one for $T_{max} \leq 100$ as shown in equ.22. As value of α is greater than one, slope of utility function corresponding to the inelastic traffic does not increase sharply, so elastic traffic running users will get rate allocation after QoS satisfaction of inelastic traffic running users. Table II shows optimized allocated rate provided by ARA technique and PUF technique at $T_{max} = 100$. Result fig.3 shows throughput allocation to users by proposed technique and result fig.4 shows price demand of users with respect to number of iterations. Result fig.4 discloses that as price demand increases rate allocation for corresponding also increases. Performance of proposed ARA technique at $T_{max} = 150$: Result fig.5 shows throughput allocation to users as per existing PUF technique at $T_{max} = 150$. It shows that at higher value of T_{max} , maximum throughput rate is allocated to elastic traffic corresponding users while marginal throughput rate is allocated to inelastic traffic corresponding users only around b_n . Result fig.6 shows throughput allocation to users as per existing PUF technique at $T_{max} = 150$. It clearly shows that more throughput rate is allocated to inelastic traffic corresponding users in comparison to elastic traffic corresponding users. Because, it is considered that value of α parameter is less than one for $T_{max} = 100$ as shown in equ.22. As value of α is less than one, slope of utility function corresponding to the inelastic traffic increases. Thereby, elastic traffic running users will get throughput rate marginally, but inelastic traffic running users will also get rate even after QoS satisfaction of inelastic traffic running users. In fig.6 at $T_{max} = 150$, current simulation has taken particular value of α which is less than one to show effect of α on rate allocation. Thus, current scheduling technique is adaptively allocates rates as increasing in T_{max} .

5. Conclusion

Current paper discusses a rate allocation technique for mixed traffic users in relay based network. It utilizes network utility maximization approach for resource allocation to users, which are running network application either elastic traffic type or inelastic traffic type. Proposed technique treats relay node as an elastic traffic user and allocates rate to relay node when maximum rate constraint increases. Relay node provides utility proportional fairness in cell edge users. Results of the proposed rate allocation technique show improved communication performance. Current technique does not consider other scheduling parameters for resource, such as queue size, packet delay. In future, rate allocation in relay based network with considering scheduling parameters will be analyzed.

Table I

Parameters for Utility functions	
Parameter	Value
T_{max}	>60
T_{min}	0
α	Value with T_{max}
β	[10 20 30]
τ	[15 3 0.5]

Table II

Allocated throughput rates to users	
PUF technique	ARA technique
11.5	11.9
20.9	21.5
33.2	33.5
7.8	6.8
11.1	9.9
15.0	13.2

References

- [1] F. Capozzi, G. Piro, L. Grieco, G. Boggia, and P. Camarda, "Downlink packet scheduling in LTE cellular networks: Key design issues and a survey", IEEE Communications Surveys Tutorials, vol. 15, no. 2, pp. 678-700, 2013.
- [2] ayyalBendaoud, MarwenAbdennebi, Fedoua Didi, "Survey On Scheduling And Radio Resources Allocation In LTE", International Journal of Next Generation Network (IJNGN), vol. 6, March 2014.
- [3] Abu-Ali, Najah; Taha, Abd-Elhamid M.; Salah, Mohamed; Hassanein, Hossam, "Uplink Scheduling in LTE and LTE-Advanced: Tutorial, Survey and Evaluation Framework", IEEE Communications Surveys Tutorials, vol. 16, no. 3, pp.1239-1265, 2014.
- [4] Mohammad T. Kawser, Hasib M. A. B. Farid, Abduhu R. Hasin, Adil M. J. Sadik, and Ibrahim K. Razu, "Performance Comparison between Round Robin and Proportional Fair Scheduling Methods for LTE", International Journal of Information and Electronics Engineering, Vol. 2, No. 5, Sept 2012.
- [5] J.-G. Choi and S. Bahk, "Cell-throughput analysis of the proportional fair scheduler in the single-cell environment," IEEE Trans. Veh. Technol., Mar. 2007, vol. 56, no. 2, pp. 766-778.
- [6] J.-H. Rhee, J. M. Holtzman, and D. K. Kim, "Performance Analysis of the Adaptive EXP/PF Channel Scheduler in an AMC/TDM System," IEEE Communications Letters, Aug. 2004, vol. 8, pp. 4978-4980.
- [7] Jani P., Niko K., Tero H., Martti M. and Mika R., "Mixed Traffic Packet Scheduling in UTRAN Long Term Evaluation Downlink" IEEE 2008, pp 978-982.
- [8] R. Kausar and Y. Chen and K. K. Chai, "An intelligent scheduling architecture for mixed traffic in LTE-Advanced", IEEE 23rd International Symposium on Personal, Indoor and Mobile Radio Communications - (PIMRC), 2012, pp. 565-570.
- [9] Li Chen and Bin Wang and X. Chen and Xin Zhang and Dacheng Yang, "Utility-based resource allocation for mixed traffic in wireless networks" IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), 2011, pp 91-96.

[10] G. Tychogiorgos, A. Gkelias and K. K. Leung, "Utility-proportional fairness in wireless networks," 2012 IEEE 23rd International Symposium on Personal, Indoor and Mobile Radio Communications - (PIMRC), Sydney, NSW, 2012, pp. 839-844.

[11] A. Abdel-Hadi and C. Clancy, "A robust optimal rate allocation algorithm and pricing policy for hybrid traffic in 4G-LTE," IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), 2013, pp 2185-2190.

[12] A. Abdel-Hadi and C. Clancy, "A utility proportional fairness approach for resource allocation in 4G-LTE," 2014 International Conference on Computing, Networking and Communications (ICNC), Honolulu, HI, 2014, pp. 1034-1040.

[13] M. Rahman and H. Yanikomeroglu, "Enhancing cell-edge performance: a downlink dynamic interference avoidance scheme with inter-cell coordination," in IEEE Transactions on Wireless Communications, vol. 9, no. 4, pp. 1414-1425, April 2010.

[14] M. S. Alam, J. W. Mark and X. S. Shen, "Relay Selection and Resource Allocation for Multi-User Cooperative OFDMA Networks," in IEEE Transactions on Wireless Communications, vol. 12, no. 5, pp. 2193-2205, May 2013.

[15] <https://en.wikipedia.org/wiki/Utility>.

[16] Springer International Publishing Switzerland 2017 written by M. Ghorbanzadeh et al., Cellular Communications Systems in Congested Environments.

[17] F. P. Kelly, A. Maulloo, and D. Tan, "Rate control in communication networks: Shadow prices, proportional fairness and stability," Journal of the Operational Research Society, pp. 237-252, 1998.

[18] S. Boyd and L. Vandenberghe, Convex Optimization. New York, NY, USA: Cambridge University Press, 2004.

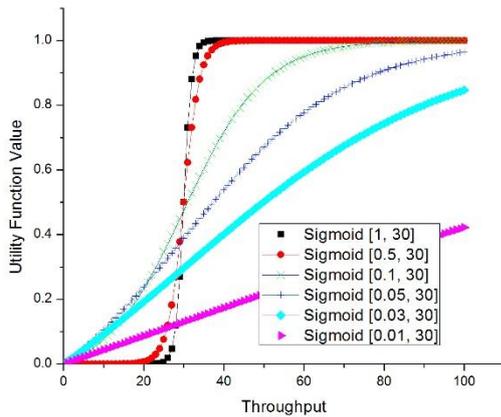


Fig. 2: Plots of sigmoidal-like utility function at increasing in x_n at different values α

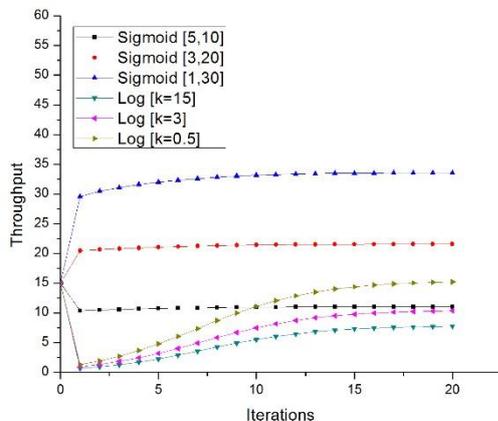


Fig. 3: Plot of throughput allocation v/s number of iterations at $T_{max} = 100$ by ARA technique

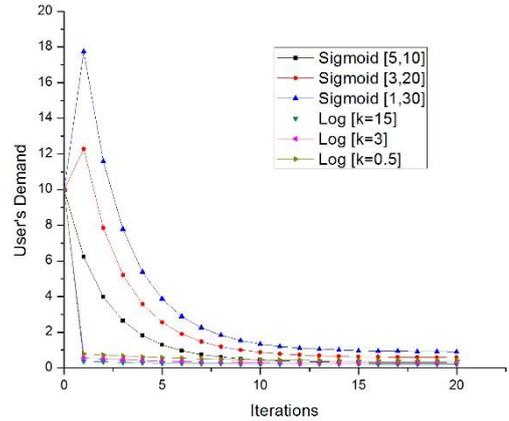


Fig. 4: Plot of price demand v/s number of iterations at $T_{max} = 100$ by ARA technique

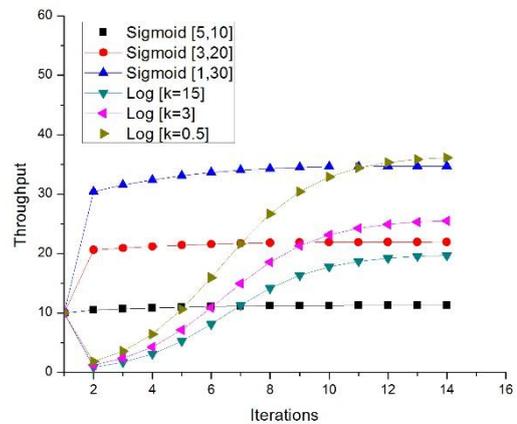


Fig. 5: Plot of throughput allocation v/s number of iterations at $T_{max} = 150$ by PUF technique

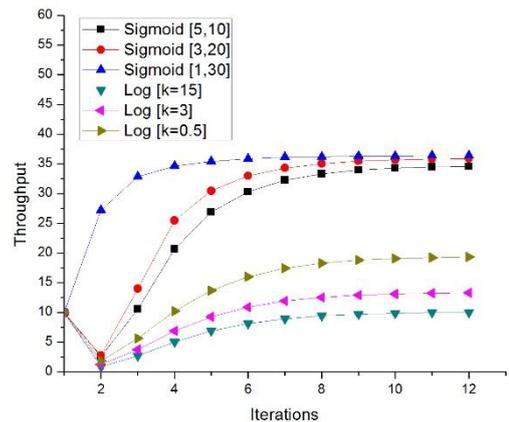


Fig. 6: Plot of throughput allocation v/s number of iterations at $T_{max} = 150$ by ARA technique