

 Open access • Proceedings Article • DOI:10.1109/JCSSE.2012.6261935

The most promising scheduling algorithm to provide guaranteed QoS to all types of traffic in multiservice 4G wireless networks — [Source link](#)

Mohsin Iftikhar, Nada Al Zaben, Suparek Manitpornsut

Institutions: King Saud University, University of the Thai Chamber of Commerce

Published on: 09 Aug 2012 - Computer Science and Software Engineering

Topics: Fair-share scheduling, Dynamic priority scheduling, Two-level scheduling, Round-robin scheduling and Rate-monotonic scheduling

Related papers:

- [An analytical G/M/1 queueing model for providing guaranteed QoS in multimedia surveillance](#)
- [Modelling and Analysis of an Integrated Scheduling Scheme with Heterogeneous LRD and SRD Traffic](#)
- [A flexible self-adaptable scheduling model for QoS multimedia provision](#)
- [Providing deterministic quality of service guarantee in a wireless environment](#)
- [Non-real-time content scheduling algorithms for wireless data networks](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/the-most-promising-scheduling-algorithm-to-provide-2t89597xlj>

The Most Promising Scheduling Algorithm to Provide Guaranteed QoS to all Types of Traffic in Multiservice 4G Wireless Networks

Mohsin Iftikhar
Computer Science Department
King Saud University
Riyadh, Kingdom of Saudi Arabia (KSA)
miftikhar@ksu.edu.sa

Suparek Manitpornsut
Department of Computer Engg. and Multimedia,
University of the Thai Chamber of Commerce,
Bangkok 10400 Thailand
suparek_man@utcc.ac.th

Abstract- Dissimilar wireless infrastructures such as WiMAX and LTE are fabulously expand in last decade. In this heterogeneous environment of wireless networks and architectures, one of the major concerns is how to allocate network resources efficiently to diverse traffic classes with different QoS constraints. Further it has been convincingly demonstrated through numerous high quality studies that multimedia traffic found in modern wireless networks exhibits Long-Range Dependence (LRD) and self-similarity, a phenomenon which cant' be captured by traditional traffic modeling based on simplistic Poisson model.

Unlike most existing studies that are primarily based on simplistic Poisson model and traditional scheduling algorithms, this research presents an analytical performance model for multiple queue systems with self-similar traffic input scheduled by a novel and promising scheduling mechanism. Our proposed model is substantiated on G/M/1 queuing system that considers multiple classes of traffic exhibiting long-range dependence and self-similar characteristics. We analyze the model on the basis of newly proposed scheduling scheme. We present closed form expressions of expected waiting times for multiple traffic classes. We develop a finite queue Markov chain for the proposed scheduling scheme. We simulated the various patterns of wireless traffic through our own designed discrete simulator. The results indicate that our proposed scheduling algorithm provides preferential treatment to real-time applications such as voice and video but not to that extent that data applications are starving for bandwidth and outperforms all scheduling schemes that are available in the market.

Keywords: QoS, LTE, WiMAX, 4G, Self-Similar

I. INTRODUCTION

In the Internet, Quality of Service (QoS) management allows different types of traffic to contend inequitably for network resources. Bandwidth is the key heuristic to manage real life network utilities like video and voice over remote locations. Mostly MPLS, IntServ and DiffSer are utilized to facilitate sensitive applications [1]. The differentiation of digital traffic is fundamentally relied on these frameworks that utilize scheduling and queuing methodologies for separating each class type. The traffic separation is categorized

under specific parameters like peak bandwidth, average bandwidth and delay. The different arrangements of these parameters can be bundled under variety of queuing and scheduling methods. It is therefore vital to QoS frameworks that modeling of traffic behaviour through network domains is accurate so that resources can be optimally assigned.

Traffic measurements have been remained a hot issues in past years. It has been noticed that packet based traffic patterns show Long-Range Dependency (LRD), hence when viewed in exact time span the patterns exhibited self-similar behavior [2-4]. Despite the recent findings of self-similarity and LRD in measured traffic from data networks, much of the current understanding of IP traffic modeling is still based on simplistic Poisson distributed traffic. In this paper, we add to a more realistic modeling of network domains through the following main contributions: (1) the presentation of an analytical approach and closed form expressions to model the accurate behaviour of multiple classes of wireless IP traffic based on a G/M/1 queuing system under self-similar assumptions, (2) the derivation of expected waiting times of corresponding self-similar traffic classes and formulation of an embedded Markov Chain (MC) and (3) the detailed simulation results to give exact QoS parameter bounds to validate the mathematical model. The current work is the extension of our prior work [28]. In [28], we have analyzed the traditional scheduling schemes based on G/M/1 queuing system, whereas in current study, we analyze the newly proposed scheduling scheme to guarantee tight bound QoS to all kind of traffic in multiservice wireless internet.

The rest of the paper is structured as follows. Section II summarizes related work. Estimation of interval times and self similar nature traffic under various classes has been discussed in Section III. Section IV explains the procedure of formulating the embedded Markov chain along with derivation of packet delays. The simulation results are given in Section V. Finally, Section VI concludes this work.

II. RELATED WORK

Queuing theory is the backbone of telecommunication systems. Major issues faced by the internet traffic are lies under this question: how burstiness can be managed in multiple time spans, commonly known as long-range dependence and fractal behavior [5]. This study actively considers the queuing theory in terms of self-similar traffic. The experimental queuing analysis and simulation studies with long-range dependent packet data traffic have been performed in [7] and [8] respectively. These studies merely indicate that providing guarantees on maximum delay, delay-jitter and cell-loss probabilities in the presence of LRD traffic is nontrivial especially if the coefficient of variation of the marginal distribution is large. This is because such traffic significantly increases queue length statistics at a multiplexer. The readers are referred to [9-13] for an overview of other queuing based results available in the presence of self-similar traffic. The core discrepancy of current queuing schemes is the implementation of FIFO logic.

The desire to provide different QoS guarantees to different classes of customers in wireless Internet is leading to the use of priorities in the allocation of resources. Multiple priority based classes are supported by the IP routers and ATM switches. Researchers have begun to examine the impact of non-FIFO queuing with self-similar input. The authors of study [14] used Matrix Geometric (analytical) Model by providing inputs numerical results of two class computed under Markovian Modulated Poisson Process (MMPP). A notable discrepancy of MMPP is the estimation of a outsized set of parameters. Methods for end-to-end delay estimation require Weighted Fair Queuing (WFQ) and Virtual Clock necessities [6]. Furthermore, Markov-Modulated rate Process (MMRP) can prioritize each class in its own buffer as reported in study [15]. The flow control management based on the computation of probability of various types of traffic belongs to multiple classes has been discussed in study [27]. The other work related to this study can be found in studies [16-18]. In previous work, the issue of providing QoS guarantees to the end-user based on tight bound QoS parameters has not been properly addressed.

In addition, we refer the readers to [19-22] regarding work that has been carried out in terms of IP network performance evaluation. The diversity of analysis conducted in [19-22] is that, their reported queuing models cannot employ self-similar network traffic and other disadvantage is that, they used only single class traffic for conducting analysis by neglecting the performance affect of other subsequent traffic classes.

To overcome the limitations of prior work, we presented a novel analytical framework [28-29] based on G/M/1 queuing system, that takes into account multiple classes of self-similar traffic. In our prior

work, we analyzed the traditional scheduling schemes such as priority and round robin. It is well known that traditional scheduling schemes can't provide the required QoS to all types of traffic found in modern wireless networks. Hence, in this current study, we analyze the G/M/1 model on the basis of a novel and most promising scheduling mechanism titled as, "Best Scheduling Algorithm (BSA)" and find exact packet delays for the corresponding classes of self-similar traffic. The results indicate that BSA completely outperforms all traditional and other available scheduling schemes. To date, no closed form expressions have been presented for G/M/1 with such scheduling scheme.

III. SELF-SIMILAR TRAFFIC WITH SEVERAL CLASSES

The traffic model considered here [23] belongs to a particular class of self-similar traffic models also called telecom process in [24], recently. The model depicts the facility of packet creation while considering the scaling characteristics of telephony network traffic. Such models, also called infinite source models, are similar to on/off processes with heavy tailed on and/or off times. Our proposed model investigates the arrival process of packet through queuing analysis.

In Poisson point process, our model embodies an infinite set of prospective resources. The traffic count calculation depends on the number of packets produced under such resources. Each resource has a ability to create a session with a weighty-tailed distribution, significantly, a Pareto distribution whose density is given by $g(r) = \delta b^\delta r^{-\delta-1}$, $r > b$

where δ is related to the Hurst parameter by $H = 3 - \delta / 2$. The sessions appear with respect to a Poisson process with rate λ . The packets pull in with respect to a Poisson process with rate α , locally, over each session.

For each class, the traffic $Y(t)$ measured as the total number of packets injected in $[0, t]$ is found by

$$Y(t) = \sum_{S_i \leq t} U_i(R_i \wedge t - S_i)$$

where U_i, R_i, S_i denote the local Poisson process, the duration and the arrival time of session i , respectively. Hence, $Y(t)$ corresponds to the sum of packets generated by all sessions initiated in $[0, t]$ until the session expires if that happens before t , and until t if it does not. The inactive edition of this model depends on an infinite past is selected in below estimations. Our model requires fixed sized packets because every queue links to a specific kind of application.

The traffic model Y is long-range dependent and almost second-order self-similar; the auto covariance function of its increments is that of fractional

Gaussian noise. Three varieties of weighty traffic limits are achievable depending on the rate enlargement of traffic constraints [23, 24]. Two of these limits parameters are common self-similar processes, fractional Brownian motion and Levy process, that do not consider the packet characteristics particularly.

Interarrival Times Calculations

Inter arrival time allocations for the specific self-similar traffic system are computed for the first time in [25]. We assume a solitary kind of packet first. The allocations of cross interarrival time among different kinds of packets are drawn on the root of lone packet results. In this study we just report the concluding results of interarrival time estimations between four kinds of traffic.

$$f_{T_{ii}}(t) = f_{T_i}(t) \bar{F}_j^0(t) \bar{F}_k^0(t) \bar{F}_l^0(t)$$

$$f_{T_{ij}}(t) = f_{T_j}^0(t) \bar{F}_{T_i}(t) \bar{F}_{T_k}^0(t) \bar{F}_{T_l}^0(t)$$

We refer the readers to [28] for the detailed derivation.

IV. SS/M/1 WITH FOUR CLASSES

We consider a model of four queues based on G/M/1 by considering four different classes of self-similar input traffic denoted by SS/M/1, and analyze it on the basis of our proposed scheduling scheme.

The scheduler serves the 4 queues according to the following logic. The queue at number one possesses peak priority and the scheduler serves prioritized queue all the time, in case of zero packet waiting in queue 1, it can deal with queue 2, 3 and 4 according to definite byte-count in a round robin way. We state the byte-count for queue 2, 3 and 4 as mentioned. The scheduler can process two packets from queue 2, one packet from queue 3 and one packet from queue 4 in each round if and only if packet waiting in queue one is zero.

Suppose the service time allocations have rate μ_1 , μ_2 , μ_3 and μ_4 for type 1, type 2, type 3 and type 4 packets, respectively, and let type 1 packets have the highest priority and type 4 packets have the lowest priority.

The standard embedded Markov chain [26] formulation of G/M/1 requires the examination of the queueing model at the time of coming instants, just before an arrival. At such instants, the count in the model is the amount of packets appeared in the queue plus packets in service, if any, not including packet appeared itself. We denote the positions and the transition probability matrix P of the Markov

chain with the self-similarity based model of four kinds of traffic.

Let $\{X_n : n \geq 1\}$ denote the embedded Markov chain at the time of coming events. As the service is depends on the priority, packet type in the service so each coming instant determines the queuing time. Therefore, we describe the state space as:

$$S = (i_1, i_2, i_3, i_4, a, s) : a \in \{a_1, a_2, a_3, a_4\},$$

$$s \in \{s_1, s_2^1, s_2^2, s_3, s_4, I\}, i_1, i_2, i_3, i_4 \in Z_+$$

where a_1, a_2, a_3, a_4 are labels to denote the type of arrival, $s_1, s_2^1, s_2^2, s_3, s_4$ are labels to denote the type of packet in service, i_1, i_2, i_3, i_4 are the number of packets in each queue including a possible packet in service, I denotes the idle state in which no packet is in service or queued and Z_+ is the set of nonnegative integers. The positions of the Markov chain and the feasible transitions probabilities can be specified by assuming each case. We will only analyze the states with non-empty queues in this paper.

Transition from

$$(i_1, i_2, i_3, i_4, a_1, s_1) \rightarrow (j_1, j_2, j_3, j_4, a_2, s_2^1)$$

This is the case where a transition occurs from an arrival of packet of class 1 to an arrival of packet of class2 where the packet of class 1 has seen that a packet of class 1 is in service, with i_1 packets in class 1, i_2 packets of class 2, i_3 packets of class 3 and, i_4 packets of class4 in the system. The transition occurs to j_1 packets of class 1, to j_2 packets of class 2, with a first packet of class 2 of some cycle in service, j_3 packets of class 3 and j_4 packets of class 4 in the system. Because of our scheduling theory an entrance of class two can view class-2 packet in queue in subsequent state only if all packets of class-1 including the most recently entered packet is exhausted at the interarrival time. Due to this reason j_1 can take only the value of 0 and exactly $i_1 + 1$ packets of class 1 are served during the interarrival time. On the contrary, the number of packets served from queue 2, let's say k , can have any value in the range between 0 and $i_2 - 1$ as at least one class 2 packet is in the system being the one in service when a new arrival happens. As we know that during each cycle, the scheduler serves 2 packets from queue 2, one packet from queue 3 and one packet from queue 4, we can consider queue 3 and queue 4 as a single queue and call it as queue I_3 . The translation probabilities hold two possibilities with the assumptions as follows:

Case (1): $i_2 < I_3$ and case (2): $i_2 \geq I_3$.

In the first case, the transition probability is:

$$P\{X_{n+} = 0, i_2 - \frac{t}{2}, i_3 - \frac{t}{2}, i_4 - \frac{t}{2}, a_2, s_2^1 | X_n = i_1, i_2, i_3, i_4, a_1, s_1\})$$

$$= \int_0^{\infty} \int_{t-x}^{\infty} \int_{S_2^1 + S_2^2 + S_2^3 + S_2^4} f_{S_2}(s) f_{S_2^1 + S_2^2 + S_2^3 + S_2^4}(x) f_{T_{12}}(t) ds dx dt$$

Whereas, in the second case, queue 3 and queue 4 might be exhausted as well and the transition probability will be:

$$= \int_0^{\infty} \int_{t-x}^{\infty} \int_{S_2^1 + S_2^2 + S_2^3 + S_2^4} f_{S_2}(s) f_{S_2^1 + S_2^2 + S_2^3 + S_2^4}(x) f_{T_{12}}(t) ds dx dt$$

Extraction of Packet Delay

Steady state distribution π as seen by an arrival can be found by solving $\pi P = \tau$ using the transition matrix P of the Markov chain analyzed above. In practice, the queue capacity is limited in a router. So, the steady state distribution exists. Our investigation depends on minimizing the allocation of the state of the queue at the coming instances, which can be estimated by employing the analysis given above for our self-similar traffic system. In general, the following analysis is valid for any G/M/1 queueing system where the limiting distribution π at the arrival instances can be computed. The expected waiting time for the highest priority queue can be found as

$$E[W_1] = \sum_{j_1=1}^{J_1-1} \sum_{j_2=0}^{J_2} \sum_{j_3=0}^{J_3} \sum_{j_4=0}^{J_4} j_1 \pi(j_1, j_2, j_3, j_4, a_1, s_1) +$$

$$\sum_{j_1=0}^{J_1-1} \sum_{j_2=1}^{J_2} \sum_{j_3=0}^{J_3} \sum_{j_4=0}^{J_4} \left(j_1 + \frac{1}{\mu} \right) \pi(j_1, j_2, j_3, j_4, a_1, s_2) +$$

$$\sum_{j_1=0}^{J_1-1} \sum_{j_2=0}^{J_2} \sum_{j_3=1}^{J_3} \sum_{j_4=0}^{J_4} \left(\frac{1}{\mu} + \frac{1}{\mu} \right) \pi(j_1, j_2, j_3, j_4, a_1, s_3) +$$

$$\sum_{j_1=0}^{J_1-1} \sum_{j_2=0}^{J_2} \sum_{j_3=0}^{J_3} \sum_{j_4=1}^{J_4} \left(\frac{1}{\mu} + \frac{1}{\mu} \right) \pi(j_1, j_2, j_3, j_4, a_1, s_4)$$

where J_1, J_2, J_3 and J_4 are the respective capacities of each queue. This clearly invokes that an higher priority coming packet will wait until all packets of this priority and the packet in the queue are served. On considering the category of the packet in queue, we have drawn the constituent expressions in the sum.

Furthermore, we got the expected waiting time for the low priority queues by investigating the instances that compose this delay. We assumed two factors (the impact of high priority queue and the effect of round-robin service) to find out the expected waiting time of a packet (class 2, 3 and class 4) arriving to non-priority queues (queue 2, 3 and queue 4). The

exact bounds on the expected waiting time for a class 2 packet can be computed as follows:

$$C_2 < E[W_2] < C_2'$$

$$c_2 = \sum_{j_1=1}^{J_1} \sum_{j_2=0}^{J_2-1} \sum_{j_3=0}^{\lfloor J_2/2 \rfloor} \sum_{j_4=0}^{\lfloor J_2/2 \rfloor} \left(\frac{1}{\mu_1} + \frac{j_1-1}{\mu_1} + \frac{j_2}{\mu_2} + \frac{j_3}{\mu_3} + \frac{j_4}{\mu_4} \right) \pi(j_1, j_2, j_3, j_4, a_2, s_1) +$$

$$\sum_{j_1=1}^{J_1} \sum_{j_2=0}^{J_2-1} \sum_{j_3=\lfloor J_2/2 \rfloor}^{J_3} \sum_{j_4=\lfloor J_2/2 \rfloor}^{J_4} \sum_{m=0}^1 \left(\frac{j_1}{\mu_1} + \frac{j_2}{\mu_2} + \frac{\lfloor J_2/2 \rfloor + m}{\mu_3} + \frac{\lfloor J_2/2 \rfloor + m}{\mu_4} \right) \pi(j_1, j_2, j_3, j_4, a_2, s_1) +$$

$$\sum_{j_1=0}^{J_1} \sum_{j_2=1}^{J_2-1} \sum_{j_3=0}^{\lfloor J_2/2 \rfloor} \sum_{j_4=0}^{\lfloor J_2/2 \rfloor} \left(\frac{1}{\mu_2} + \frac{j_1}{\mu_1} + \frac{j_2-1}{\mu_2} + \frac{j_3}{\mu_3} + \frac{j_4}{\mu_4} \right) \pi(j_1, j_2, j_3, j_4, a_2, s_2) +$$

$$\sum_{j_1=0}^{J_1} \sum_{j_2=1}^{J_2-1} \sum_{j_3=\lfloor J_2/2 \rfloor}^{J_3} \sum_{j_4=\lfloor J_2/2 \rfloor}^{J_4} \left(\frac{1}{\mu_2} + \frac{j_1}{\mu_1} + \frac{j_2-1}{\mu_2} + \frac{\lfloor J_2/2 \rfloor}{\mu_3} + \frac{\lfloor J_2/2 \rfloor}{\mu_4} \right) \times \pi(j_1, j_2, j_3, j_4, a_2, s_2) +$$

$$\sum_{j_1=0}^{J_1} \sum_{j_2=1}^{J_2-1} \sum_{j_3=0}^{\lfloor J_2/2 \rfloor} \sum_{j_4=0}^{\lfloor J_2/2 \rfloor} \left(\frac{1}{\mu_2} + \frac{j_1}{\mu_1} + \frac{j_2-1}{\mu_2} + \frac{j_3}{\mu_3} + \frac{j_4}{\mu_4} \right) \pi(j_1, j_2, j_3, j_4, a_2, s_2) +$$

$$\sum_{j_1=0}^{J_1} \sum_{j_2=1}^{J_2-1} \sum_{j_3=\lfloor J_2/2 \rfloor}^{J_3} \sum_{j_4=\lfloor J_2/2 \rfloor}^{J_4} \left(\frac{1}{\mu_2} + \frac{j_1}{\mu_1} + \frac{j_2-1}{\mu_2} + \frac{\lfloor J_2/2 \rfloor}{\mu_3} + \frac{\lfloor J_2/2 \rfloor}{\mu_4} \right) \times \pi(j_1, j_2, j_3, j_4, a_2, s_2) +$$

$$\sum_{j_1=0}^{J_1} \sum_{j_2=0}^{J_2-1} \sum_{j_3=1}^{\lfloor J_2/2 \rfloor} \sum_{j_4=0}^{\lfloor J_2/2 \rfloor} \left(\frac{1}{\mu_3} + \frac{j_1}{\mu_1} + \frac{j_2}{\mu_2} + \frac{j_3-1}{\mu_3} + \frac{j_4}{\mu_4} \right) \pi(j_1, j_2, j_3, j_4, a_2, s_3) +$$

$$\sum_{j_1=0}^{J_1} \sum_{j_2=0}^{J_2-1} \sum_{j_3=\lfloor J_2/2 \rfloor}^{J_3} \sum_{j_4=\lfloor J_2/2 \rfloor}^{J_4} \left(\frac{1}{\mu_3} + \frac{j_1}{\mu_1} + \frac{j_2}{\mu_2} + \frac{\lfloor J_2/2 \rfloor}{\mu_3} + \frac{\lfloor J_2/2 \rfloor}{\mu_4} \right) \times \pi(j_1, j_2, j_3, j_4, a_2, s_3) +$$

$$\sum_{j_1=0}^{J_1} \sum_{j_2=0}^{J_2-1} \sum_{j_3=0}^{\lfloor J^2/2 \rfloor} \sum_{j_4=1}^{\lfloor J^2/2 \rfloor} \left(\frac{1}{\mu_4} + \frac{j_1}{\mu_1} + \frac{j_2}{\mu_2} + \frac{j_3}{\mu_3} + \frac{j_4-1}{\mu_4} \right) \pi(j$$

$$, j_2, j_3, j_4, a_2, s_4) +$$

$$\sum_{j_1=0}^{J_1} \sum_{j_2=0}^{J_2-1} \sum_{j_3=\lfloor J^2/2 \rfloor}^{J_3} \sum_{j_4=\lfloor J^2/2 \rfloor}^{J_4} \left(\frac{1}{\mu_4} + \frac{j_1}{\mu_1} + \frac{j_2}{\mu_2} + \frac{\lfloor J^2/2 \rfloor}{\mu_3} + \frac{\lfloor J^2/2 \rfloor}{\mu_4} \right) \times \pi(j$$

$$, j_2, j_3, j_4, a_2, s_4) +$$

$$\frac{C_1}{\mu_1} \times c_2$$

$$c'2=$$

$$\sum_{j_1=1}^{J_1} \sum_{j_2=0}^{J_2-1} \sum_{j_3=\lfloor J^2/2 \rfloor}^{J_3} \sum_{j_4=\lfloor J^2/2 \rfloor}^{J_4} \sum_{m=0}^1 \left(\frac{j_1}{\mu_1} + \frac{j_2}{\mu_2} + \frac{\lfloor J^2/2 \rfloor + m}{\mu_3} + \frac{\lfloor J^2/2 \rfloor + m}{\mu_4} \right) \pi(j$$

$$, j_2, j_3, j_4, a_2, s_1) +$$

$$\sum_{j_1=0}^{J_1} \sum_{j_2=1}^{J_2-1} \sum_{j_3=\lfloor J^2/2 \rfloor}^{J_3} \sum_{j_4=\lfloor J^2/2 \rfloor}^{J_4} \left(\frac{1}{\mu_2} + \frac{j_1}{\mu_1} + \frac{j_2-1}{\mu_2} + \frac{\lfloor J^2/2 \rfloor}{\mu_3} + \frac{\lfloor J^2/2 \rfloor}{\mu_4} \right) \times \pi(j$$

$$, j_2, j_3, j_4, a_2, s_2^1) +$$

$$\sum_{j_1=0}^{J_1} \sum_{j_2=1}^{J_2-1} \sum_{j_3=\lfloor J^2/2 \rfloor}^{J_3} \sum_{j_4=\lfloor J^2/2 \rfloor}^{J_4} \left(\frac{1}{\mu_2} + \frac{j_1}{\mu_1} + \frac{j_2-1}{\mu_2} + \frac{\lfloor J^2/2 \rfloor}{\mu_3} + \frac{\lfloor J^2/2 \rfloor}{\mu_4} \right) \times \pi(j$$

$$, j_2, j_3, j_4, a_2, s_2^2) +$$

$$\sum_{j_1=0}^{J_1} \sum_{j_2=0}^{J_2-1} \sum_{j_3=\lfloor J^2/2 \rfloor}^{J_3} \sum_{j_4=\lfloor J^2/2 \rfloor}^{J_4} \left(\frac{1}{\mu_3} + \frac{j_1}{\mu_1} + \frac{j_2}{\mu_2} + \frac{\lfloor J^2/2 \rfloor}{\mu_3} + \frac{\lfloor J^2/2 \rfloor}{\mu_4} \right) \times \pi(j$$

$$, j_2, j_3, j_4, a_2, s_3) +$$

$$\sum_{j_1=0}^{J_1} \sum_{j_2=0}^{J_2-1} \sum_{j_3=\lfloor J^2/2 \rfloor}^{J_3} \sum_{j_4=\lfloor J^2/2 \rfloor}^{J_4} \left(\frac{1}{\mu_4} + \frac{j_1}{\mu_1} + \frac{j_2}{\mu_2} + \frac{\lfloor J^2/2 \rfloor}{\mu_3} + \frac{\lfloor J^2/2 \rfloor}{\mu_4} \right) \times \pi(j$$

$$, j_2, j_3, j_4, a_2, s_4) +$$

$$\frac{C_1}{\mu_1} \times c'2$$

Similarly the exact bounds on expected waiting time of a class 3/4 packet can be written. Due to the symmetry of service between queue 3 and queue 4, it is same. Hence

the expected waiting time for a class 3 packet can be written as follow:

$$c_3 \leq E[W_3] \leq c'3$$

$$c_3 = \sum_{j_1=1}^{J_1} \sum_{j_2=0}^{(J_3)} \sum_{j_3=0}^{J_3-1} \sum_{j_4=0}^{j_3} \left(\frac{1}{\mu_1} + \frac{j_1-1}{\mu_1} + \frac{j_2}{\mu_2} + \frac{j_3}{\mu_3} + \frac{j_4}{\mu_4} \right) \pi(j$$

$$, j_2, j_3, j_4, a_3, s_1) +$$

$$\sum_{j_1=1}^{J_1} \sum_{j_2=(J_3)+1}^{J_2} \sum_{j_3=0}^{J_3-1} \sum_{j_4=0}^{j_3} \sum_{m=0}^1 \left(\frac{j_1}{\mu_1} + \frac{(j_3)+m}{\mu_2} + \frac{j_3}{\mu_3} + \frac{j_3}{\mu_4} \right) \pi(j$$

$$, j_2, j_3, j_4, a_3, s_1) +$$

$$\sum_{j_1=0}^{J_1} \sum_{j_2=1}^{(J_3)+2} \sum_{j_3=0}^{J_3-1} \sum_{j_4=0}^{j_3} \left(\frac{1}{\mu_2} + \frac{j_1}{\mu_1} + \frac{j_2-1}{\mu_2} + \frac{j_3}{\mu_3} + \frac{j_4}{\mu_4} \right) \pi(j$$

$$, j_2, j_3, j_4, a_3, s_2^1) +$$

$$\sum_{j_1=0}^{J_1} \sum_{j_2=(J_3)+3}^{J_2} \sum_{j_3=0}^{J_3-1} \sum_{j_4=0}^{j_3} \left(\frac{1}{\mu_2} + \frac{j_1}{\mu_1} + \frac{(j_3)+1}{\mu_2} + \frac{j_3}{\mu_3} + \frac{j_3}{\mu_4} \right) \times \pi(j$$

$$, j_2, j_3, j_4, a_3, s_2^1) +$$

$$\sum_{j_1=0}^{J_1} \sum_{j_2=1}^{(J_3)+1} \sum_{j_3=0}^{J_3-1} \sum_{j_4=0}^{j_3} \left(\frac{1}{\mu_2} + \frac{j_1}{\mu_1} + \frac{j_2-1}{\mu_2} + \frac{j_3}{\mu_3} + \frac{j_4}{\mu_4} \right) \pi(j$$

$$, j_2, j_3, j_4, a_3, s_2^2) +$$

$$\sum_{j_1=0}^{J_1} \sum_{j_2=(J_3)+2}^{J_2} \sum_{j_3=0}^{J_3-1} \sum_{j_4=0}^{j_3} \left(\frac{1}{\mu_2} + \frac{j_1}{\mu_1} + \frac{(J_3)}{\mu_2} + \frac{j_3}{\mu_3} + \frac{j_3}{\mu_4} \right) \times \pi(j$$

$$, j_2, j_3, j_4, a_3, s_2^2) +$$

$$\sum_{j_1=0}^{J_1} \sum_{j_2=0}^{(J_3)} \sum_{j_3=1}^{J_3-1} \sum_{j_4=0}^{j_3} \left(\frac{1}{\mu_3} + \frac{j_1}{\mu_1} + \frac{j_2}{\mu_2} + \frac{j_3-1}{\mu_3} + \frac{j_4}{\mu_4} \right) \pi(j$$

$$, j_2, j_3, j_4, a_3, s_3) +$$

$$\sum_{j_1=0}^{J_1} \sum_{j_2=(J_3)+1}^{J_2} \sum_{j_3=1}^{J_3-1} \sum_{j_4=0}^{j_3} \left(\frac{1}{\mu_3} + \frac{j_1}{\mu_1} + \frac{(J_3)}{\mu_2} + \frac{j_3-1}{\mu_3} + \frac{j_3}{\mu_4} \right) \times \pi(j$$

$$, j_2, j_3, j_4, a_3, s_3) +$$

$$\sum_{j_1=0}^{J_1} \sum_{j_2=0}^{(J_3)} \sum_{j_3=0}^{J_3-1} \sum_{j_4=1}^{j_3} \left(\frac{1}{\mu_4} + \frac{j_1}{\mu_1} + \frac{j_2}{\mu_2} + \frac{j_3}{\mu_3} + \frac{j_4-1}{\mu_4} \right) \pi(j$$

$$, j_2, j_3, j_4, a_3, s_4) +$$

$$\sum_{j_1=0}^{J_1} \sum_{j_2=(j_3)+1}^{J_2} \sum_{j_3=0}^{J_3-1} \sum_{j_4=1}^{J_4} \left(\frac{1}{\mu_4} + \frac{j_1}{\mu_1} + \frac{(j_3)}{\mu_2} + \frac{j_3}{\mu_3} + \frac{j_3}{\mu_4} \right) \times \pi(j_{1,j_2,j_3,j_4,a_3,s_4}) +$$

$$\frac{C1}{\mu_1} \times c_3$$

And

$$c'3=$$

$$\sum_{j_1=1}^{J_1} \sum_{j_2=(j_3)+1}^{J_2} \sum_{j_3=0}^{J_3-1} \sum_{j_4=0}^{J_4} \sum_{m=0}^1 \left(\frac{j_1}{\mu_1} + \frac{(j_3)+m}{\mu_2} + \frac{j_3}{\mu_3} + \frac{j_3}{\mu_4} \right) \pi(j_{1,j_2,j_3,j_4,a_3,s_1}) +$$

$$\sum_{j_1=0}^{J_1} \sum_{j_2=(j_3)+3}^{J_2} \sum_{j_3=0}^{J_3-1} \sum_{j_4=0}^{J_4} \left(\frac{1}{\mu_2} + \frac{j_1}{\mu_1} + \frac{(j_3)+1}{\mu_2} + \frac{j_3}{\mu_3} + \frac{j_3}{\mu_4} \right) \times \pi(j_{1,j_2,j_3,j_4,a_3,s_2^1}) +$$

$$\sum_{j_1=0}^{J_1} \sum_{j_2=(j_3)+2}^{J_2} \sum_{j_3=0}^{J_3-1} \sum_{j_4=0}^{J_4} \left(\frac{1}{\mu_2} + \frac{j_1}{\mu_1} + \frac{(j_3)}{\mu_2} + \frac{j_3}{\mu_3} + \frac{j_3}{\mu_4} \right) \times \pi(j_{1,j_2,j_3,j_4,a_3,s_2^2}) +$$

$$\sum_{j_1=0}^{J_1} \sum_{j_2=(j_3)+1}^{J_2} \sum_{j_3=1}^{J_3-1} \sum_{j_4=0}^{J_4} \left(\frac{1}{\mu_3} + \frac{j_1}{\mu_1} + \frac{(j_3)}{\mu_2} + \frac{j_3-1}{\mu_3} + \frac{j_3}{\mu_4} \right) \times \pi(j_{1,j_2,j_3,j_4,a_3,s_3}) +$$

$$\sum_{j_1=0}^{J_1} \sum_{j_2=(j_3)+1}^{J_2} \sum_{j_3=0}^{J_3-1} \sum_{j_4=1}^{J_4} \left(\frac{1}{\mu_4} + \frac{j_1}{\mu_1} + \frac{(j_3)}{\mu_2} + \frac{j_3}{\mu_3} + \frac{j_3}{\mu_4} \right) \times \pi(j_{1,j_2,j_3,j_4,a_3,s_4}) +$$

$$\frac{C1}{\mu_1} \times c'3$$

V. SIMULATION RESULTS

A discrete event simulator for queuing systems was used to observe and evaluate the QoS behavior of self-similar traffic. The simulator is highly modular equipped with self-similar traffic generator and also allowing free customization of any newly designed scheduling logic. Further, it enables the evaluation of any chosen scheduling discipline under any type of input traffic. The *scheduler* class holds the key functions for the schedule logic which enables any scheduling algorithm used to be loosely coupled but easily integrated. The *BSAScheduler* was implemented to analyze the behavior of the corresponding QoS of different traffic classes. A

traffic generator was also implemented that reflects the traffic model described in previous traffic model section. Furthermore, a number of other classes also were implemented to facilitate program function such as:

- *Simulation*: this class acts as the simulation engine where it moves time forward and updates the event list etc.
- *Random-Number*: a class that generate a random number with each specific distribution.
- *Packet*: a class which stores the system states as encountered by every packet.
- *Traffic Parameter*: a header defined to hold numerical values needed for computing the divergence to resist the instability in the numerical results in the cumulative QoS statistics.

For the higher priority queue (class 1 packet) we set the session arrival rate to $\lambda_1 = \frac{1}{5} s$, the in-session packet arrival rate to $\alpha_1 = \frac{1}{50} s$ which is the characteristics of VoIP traffic. Also, the service rate to $\mu_1 = \frac{1}{2500} s$

For the rest of the queues (queue2, 3, and 4) we set the session arrival rate same as in queue 1 $\lambda_2 = \lambda_3 = \lambda_4 = \frac{1}{50} s$, the in-session packet arrival rate to $\alpha_2 = \alpha_3 = \alpha_4 = (1/5)$ and the service rate to $\mu_2 = \mu_3 = \mu_4 = \mu_1$. Fig. 1 below shows the mean delay in (ms) for each queue at different values of Hurst parameter value (H=0.5, H=0.75, H= 0.9). We can notice particularly in the case of non- priority queues that as the value of H increases, the queueing delay also increases. Further, we can notice the effect of our newly designed scheduler (BSA), in which queue 1 packets have the highest priority and queue 2, 3 and 4 packets are being treated by the scheduler as non-priority packets. Also we can notice the symmetry of service for queue 3 and queue 4 packets because the packets in both queues are experiencing the same expected delay.

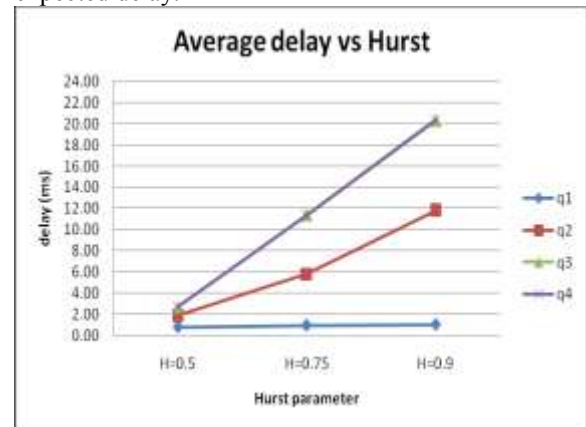


Fig. 1: Average delay vs. Hurst Parameter for BSA

VI. CONCLUSION AND FUTURE WORK

In this paper, we have extended our priory work based on G/M/1 queueing system for accurate modeling of wireless IP traffic behavior through presenting a novel scheduling scheme called as BSA. The simulation results clearly indicate that our proposed scheduling algorithm outperforms the traditional scheduling schemes such as priority and round-robin. The BSA provides a preferential treatment to real time applications by offering a very low delay but at the same time, this preference is not up to that extent that generic data applications are starving for bandwidth. In our future work, we are intending to explore the possibility of practical implementation of proposed BSA in different 4G wireless networks.

ACKNOWLEDGEMENT

This research work has been supported by National Plan for Science and Technology (NPST) Grant (10-INF1112-02) at King Saud University, Riyadh, KSA.

REFERENCES

- [1] G. Armitage, "Quality of Service in IP Networks", MTP, 2004, pp. 105-138
- [2] M. S. Taqqu, Will E. Leland, W. Willinger and D Wilson, "On the Self-Similar nature of Ethernet traffic", in *IEEE/ACM Transactions on Networking*, vol. 2, issue 1, Feb. 1994, pp. 1-15
- [3] W. Willinger, M. W. Garrett, "Analysis, Modeling and Generation of Self-Similar VBR Video Traffic", *ACM SIGCOMM Symposium*, Oct. 1994, ACM Computer Communication Review, vol. 24
- [4] M. Corvella, A. Bestavros, "Explaining World Wide Web Traffic Self-Similarity", Tech. Rep. TR-95-015, Boston University, CS Dept, Boston, MA 02215, Aug 1995
- [5] M. S. Taqqu et al, "Theory and Applications of Long-Range Dependence", Birkkhauser Boston, 2003, pp. 409-415
- [6] H. Zhang, "Service disciplines for packet switching network protocols", *Proc. of IEEE*, 83(10): October 1995
- [7] A. Erramilli, O. Narayan and W. Willinger, "Experimental queueing analysis with long-range dependent packet traffic", *IEEE/ACM Trans. Network Comput.*, 1996, 4: pp. 209-223
- [8] B. Melemed, M. Livny and A. K. Tsilois, "The impact of autocorrelations on queueing systems", *Manage. Sci.*, 1993, 39, pp. 322-339
- [9] A. Adas and A. Mukherjee, "On Resource Management and QoS guarantees for long-range dependant traffic", In proc IEEE INFOCOMM 95, pp. 779-787
- [10] M. Parulekar and A. Makowski, "Tail Probabilities for a multiplexer with self-similar input", In proc IEEE INFOCOMM 96, pp. 1452-1459
- [11] I. Norros, "A Storage model with self-similar input", *Queueing System*, 16, 1994, pp. 387-396
- [12] B. Tsybakov and N. D. Georganas, "Self-Similar traffic and upper bounds to buffer overflow in ATM queue", *Performance Evaluation*, 36, 1998, pp. 57-80
- [13] R. Addie, M. Zukerman and T. Naeme, "Fractal traffic: measurements, modeling and performance evaluation", In proc IEEE INFOCOMM 95, pp. 977-984
- [14] M. Zukerman et al, "Analytical Performance Evaluation of a Two Class DiffServ Link", *IEEE ICS*, 25-28 Nov. 2002, vol. 1, pp. 373-377
- [15] J. Zhang, "Performance study of Markov modulated fluid flow models with priority traffic", in Proc. IEEE INFOCOMM 93, San Francisco, CA, Mar. 30-Apr. 1, 1993, pp. 10-17
- [16] A. I. Elwalid and D. Mitra, "Fluid Models for analysis and design of statistical multiplexing with loss priorities on multiple classes of bursty traffic" in Proc. IEEE INFOCOMM 92, Florence, Italy, 1992, pp. 415-425
- [17] G. L. Choudhury, K. K. Leung and W. Whitt, "An algorithm to compute blocking probabilities in multi-rate multi-class multi-resource loss models", *Adv. Appl. Prob.* vol. 27. pp. 1104-1143. 1995
- [18] G. L. Choudhury, K. K. Leung and W. Whitt, "An inversion algorithm to compute blocking probabilities in loss networks with state-dependant rates", *IEEE/ACM Transactions on Networking*, vol. 3. pp. 585-601, 1995
- [19] C. F. Chou et al, "Low Latency and efficient packet scheduling for streaming applications" *IEEE International Conference on Communications*, 20-24 June, 2004, Vol. 4, pp. 1963-1967
- [20] A. Kos, B. Klepec, "Performance of VoIP Applications in a Simple Differentiated Services Network Architecture", *IEEE EUROCON*, 4-7 July, 2001, Vol. 1, pp. 214-217
- [21] N. F. MIR, A. Chien, "Simulation of Voice over MPLS communication networks", *IEEE ICCS*, 25-28 Nov. 2002, Vol. 1, pp. 389-393
- [22] J. M. Chung, H. M. Soo, "Analysis of Non Preemptive Priority Queueing of MPLS networks with Bulk Arrivals", *IEEE MWSCAS*, 4-7 Aug. 2002, vol. 3. pp. 81-84
- [23] M. Caglar, "A Long-Range Dependent workload model for packet data traffic", *Mathematics of Operations Research*, 29, 2004, pp. 92-105
- [24] I. Kaj, "Limiting fractal random processes in heavy-tailed systems" in *Fractals in Engineering, New Trends in Theory and Applications*, Eds. J. Levy-Lehel, E. Lutton, Springer-Verlag London 2005, pp. 199-218
- [25] M. Iftikhar, B. Landfeldt and Mine Caglar, "An analytical model based on G/M/1 with Self-Similar input to provide End-to-End QoS in 3G networks" in press *ACM Mobiwac 2006 (MSWIM)*
- [26] E. Cinlar, "Introduction to Stochastic Processes", 1975. pp. 178
- [27] Do. Young. Eun and Ness. B. Shroff, "A measurement-analytical approach for QoS estimation in a network based on dominant time scale" in *IEEE/ACM Trans. on Networking*, vol. 11, No. 2, April 2003, pp. 222-235
- [28] M. Iftikhar et al, "SLAs parameter negotiation between heterogeneous 4G wireless network operators", *Elsevier Journal of Pervasive and Mobile Computing*, vol. 7, issue 5, October 2011, pp. 525-544
- [29] M. Iftikhar et al, "Towards the formation of comprehensive SLAs between heterogeneous wireless DiffServ domains", *Springer journal of Telecommunication Systems* (2009) 42: 179-199