NEW SCHEDULING POLICY FOR ESTIMATION OF STATIONARY PERFORMANCE CHARACTERISTICS IN SINGLE SERVER QUEUES WITH INACCURATE JOB SIZE INFORMATION

Lusine Meykhanadzhyan
Peoples' Friendship University of Russia,
Moscow, Russia
Email: lameykhanadzhyan@gmail.com

Rostislav Razumchik
Institute of Informatics Problems
of the FRC CSC RAS, Moscow, Russia,
Peoples' Friendship University of Russia,
Moscow, Russia
Email: rrazumchik@ipiran.ru

KEYWORDS

inaccurate job size, long tails, service policy, size-based scheduling

ABSTRACT

The study of size-based and size-oblivious scheduling policies with inaccurate job size information appears nowadays to be an important direction of scientific studies because as recent research results show advantages of size-based policies can be saved even when the job sizes are not perfectly known a priori. This paper is focused on the same topic but touches upon a different question: is it possible to predict such estimates of system's performance characteristics (for example, job's mean sojourn time), that will be close to those which will be observed in practice, if the scheduler is provided only with the inaccurate information about the job size distribution? It is shown here that there are conditions under which the answer to the question is positive. A simple mathematical model ($M/G/1$ queueing system) of a top level view of a data-intensive execution engine is being proposed. It is shown that, in case of long-tailed service time distribution, a special service policy – Preemptive-Last-Come-First-Served with service time re-generation on arrival instants – allows one to obtain better upper bounds for job's mean sojourn time than those achieved by common work conserving policies. Extensive numerical examples are presented.

INTRODUCTION

The topic of this paper concerns performance evaluation of data-intensive systems in the presence of job size estimation errors. It is known that celebrated size-based scheduling policies which allow one to increase system's performance by changing the service policy start to perform poor (compared to size-oblivious policies like Processor sharing) as the uncertainty about the job size distribution increases (see Lu et al. (2004)). Since the inaccuracy in job size information is reported to be quite a common issue in practice (see Dell'Amico (2013); Dell'Amico et al. (2015); Wierman (2008)) there appeared a number of research papers appeared that focused on the behaviour of size-based scheduling policies with inaccurate size estimations. To our knowledge one can find in Dell'Amico et al. (2015); Wierman (2008); Harchol-Balter et al. (2003); Chang et al. (2011) the most recent results on this topic (including short reviews).

An important question which is being faced in systems with inaccurate job size information is whether it is possible to devise a scheduler which performs as good as the scheduler fed with the accurate information. A number of studies show, that sometimes it is possible (Dell'Amico et al. (2015)). In this paper we try to look at the performance of such systems from a different point of view and raise another question. We start with an example. Assume that one has a simulation or a physical copy of the data-intensive system. Each job is admitted into the system and after being served departs from it, and never comes back. In practice this system will be fed for a long time with the flow of jobs of random size s, exact distribution of which is unknown to the system but is known to be long-tailed. Denote the mean sojourn time which will be seen in the future by v^*. Before launching the system the owner is interested in job's mean sojourn time which can be guaranteed by the system. In order to find it out one can run simulation experiments with one (or more) scheduler and some job's traces which contain *user's estimated* job's sizes which we denote by $\hat{s}$. Denote by v the estimated mean sojourn time. Clearly v is only the estimate for the unknown value v^*. The questions are: (i) is it possible to improve the estimate v by using only the available information about the distribution of $\hat{s}$ and having control over the scheduler? (ii) if it is possible then what is the quality of the estimate? In this paper one shows that in some cases (when the distribution of $\hat{s}$ is long-tailed) the answer to the question (i) is positive. The quality of estimate is investigated numerically.

As the answer to the question (i) may depend on many technical aspects related to real-life systems we will restrict ourselves to a probably the most simple case: a system is modelled by a single-server queue of $M/G/1$ type. An extent to which such assumption is an oversimplification of the real-life systems can be seen, for example, from papers Lu et al. (2004); Qiao et al. (2004); Dell'Amico et al. (2015) where it was used to evaluate

Proceedings 30th European Conference on Modelling and Simulation ©ECMS Thorsten Claus, Frank Herrmann, Michael Manitz, Oliver Rose (Editors)
ISBN: 978-0-9932440-2-5 / ISBN: 978-0-9932440-3-2 (CD)

size-based scheduling policies with inaccurate job size information. Coming back to the question (i), we suggest to use a special service policy which utilizes the assumption[1] that job's true service time distribution is long-tailed. The idea of the policy is the following. Suppose that upon each arrival of a job the processing of current running job is interrupted and one re-generates the service time of both customers (depending on their current service times and according to the distribution of $\hat{s}$) and then the service is resumed with the new service time and the arrived customer occupies one place in the queue. A number of questions arise here. Will such service policy lead to a better estimate of v, than the ordinary work conserving policy like processor sharing? If yes then under what conditions? One can expect this policy to be efficient sometimes because long-tailed distributions "tend" to have decreasing hazard rates and increasing mean residual lives. As the numerical experiments show this is really so.

In this paper we propose a simple mathematical model, which is a top level view of a data-intensive execution engine and thus does not take into consideration most of its technical details. Specifically one considers an execution engine as an $M/G/1$ queueing system, with an infinite queue, a single flow of jobs, i.i.d. execution (service) times and pre-emptive last come first served discipline which allows service time re-generation on arrival instants. The comments to assumptions allowing one to take such a simplified view can be found in papers Dell'Amico (2013); Dell'Amico et al. (2015), devoted to the design of the new scheduler for Handoop execution engine for data-intensive systems. At first we present some analytical results concerning the analysis of system's stationary characteristics, concentrating on sojourn time distribution. These results heavily rely on Meykhanadzhyan et al. (2014) and thus are presented in short. Then we show that if the service time distribution is long-tailed (either s or $\hat{s}$) PLCFS-re policy allows one to obtain upper bounds for the true value v^*, which can be much better than the values obtained using common work conserving disciplines. The comparison is done with classic size-based and size-oblivious service policies, which use the inaccurate job size information for scheduling.

In the next section the detailed description of the mathematical model is presented, which is followed by some results concerning the system's stationary performance characteristics. Section 3 is devoted to numerical results. In the conclusion one discusses in short the obtained results and directions of further research.

MATHEMATICAL MODEL

Consideration is given to a queueing system with one queue on infinite capacity, one server and Poisson arrival flow of rate λ. Service times of customers are i.i.d. random variables with known cumulative distribution function $B(x)$ and density $b(x) = B'(x)$. We assume that the customer's service time becomes known upon its arrival at the system and at any time instant the service time (remaining service time) of each customer in the system (both in the server and in the queue) is known. Newly arriving customers and those which are in the queue obey the special service policy which we will call Preemptive-Last-Come-First-Served with re-service (PLCFS-Re). It implies the following service rule. When a customer arrives at the system it interrupts the service process of the customer in server (if any) and compares its service time u with the (remaining) service time v of the customer in server. Then the arrived customer receives new service time U and the customer in service receives new service time V according to the known distribution

$$D(x, y|u, v) = \mathbf{P}\{U < x, V < y|u, v\}$$

which can depend on u and v. After that the service of the customer in server is resumed with the updated service time V and the arrived customer updates its service time with U and occupies the last place in the queue. When the service time of a customer in service becomes zero it leaves the system and one customer from the last place in the queue enters server. For the sake of convenience it is assumed that the density $d(x, y|u, v) = \partial^2 D(x, y|u, v)/(\partial x\, \partial y)$ is continuous and bounded. Note that for each u and v the following identity holds:

$$\int_0^\infty \int_0^\infty d(x, y|u, v)\, dx\, dy = D(\infty, \infty|u, v) = 1. \qquad (1)$$

Let the stationary regime of the system exist. The main performance characteristic under study in this paper is the system's stationary mean sojourn time. Denote by $\beta(s)$ the Laplace-Stieltjes transform (LST) of $B(x)$ i.e.

$$\beta(s) = \int_0^\infty e^{-sx} dB(x).$$

For example, if $d(x, y|u, v) = b(x)b(y)$ i.e. the new service times of the arriving customer and the customer in server are chosen (upon arrival) independently of their previous service times u and v the stationary regime (and the mean sojourn time) exists if and only if the inequalities $1/2 < \beta(\lambda) < 1$ hold.

Stationary probabilities

Denote by $\nu(t)$ the number of customers in the system at instant t, and by $\vec{\xi}(t) = (\xi_1(t), \ldots, \xi_{\nu(t)}(t))$ — the row vector, in which $\xi_1(t)$ is the (remaining) service time of the customer in server, $\xi_2(t)$ — the service time of the 1^{st} customer in the queue,..., $\xi_{\nu(t)-1}(t)$ — the service time of the last, $(\nu(t) - 1)$, customer in the queue. If $\nu(t) = 0$ the vector $\vec{\xi}(t)$ is not defined. Then the process $\eta(t) = (\nu(t), \vec{\xi}(t))$ describing the evolution of the number of customers in the system is a continuous-time Markov chain.

[1]The practical evidences for such an assumption can be found, for example, in Ren et al. (2012); Crovella (2001).

Let us introduce the stationary distribution of the chain $\eta(t)$:

$$p_0 = \lim_{t\to\infty} \mathbf{P}\{\nu(t) = 0\},$$

$$P_n(x_1, \ldots, x_n) = \lim_{t\to\infty} \mathbf{P}\{\, \nu(t) = n, \xi_1(t) < x_1, \ldots, \xi_n(t) < x_n\}, \;\; n \geq 1.$$

From the system's description and introduced assumptions it can be shown that density functions

$$p_n(x_1, \ldots, x_n) = \frac{\partial^n}{\partial x_1 \cdots \partial x_n} P_n(x_1, \ldots, x_n), \;\; n \geq 1,$$

are continuous and bounded. Using the properties of the restricted Markov chains and the properties of the PLCFS-Re service policy one can write out the system of integro-differential equations for the stationary density functions $p_n(x_1, \ldots, x_n)$. The details of this approach can be found in Meykhanadzhyan et al. (2014). But in order to compute the mean sojourn time it is enough to find the marginal density functions $p_1(x)$ and

$$p_n(x) = \int \ldots \int\limits_{x_2,\ldots,x_n>0} p_n(x, x_2, \ldots, x_n)\, dx_2 \cdots dx_n, \;\; n \geq 2,$$

which take intro consideration only the number of customers in the system and the remaining service time of the customer in server. Referring again to the approach in Meykhanadzhyan et al. (2014), one can obtain the following system of integro-differential equations for the functions $p_n(x)$:

$$-p_n'(x) = a_n(x) - \lambda p_n(x) + \int\limits_0^\infty K_n(x, v) p_n(v)\, dv, \; n \geq 1, \quad (2)$$

where $a_1(x) = \lambda b(x) p_0$,

$$a_n(x) = \lambda \int\limits_0^\infty p_{n-1}(v)\, dv \int\limits_0^\infty b(u)\, du \int\limits_0^\infty d(y, x|u, v)\, dy, \;\; n \geq 2,$$

$$K_n(x, v) = \lambda \int\limits_0^\infty b(u)\, du \int\limits_0^\infty d(x, y|u, v)\, dy, \;\; n \geq 1.$$

The initial conditions for the system (2) follow from the properties of density functions and have the form

$$\lim_{x\to\infty} p_n(x) = 0, \;\; n \geq 1.$$

Note that for an arbitrary function $d(x, y|u, v)$ the system (2) can be solved numerically. Substitution of $p_n(x) = e^{\lambda x} q_n(x)$ into (2) and subsequent integration of the new system for $q_n(x)$ lead to Fredholm equations of the second kind with non-negative kernels, for which standard approaches are still feasible (for example, iteration with first iteration equal to zero). The only unknown probability left is the probability of the empty system p_0, which can be computed from the normalization condition

$$\sum_{n=0}^{\infty} p_n = 1,$$

where

$$p_n = \int_0^\infty p_n(x)\, dx, \;\; n \geq 1,$$

is the stationary probability of n customers in the system.

In some cases, which are relevant for practice (see the numerical section), one can easily compute some useful quantities from (2). Consider the following special case of the PLCFS-Re service policy. Let upon arrival of a new customer its service time u and the (remaining) service time v of the customer in server (if any) be regenerated independently of u and v (i.e. independently of how long the customer has been already served) according to $B(x)$ (i.e. $d(x, y|u, v) = b(x)b(y)$). Then the system (2) can be simplified to the following form:

$$-p_n'(x) = \lambda b(x) p_{n-1} - \lambda p_n(x) + \lambda b(x) p_n, \;\; n \geq 1. \quad (3)$$

Remarks to the solution of this system remain the same as to the system (2). But here one is able to find the explicit expression for the mean number of customers in the system without finding all the unknown functions $p_n(x)$. Indeed, if one introduces the generating function

$$\pi(z, x) = \sum_{n=1}^{\infty} p_n(x) z^n,$$

then, if one treats z as a parameter, the solution of the system (3) in terms of $\pi(z, x)$ can be written in the form

$$\pi(z, x) = \int\limits_x^\infty e^{-\lambda(u-x)} [\lambda z b(u) p_0 + \lambda(1+z) b(u) \pi(z)] du, \;\; (4)$$

where $\pi(z) = \int_0^\infty \pi(z, v)\, dv$. By integrating (4) from zero to infinity and putting $z = 1$, one obtains the equation for the determination of $\pi(1)$. Its solution is $\pi(1) = [\beta(\lambda)]^{-1} - 1$ and thus $p_0 = 1 - \pi(1) = 2 - [\beta(\lambda)]^{-1}$. Next, by differentiating (4) and then integrating out x, one obtains the equation for the mean number of customers in the system N, which solution is $N = [1 - \beta(\lambda)]/[2\beta(\lambda) - 1]$.

Stationary sojourn time

As one is unable to use the Little's law without the prior check for the considered system, we now dwell on the stationary sojourn time distribution. Denote by $\chi(s)$ the LST of the customer's sojourn time and by $u(s)$ the LST of the system's busy period.

For the sake of simplicity we will consider only the special case of the PLCFS-Re service policy when $d(x, y|u, v) = b(x)b(y)$, because this allows one to obtain most of the expressions in explicit form. The general case of $d(x, y|u, v)$ can be handled in a similar manner, yet the LST $\chi(s)$ and $u(s)$ can be obtained only as solutions of certain functional equations.

Following again the approach from Meykhanadzhyan et al. (2014) and using the properties of the PLCFS-Re service policy one can verify that the LST of the busy period $u(s)$ satisfies the equation

$$u(s) = \beta(s + \lambda) + \frac{\lambda}{\lambda + s} [1 - \beta(s + \lambda)]\, u^2(s),$$

and its appropriate solution has the form

$$u(s) = \frac{\lambda + s - \sqrt{[\lambda + s]^2 - 4\lambda[1 - \beta(s+\lambda)]\beta(s+\lambda)[\lambda + s]}}{2\lambda[1 - \beta(s+\lambda)]}.$$

It can be verified that the busy period is finite with probability 1 (i.e. $u(0) = 1$) if and only if $\beta(\lambda) > 1/2$ and the mean length of the busy period $-\beta'(0)$ is finite if and only if $1/2 < \beta(\lambda) < 1$.

Denote by $\psi(s)$ the LST of the total time which the customer spends in service once it enters server (this includes all possible re-generations of its service time). Using the first step analysis it can be shown that $\psi(s)$ satisfies the equation

$$\psi(s) = \beta(\lambda + s) + \frac{\lambda}{\lambda + s}\psi(s)[1 - \beta(\lambda + s)],$$

and thus has the form

$$\psi(s) = \frac{(\lambda + s)\beta(\lambda + s)}{s + \lambda\beta(\lambda + s)}.$$

Remembering that customers from the queue are served according to LCFS order then, using the law of total probability, one obtains the following expression for the LST $\chi(s)$ of the customer's sojourn time distribution:

$$\chi(s) = p_0\psi(s) + (1 - p_0)\psi(s)u(s). \tag{5}$$

Further computation of the mean and higher moments (if they exist) of the stationary sojourn time is straightforward.

NUMERICAL RESULTS

In this section we demonstrate through numerical examples that in certain cases PLCFS-Re service policy is able to eliminate job-size estimation errors and produce good upper bounds for values of the mean sojourn time that are seen in the system with true job-sizes.

In order to be able to make comparisons we assumed that the true job sizes (service times), say S, have a Weibull distribution with the shape parameter k and the scale parameter α. Notice that such assumption is justified by the empirical measurements (see, for example, Dell'Amico et al. (2015) and references therein). Here we present results both for long-tailed and light-tailed service time distributions achieved by varying only the shape parameter k. The values of the scale parameter α were always set to guarantee that the true mean service time $\mathrm{E}(S)$ is equal to 1.

In the experiments we used the assumption, which has already been adopted in a number of recent papers on the topic (see Dell'Amico (2013); Dell'Amico et al. (2015)), when the job-size estimation error, say X, is log-normally distributed with zero mean and variance σ^2 and a job, having true size S, is estimated as $\hat{S} = SX$. Denote by $B(x)$ the cumulative distribution of the random variable $\hat{S}$. The mean $\mathrm{E}(\hat{S})$ of the distribution is always greater that 1 and as σ grows the sizes of the jobs are biased towards overestimating.

To summarize, one considers $M/G/1$ infinite capacity queueing system with Poisson arrivals of rate λ and i.i.d. service times with distribution $B(x)$. Assume that four different service policies are implemented in the system: Processor Sharing (PS), Shortest-Preemptive-Last-Come-First-Served (SPLCFS). First-Come-First-Served (FCFS) and PLCFS-Re. Denote by v^{PS}, v^{SPLCFS}, v^{FCFS}, $v^{\mathrm{PLCFS-Re}}$ the values of customer's stationary mean sojourn time under each of the four policies. We are interested in two aspects:

- understanding the relationships between $v^{\mathrm{PLCFS-Re}}$ and v^{PS}, v^{SPLCFS}, v^{FCFS} in the presence of estimation errors i.e. when $\sigma > 0$;
- understanding how close the values of $v^{\mathrm{PLCFS-Re}}$ are to the values of true stationary mean sojourn time, which we denote by[2] $v^{*\mathrm{PS}}$, $v^{*\mathrm{SPLCFS}}$ and $v^{*\mathrm{FCFS}}$.

As the considered model is of $M/G/1$ type, for the computation of v^{PS}, v^{SPLCFS} and v^{FCFS} one can use known analytic expressions.

According to the specification of the PLCFS-Re service policy (see Section 2), in order to use it one has to specify the function $d(x, y|u, v)$. The flexibility of the definition of $d(x, y|u, v)$ gives a plenty of choices but we will choose probably the most simple one. Let $d(x, y|u, v) = b(x)b(y)$ i.e. the service time of the newly arriving customer and the (remaining) service time of the customer in server (if any) are re-generated independently of their current values. Then the value of $v^{\mathrm{PLCFS-Re}}$ is equal to $-\chi'(0)$, which can be computed using (5). Notice that being defined in such a way, this policy is completely memoryless and not work conserving.

In Fig. 1, Fig. 2, Fig. 3 and Fig. 4 one can see the relationship between the values of $v^{\mathrm{PLCFS-Re}}$, v^{PS} and true mean sojourn time $v^{*\mathrm{PS}}$. Along the x-axis one shows the values of load $\lambda\mathrm{E}(S) = \lambda$ in the system without estimation errors. The values of mean sojourn time are given along the y-axis.

From Fig. 1 one can see that if the shape parameter is $k = 1$ (i.e. the true service time distribution is exponential) then $v^{*\mathrm{PS}} < v^{\mathrm{PLCFS-Re}} < v^{\mathrm{PS}}$ over the whole range of load values. In the presence of errors $v^{\mathrm{PLCFS-Re}}$ provides a good upper bound for the true mean sojourn time $v^{*\mathrm{PS}}$ under PS policy. Another interesting observation from Fig. 1 is the following. If one fixes the arrival rate, say $\lambda = 0.9$, then starting from a certain value of σ, one is already unable to compute v^{PS} because system's load exceeds 1 (in case $\sigma = 1$ and $\lambda = 0.9$ system's load is ≈ 1.48). But the value $v^{\mathrm{PLCFS-Re}}$ still can be computed and provides a good upper bound for the true value $v^{*\mathrm{PS}}$. Thus sometimes PLCFS-Re policy allows one to make estimations of the true values of mean sojourn time for wider ranges of system's load. Almost similar dynamics can be observed for higher values of k (see Fig. 2).

[2] These denote the values of customer's stationary mean sojourn time in case of no estimation errors i.e. when the service time distribution is Weibull.

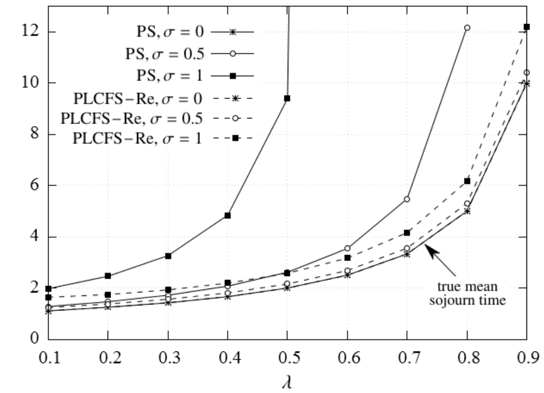

Figure 1: Mean sojourn time versus load for PS and PLCFS-Re service policies. True service time distribution is exponential ($k = 1$). The inequalities $v^{*\mathrm{PS}} < v^{\mathrm{PLCFS-Re}} < v^{\mathrm{PS}}$ always hold.

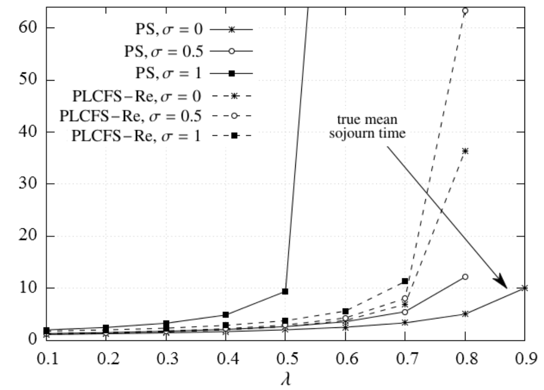

Figure 2: Mean sojourn time versus load for PS and PLCFS-Re service policies. True service time distribution is light-tailed ($k = 1.5$). For high values of estimation errors $v^{*\mathrm{PS}} < v^{\mathrm{PLCFS-Re}} < v^{\mathrm{PS}}$.

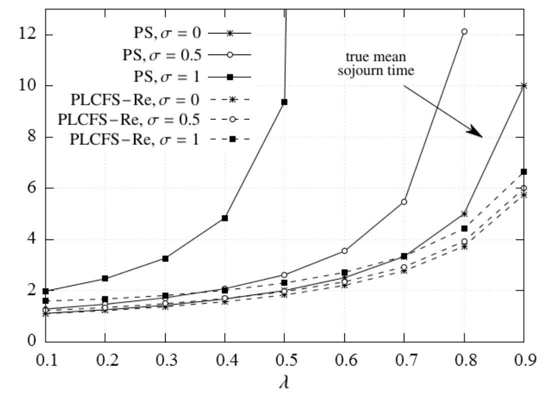

Figure 3: Mean sojourn time versus load for PS and PLCFS-Re service policies. True service time distribution is long-tailed ($k = 0.9$) with joint ratio $\approx 30/70$. The values of load for which $v^{*\mathrm{PS}} < v^{\mathrm{PLCFS-Re}} < v^{\mathrm{PS}}$ depend on the estimation error σ.

As the tail of the true service time distribution becomes heavier the situation changes. From Fig. 3 one can see that for $k = 0.9$ in the presence of errors the inequal-

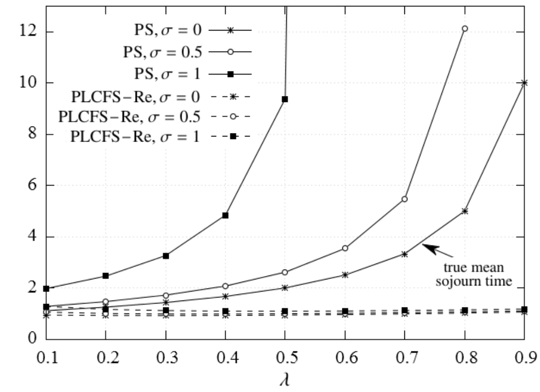

Figure 4: Mean sojourn time versus load for PS and PLCFS-Re service policies. True service time distribution is long-tailed ($k = 0.5$) with joint ratio $\approx 20/80$. PLCFS-Re with $d(x, y|u, v) = b(x)b(y)$ almost always underestimates the true values of the mean sojourn time.

ity $v^{\mathrm{PLCFS-Re}} < v^{\mathrm{PS}}$ almost always holds. But now the range of load values for which $v^{\mathrm{PLCFS-Re}}$ provides an upper bound for the true value $v^{*\mathrm{PS}}$ depends on the estimation error σ: the greater σ the wider range. Yet for quite low values of k (see Fig. 4) the values $v^{\mathrm{PLCFS-Re}}$ significantly underestimate mean sojourn time for PS service policy and are thus useless.

Qualitatively the above observations are valid for FCFS and SPLCFS policies as well (see Fig. 5 for SPLCFS policy).

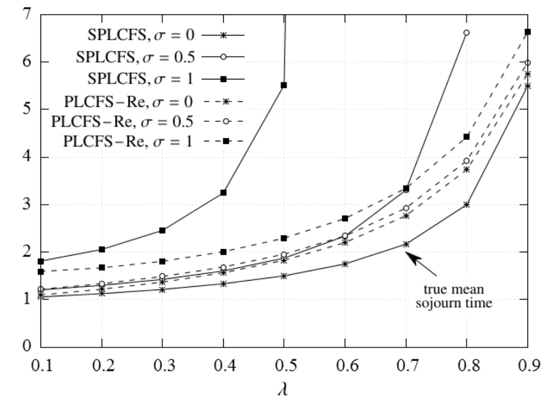

Figure 5: Mean sojourn time versus load for SPLCFS and PLCFS-Re service policies. True service time distribution is long-tailed ($k = 0.9$) with joint ratio $\approx 30/70$. PLCFS-Re with $d(x, y|u, v) = b(x)b(y)$ for *high* values of load provides better estimates of v than ordinary SPLCFS.

From Fig. 1–4 one can observe that for some values of k and σ the inequalities $v^{*\mathrm{PS}} < v^{\mathrm{PLCFS-Re}} < v^{\mathrm{PS}}$ are held and for others are not. More thorough study has revealed that these inequalities also depend on the value of λ (see Fig. 6).

According to our experiments the PLCFS-Re policy can provide good upper bounds for the true values of the mean sojourn time only if the service time distribu-

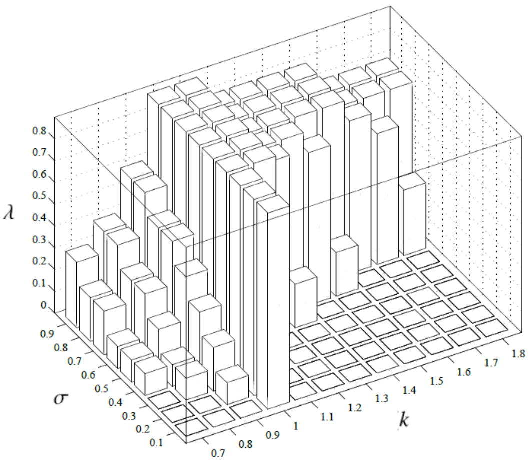

Figure 6: A visual presentation of the value range of λ, σ and k within which $v^{*\text{PS}} < v^{\text{PLCFS-Re}} < v^{\text{PS}}$ i.e. when PLCFS-Re policy always outperforms PS. The presence of the vertical bar and its height indicates the set of values of λ, σ and k for which $v^{*\text{PS}} < v^{\text{PLCFS-Re}} < v^{\text{PS}}$. The absence of the bar for the values of λ, σ and k indicates that PLCFS-Re policy is worse than PS. For the ease of presentation it is assumed that λ, σ and k take only discrete values (with step 0.1). We note that not the full range of possible values of λ, σ and k is displayed.

tion $B(x)$ is long-tailed and thus exhibits more or less the mass-count disparity[3]. If one chooses the joint ratio and median-to-median distance[4] as a characterisation of the service time distribution $B(x)$ then from Table 1 it can be seen that PLCFS-Re policy is useful (i.e. guarantees that $v^{*\text{PS}} < v^{\text{PLCFS-Re}} < v^{\text{PS}}$) whenever the joint ratio is approximately less than 32. But even for joint ratio < 32 the PLCFS-Re policy can underestimate the values of the true mean sojourn time, because there is also a dependency on the values of λ (see Fig. 6).

As the mass-count disparity becomes stronger (i.e. the tail of the service time distribution becomes more pronounced) PLCFS-Re policy with $D(x, y|u, v) = B(x)B(y)$ provides very bad (highly underestimated) values of the true mean sojourn time for all three policies PS, SPLCFS and FCFS. But simple changes in the definition of the function $D(x, y|u, v)$ allow one to broaden the the range of values k, σ and λ for which PLCFS-Re policy is good. For example, fix $M > 0$ and let the $d(x, y|u, v)$ be equal to

$$d(x,y|u,v) = \begin{cases} b(x)\delta(y-v), & u > M, v < M, \\ b(y)\delta(x-u), & u < M, v > M, \\ \delta(x-u)\delta(y-v), & u < M, v < M, \\ b(x)b(y), & u > M, v > M, \end{cases} \quad (6)$$

[3]A small number of samples account for the majority of mass, whereas all small samples together only account for negligible mass.

[4]Joint ratio is the generalization of the Pareto principle: p% of customers account for (100-p)% of the service time, and (100-p)% of service times account for p% of customers. Median-to-median distance is the result of the division of the median of the mass distribution by the median of count distribution. See, for example, (Feitelson, 2015, Chapter 5).

where $\delta(x)$ is the Dirac delta function. According to Fig. 6 PLCFS-Re policy with $d(x, y|u, v) = b(x)b(y)$ is worse than PS in the presence of errors when $k = 0.7$, $\sigma = 0.9$ and $\lambda = 0.4$. But if one uses the new definition (6) for $d(x, y|u, v)$ then it holds[5] that $v^{*PS} < v^{PLCFS-Re} < v^{PS}$ if $M = 244$.

CONCLUSION

The function $d(x, y|u, v)$, governing the PLCFS-Re service policy, presents a sort of control over the system. Indeed, the solution of the system (2) leads to Fedholm's equations of the second kind and each of them can be seen as a dynamic programming equation (Bellman's equation) but without the control variable. Numerical results show that even simple control without any memory (i.e. when $d(x, y|u, v) = b(x)b(y)$) allows one to eliminate errors in the estimations of the job sizes and obtain better upper bounds for the true mean sojourn time in $M/G/1$ type systems with inaccurate job size information. Here by saying "better" we mean "better than the values of the mean sojourn time that can be obtained using ordinary work conserving policies PS, LCFS, FCFS, PLCFS and some others". But such simple control does not work for too long-tailed service time distributions, exhibiting strong mass-count disparity (for example, exhibiting 20/80 Pareto principle). But by changing the definition of the function $d(x, y|u, v)$ (for example, by introducing a threshold) one can broaden the ranges of mass-count disparity in which the PLCFS-Re service policy outperforms ordinary policies. Up to now we were unable to find a general rule which specifies when the PLCFS-Re service policy can provide a good upper bound in the presence of errors for the true mean sojourn time. For each combination of system's initial parameters one has to compare the values manually. But in the experiments we have observed such properties as linear order (with respect to the policies) and monotonicity with respect to initial parameters. Even though the considered model can be considered as oversimplified, it allows one to reach the problem from an analytic point of view. In the presence of long-tailed distributions, which complicate the simulation experiments, this can be seen as an advantage. We see the following further directions of research: searching for the best form of the function $d(x, y|u, v)$; checking the feasibility of the proposed approach for other performance characteristics (such as variance of the mean sojourn time) and multi-server systems; evaluation of PLCFS-Re service policy when other approach for modelling of job's inaccurate size is used (as discussed in Tsafrir et al. (2005)).

ACKNOWLEDGEMENTS

This work was supported by the Russian Foundation for Basic Research (grants 15-07-03007 and 15-07-03406).

[5]The exact values are $v^{*\text{PS}} = 1.66$, $v^{\text{PLCFS-Re}} = 1.73$, $v^{\text{PS}} = 3.75$.

Table 1: Values of the joint ratio and median-to-median distance for different values of k and σ. The scale parameter of the Weibull distribution is fixed and equal to $\alpha = \Gamma(1 + 1/k)^{-1}$ i.e. the mean of the Weibull distribution is 1. The values for which $v^{*\mathrm{PS}} < v^{\mathrm{PLCFS-Re}} < v^{\mathrm{PS}}$ i.e. when PLCFS-Re policy outperforms PS are in bold type.

	$k = 0.7$		$k = 1$		$k = 1.8$	
σ	joint ratio $p\%/(1-p)\%$	m-m dist.	joint ratio $p\%/(1-p)\%$	m-m dist.	joint ratio $p\%/(1-p)\%$	m-m dist.
0	26.55/73.45 [a]	4.87	31.92/68.08 [b]	2.42	38.68/61.32 [c]	1.38
0.1	26.49/73.51	4.91	**31.86/68.14**	**2.44**	38.48/61.52	1.39
0.2	26.31/73.69	5.05	**31.49/68.51**	**2.52**	38.01/61.99	1.43
0.3	25.89/74.11	5.33	**30.98/69.02**	**2.65**	37.39/62.61	1.51
0.4	**25.44/74.56**	**5.68**	**30.45/69.55**	**2.84**	36.43/63.57	1.63
0.5	**24.93/75.07**	**6.25**	**29.55/70.45**	**3.12**	35.13/64.87	1.78
0.6	**24.24/75.76**	**7.00**	**28.69/71.31**	**3.50**	33.91/66.09	1.99
0.7	**23.52/76.48**	**8.01**	**27.69/72.31**	**4.00**	**32.50/67.50**	**2.28**
0.8	**22.72/77.28**	**9.34**	**26.71/73.29**	**4.66**	**31.04/68.96**	**2.66**
0.9	**21.90/78.10**	**11.13**	**25.60/74.40**	**5.55**	**29.59/70.41**	**3.16**

[a] Long-tailed Weibull service time distribution with mean 1.
[b] Exponential service time distribution with mean 1.
[c] Light-tailed service time distribution with mean 1.

REFERENCES

Chang, H., Kodialam, M. S., Kompella, R. R., Lakshman, T. V., Lee, M., Mukherjee, S. 2011. Scheduling in MapReduce-like systems for fast completion time. In INFOCOM, Proceedings IEEE. New York. Pp.3074–3082.

Crovella, M. 2001. Performance Evaluation with Heavy Tailed Distributions. JSSPP '01 Revised Papers from the 7th International Workshop on Job Scheduling Strategies for Parallel Processing. Pp.1–10.

Dell'Amico, M. 2013. A simulator for data-intensive job scheduling. Tech. Rep. arXiv:1306.6023, arXiv.

Dell'Amico, M., Carra, D., Michiardi, P. 2015. PSBS: Practical Size-Based Scheduling. IEEE Transactions on Computers. Issue.99. Pp. 1–15.

Feitelson, D. 2015. Workload Modeling for Computer Systems Performance Evaluation. Cambridge University Press.

Harchol-Balter, M., Schroeder, B., Bansal, N., Agrawal, M. 2003. Size-based scheduling to improve web performance. ACM Transactions on Computer Systems (TOCS). Vol. 21. Issue 2. Pp.207–233.

Lu, D., Sheng, H., Dinda, P. 2004. Size-based scheduling policies with inaccurate scheduling information. Modeling, Analysis, and Simulation of Computer and Telecommunications Systems. (MASCOTS 2004). Proceedings. The IEEE Computer Society's 12th Annual International Symposium on. Pp.31–38.

Meykhanadzhyan, L. A., Milovanova, T. A., Pechinkin, A. V., Razumchik, R. V. 2014. Stationary distribution in a queueing system with inverse service order and generalized probabilistic priority. Inform. Primen. Vol. 8. Issue 3. Pp.28–38. (in Russian)

Qiao, Yi, Lu, D., Bustamante, F., Dinda, P. 2004. Looking at the server side of peer-to-peer systems. In Proceedings of the 7th workshop on Workshop on languages, compilers, and run-time support for scalable systems (LCR '04). ACM, New York, NY. USA. Pp.1–8.

Ren K., Kwon, Y., Balazinska, M., Howe, B. 2012 Hadoops adolescence: A comparative workload analysis from three research clusters. In Technical Report, CMU-PDL-12-106.

Tsafrir, D., Etsion, Y., Feitelson, D. G. 2005. Modeling user runtime estimates. In 11th Workshop on Job Scheduling Strategies for Parallel Processing (JSSPP), Springer-Verlag. Lect. Notes Comput. Sci. Vol. 3834. Pp. 1–35.

Wierman, A., Nuyens, M. 2008. Scheduling despite inexact job-size information. In ACM SIGMETRICS Performance Evaluation Review. Vol. 36. Issue 1. Pp.25–36.

AUTHOR BIOGRAPHIES

LUSINE A. MEYKHANADZHYAN is a PhD student at the department of applied probability and informatics of Peoples Friendship University of Russia. At present her research interests are focused on queuing-theoretic aspects of scheduling. Her email address is `lameykhanadzhyan@gmail.com`.

ROSTISLAV V. RAZUMCHIK received his Ph.D. degree in Physics and Mathematics in 2011. Since then, he has worked as a senior research fellow at Institute of Informatics Problems of the Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences (FRC CSC RAS). At present he holds the position of Head of the Information and Telecommunication System Modelling section at the FRC CSC RAS and associate professor position at Peoples' Friendship University of Russia. His current research activities are focused on queueing theory and its applications in performance evaluation of stochastic systems. His email address is `rrazumchik@ipiran.ru`.