BACHELOR

The blocked road problem

Sprenkels, L.N.

Award date:
2011

The Blocked Road Problem

Laura Sprenkels 0651434
Technische Universiteit Eindhoven
Bachelor technische wiskunde
Bachelor Project
Final Report

September 21, 2011

Contents

1 Introduction

The aim of this bachelor project is to develop a good waiting-time approximation of the so-called blocked road problem. There is a road with traffic coming from two directions and at one point in the road one of the two lanes is blocked for whatever reason. Now at this point only traffic from one of the two sides can pass. To avoid any accidents, traffic is regulated by two traffic lights, both with a maximum green time. A maximum green time means that the traffic light is green for a certain time, unless the lane is empty earlier, then the traffic light switches to red when there are no cars left.
To analyze this problem, the situation is modeled by a polling model. In a polling model one server serves several queues one by one with a certain server policy. For this particular situation there are only two queues, namely the traffic coming from both sides, and one server. Figure 1 shows a schematic reproduction of the model. Serving one customer is equivalent to letting one car pass the stop line. The arrivals of the cars from both sides are assumed to follow a Poisson process and the service time of the customers is fixed and the same for every car. The service policy of the server is k-limited, meaning the server will serve exactly k customers, unless the queue is empty earlier, then it will move on to the next queue. This k-limited policy is, in our model, equivalent to a time limited policy, because the service time for every car is assumed to be deterministic.
After the traffic light turned red at one queue and before it becomes green for the other queue, both traffic lights are red simultaneously for a moment. In the model this is called the switchover time, during which the server is completely idle and does not serve customers from either queue. The switchover time is also assumed to be deterministic.
For this project we are interested in the expected waiting time of the customers in both queues. In Section 2 we try to make the reading of this paper more comfortable by introducing some

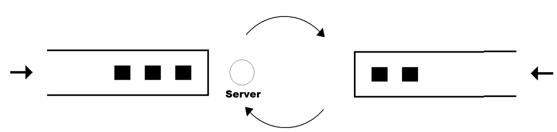

Figure 1: A two-queue polling model

notation, Section 3 is an interesting summary of all the basic queueing theory that we needed for this project, Section 4 is about why we did not find an exact solution and shows some interesting solutions for special cases were an exact solution does exist, Section 5 present a new approximation and compares it to already existing approximations and finally Section 6 will show how very well the approximation performs.

2 Notation and model description

In queueing theory there are many different notations used and there seems to be no standard way of denoting things. In this report the notation is based on a letter that stands for the quantity in question, say X. $F_X(t) = P(X \leq t)$ is the distribution of a continuous distribution and $f_X(t) = \frac{d}{dt}F_X(t)$ its density. $X(s) = \int_0^\infty e^{-ts}dF_X(t)$ denotes the Laplace-Stieltjes transform(LST) of X. In the discrete case $f_X(t) = P(X = t)$ and $F_X(t) = \sum_{k=0}^{t} f_X(k)$. Now

$X(z) = \sum_{n=0}^{\infty} z^n f_X(n)$ denotes the probability generating function (PGF) of X. The n-th moment of X is denoted by $E(X^n)$. In this report we assume service and switchover times to be deterministic, so the nth moment is equal to $E(X)^n$. In this case we write $E(X)$ as just X to make things look simpler.
Vectors, like $\mathbf{z} = (z_1, ..., z_n)$, are printed bold.

3 Queueing models

We will take a closer look at three more or less basic models and their waiting-time distributions. We start with an M/G/1 queue to see how to handle queues in general. Then we proceed with a more special M/G/1 queue, where the server goes on a holiday every once in a while. The reason why this model is concerned, is because it is a logical next step towards polling models, what our final model will be. In a polling model the vacation of a server, from the point of view of one queue, could be time spent serving the other queues and the switchover times.

3.1 M/G/1 queue

The M/G/1 queue is a system with one queue, where customers arrive according to a Poisson process with intensity λ, and with one server, serving the customers one by one with a generally distributed service time.
The total time for a customer spent in the system is the sum of the waiting time and the service time, which are independent of each other. So we can say the LST of the total time spent in the system by a single customer is the product of the LST of the waiting time and the LST of the service time. We know the LST of the sojourn time thanks to the *Pollaczek-Khinchin formula* $T(s) = \frac{(1-\rho)B(s)s}{\lambda B(s)+s-\lambda}$, where $B(s) = e^{-sB}$ is the LST of the service time sinds it is deterministic and $\rho = \lambda B$, the fraction of the total time the server is busy because the system is assumed to be in equilibrium. From this we can directly see the LST of the waiting time of a customer in the system, namely $W(s) = \frac{(1-\rho)s}{\lambda B(s)+s-\lambda}$, this is also a version of the Pollaczek-Khinchin formula [14]. For a derivation of this formula, see Section 3.3. Differentiating and substituting $s = 0$ shows that $E(W) = \frac{\rho}{1-\rho}\frac{B}{2}$.
It is also important to realize that once we know the PGF of the queue length after a departure we also know the LST of the sojourn time. When a customer leaves the system, the customers left behind are the ones that arrive during its sojourn time. Since the arrivals happen according to a Poisson process, the PGF of the number of customers arriving during a sojourn time is [1]

$$\begin{aligned}\sum_{n=0}^{\infty}\int_0^{\infty}\frac{(\lambda t)^n}{n!}e^{-\lambda t}f_T(t)dtz^n &= \int_0^{\infty}e^{-\lambda t}f_T(t)\sum_{n=0}^{\infty}\frac{(\lambda tz)^n}{n!}dt \\ &= \int_0^{\infty}e^{t\lambda(z-1)}f_T(t)dt \\ &= T(\lambda(1-z)). \qquad (1)\end{aligned}$$

The distribution of the queue length at the end of a service is the same as the distribution of the queue length at a random point in time. We know from the PASTA property, that

Poisson Arrivals See Time Averages, meaning the distribution of the queue length when a customer arrives is the same as the distribution of the queue length at a random point in time [15]. So we only have to prove that the queue length at an arrival and at a departure have the same distribution.
The number of customers in the system is a Birth and Death process, so if the system is in state n it can only hop to $n+1$ or $n-1$. Because the system is in equilibrium, we know the intensity from a state is the same as towards that state. The intensity from state n to state m is denoted with $\psi(n,m)$. So $\psi(0,1) = \psi(1,0)$ and $\psi(1,2) + \psi(1,0) = \psi(0,1) + \psi(2,1)$, therefore $\psi(1,2) = \psi(2,1)$. In general we can say the intensity from state n to state $n+1$ is the same as from state $n+1$ to n [14]. The number of times going from n to $n+1$ per time unit is λ times the probability of having n customers when a customer arrives. The number of times going from $n+1$ to n customers per time unit is the number of departures per unit time, also λ since the system is in equilibrium, times the probability of the queue having n customers in the system left just after a departure.
So the distribution of the queue length after a departure is the same as the distribution of the queue length at an arrival, which is the same as the distribution of the queue length at an arbitrary moment.

A busy period is a period the server is busy serving customers without any breaks. As will be seen in Section 3.3, it might be interesting to know the expected length of a busy period. Take the first customer arriving after the system was empty, the mean length of the busy period produced by this customer, $E(BP)$ is the sum of the length of its own service and the length of the busy periods produced by the customers arrived during this service, $E(BP)$, times the expected number of customers arriving during the service of the first, $\lambda * B$. Hence

$$E(BP) = B + \lambda * B * E(BP) = \frac{B}{1-\rho} \tag{2}$$

and so the mean number of customers served during one busy period is $\frac{1}{1-\rho}$.

3.2 Vacation model

In a vacation model, the system is the same as in the previous situation, except for the fact that every time the queue is empty, the server goes on a holiday for a certain time. We assume that the lengths of the holidays are deterministic. The PGF of the number of customers left behind by a random departing customer is the product of the PGF of the number of customers arriving during the residual vacation time and the PGF of the number of customers left behind by a random departure in the previous situation. [6]
Note that if we know this PGF, we also know the LST of the waiting time distribution. To understand why this is the case in a more intuitive manner, we follow the reasoning from[11]. In this article they shuffle the order of the customers, what is legitime because we are only looking at the queue *length*, so the position of each customer does not matter.

If a customer is in service, either it arrived during a vacation, then we call it a primary customer, or during the service of another customer, then it is a secondary customer. First the server serves a primary customer, then it serves all of the secondary customers arriving during the first service and during these services until there are no secondary customers left. Now the server serves the second primary customer and all of its secondary customers. The

busy period generated by a primary customer is the same as a normal busy period a customer in a M/G/1 queue would generate. This keeps repeating until the whole queue is empty and the server leaves for a break and allows the system to enter new primary customers.
Now pick a random departing customer, the customers left behind in the queue are the left primary customers, i.e., the number of customers that have arrived during the residual length of a vacation, and the customers that would be there when the system was a normal M/G/1 queue.

The distribution of the residual vacation time, $F_R(t)$, equals $\frac{1}{E(V)}\int_0^t (1 - F_V(x))dx$, with $F_V(x)$ the distribution of a vacation time. The probability of a customer arriving during a vacation of a certain length is linear with the actual length of a vacation, so it is Cxf_V. We would like $\int_0^\infty Cxf_V dx$ to be equal to one, which is only the case if $C = \frac{1}{E(V)}$. The arrival time of a customer, given an interval where the arrival takes place, is uniformly distributed. So the density of the residual vacation time after an arrival is

$$f_R(t) = \frac{1}{E(V)}\int_t^\infty \frac{1}{x} x f_V(x)dx = \frac{1 - F_V(t)}{E(V)}. \tag{3}$$

The PGF of the number of customers arriving during such a residual vacation time is $R(\lambda(1 - z))$. This is easily deduced by following the same reasoning as for the number of customers arriving during a sojourn time from Section 3.1.[1]

Since $W(s) = R(s)W_{M/G/1}(s)$ we have that $E(W) = \frac{E(V^2)}{2E(V)} + \frac{\rho}{1-\rho}\frac{B}{2}$ and if we now assume the vacations are deterministic

$$E(W) = \frac{E(V)}{2} + \frac{\rho}{1-\rho}\frac{B}{2}. \tag{4}$$

3.3 Polling model

As said before a polling model is a system with multiple queues and one server. There are numerous variations, all modeling a different situation. Variations can be made in, for example, the length of visit time per queue, the order of the queues visited by the server and the length of switchover times between queues. A typical cycle in a polling model with n queues can be seen in Figure 2.

The PGF of the joint queue length at the beginning, $Sb_i(\mathbf{z})$ and the completion $Sc_i(\mathbf{z})$

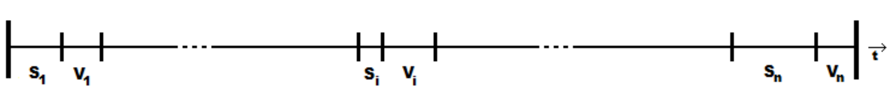

Figure 2: One cycle in a polling model

of the service of a customer at queue i are related through

$$Sc_i(\mathbf{z}) = \frac{Sb_i(\mathbf{z})}{z_i} B_i(\sum \lambda_j(1 - z_j)), \tag{5}$$

$\mathbf{z} = (z_1, ..., z_n)$ and $i \in \{1, ..., n\}$. This is because at queue i the number of customers at the service completion is the same minus the one customer being served plus the number of customers having arrived during the service. For every other queue the number of customers is the old number of customers and the additional customers who have arrived during the service.

Since the beginning of a service either coincides with a beginning of a visit or the end of a service. And since the end of a service either coincides with the end of a visit or the beginning of a visit, the following relation between different PGF's holds;

$$\frac{1}{E(L_i)} Vb_i(\mathbf{z}) + Sc_i(\mathbf{z}) = \frac{1}{E(L_i)} Vc_i(\mathbf{z}) + Sb_i(\mathbf{z}), \tag{6}$$

with L_i the number of customers served during one visit at Q_i and $Vb_i(\mathbf{z})$ and $Vc_i(\mathbf{z})$ resp. the PGF of the queue length at the beginning and ending of a visit period at queue i. [7][3] We now have two expressions for $Sb_i(\mathbf{z})$,(5)(6), so we can eliminate it, leading to

$$Sc_i(\mathbf{z}) = \frac{1}{E(L_i)} \frac{B_i(\sum \lambda_j (1 - z_j))}{z_i - B_i(\sum \lambda_j (1 - z_j))} (Vb_i(\mathbf{z}) - Vc_i(\mathbf{z})). \tag{7}$$

The marginal PGF of the number of customers after a departure in queue i is $Sc_i(\mathbf{z}_i)$, with $\mathbf{z}_i = (1, 1, ..., z_i, ..., 1)$. This is also the PGF of the marginal queue length at a random point in time as shown before.

If we apply this to the M/G/1 queue, we know that $Vb(z) = z$ because the service starts when there is one customer at the queue and we know $Vc(z) = 1$ because the service stops when the queue is empty, see Figure 3. The only thing missing is the mean number of customers served during one visit. The expected number of services during one visit, is the expected length of a busy period divided by the expected length of one service. This shows that

$$Sc(z) = (1 - \rho) \frac{B(\lambda(1 - z))}{z - B(\lambda(1 - z))} (z - 1) \tag{8}$$

and so

$$T(s) = Sc(1 - \frac{s}{\lambda}) = \frac{(1 - \rho) B(s) \frac{s}{\lambda}}{B(s) - 1 + \frac{s}{\lambda}}, \tag{9}$$

the Pollaczek-Khinchin formula.

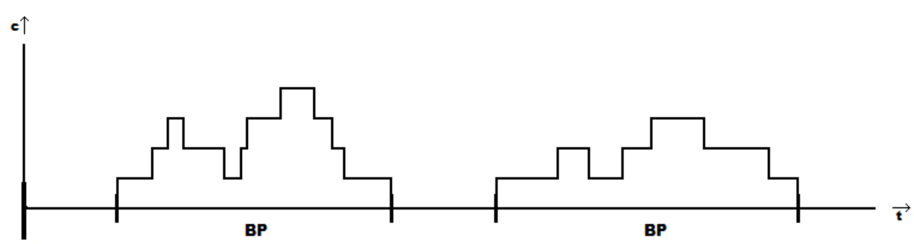

Figure 3: The Evolution of the number of customers, c, in an M/G/1 queue

So far we assumed that the number of customers at the beginning and ending of a visit are known, but this is not always the case. It is actually only the case for service disciplines that satisfy the so-called branching property.[13] This means that the customers present at queue i when the server arrives, can be replaced independently by a random number of other customers, with an n-dimensional PGF $h_i(s_1, ..., s_n)$.[12]

A very useful property to determine or estimate waiting times in polling systems with various service disciplines is the so-called pseudo conservation law (PCL). The PCL for a polling system with exhaustive service says that

$$\sum_{i=1}^{n} \rho_i E(W_i) = \rho \frac{\sum_{i=1}^{n} \lambda_i E(B_i^2)}{2(1-\rho)} + \rho \frac{E(S^2)}{2S} + \frac{S}{2(1-\rho)}(\rho^2 + \sum_{i=1}^{n} \rho_i), \tag{10}$$

where $\rho = \sum_{i=1}^{n} \rho_i$ and $S = \sum_{i=1}^{n} S_i$. For every other service discipline holds that $\sum_{i=1}^{n} \rho_i E(W_i) = rhs + \sum_{i=1}^{n} E(U_i)$, with rhs the right hand side of equation (10) and U_i the unfinished work left at queue i when the server leaves that queue for the rest of the cycle. So if the expected unfinished work per cycle per queue is known, a weighted sum of the mean waiting times is known in the system.[4]

4 Exact solutions for special cases

4.1 One-limited, one-limited

In this model there are no switchover times and both queues are served k-limited, where $k = 1$. This means that if there is a customer in the first queue, it will be served. If this is done and there are customers present at the other queue, the first customer of the second queue will be served, if there are no customers in the other queue and there still are customers in the first queue, another customer of type one will be served. If there are no customers in both queues the system is idle.
This can be modeled by a Markov chain with states $(Q_{1,n}, Q_{2,n}, C_n)$. $Q_{i,n}$ represents the number of customers in queue i after the nth departure and C_n is 1 if the nth customer left the first queue or 2 when it left the second queue. $\pi(q_1, q_2, c) = \mathrm{limP}[(\mathrm{Q_{1,n}, Q_{2,n}, C_n}) = (\mathrm{q_1, q_2, c})]$ is the probability the chain is in a certain state after a departure and $p_c(z_1, z_2) = \sum_{q_1} \sum_{q_2} \pi(q_i, q_2, c) z_1^{q_1} z_2^{q_2}$,$c \in \{1, 2\}$ is its generating function. $p_0 = p_1(0,0) + p_2(0,0)$ is the probability that there are no customers in both queues.
When the system is in equilibrium

$$p_1(z_1, z_2) = \frac{B_1(\lambda_1(1-z_1) + \lambda_2(1-z_2))}{z_1}(p_2(z_1, z_2) - p_2(0, z_1) + p_1(z_1, 0) - p_1(0,0) + p_0 \frac{\lambda_1}{\lambda_1 + \lambda_2} z_1) \tag{11}$$

and $p_2(z_1, z_2)$ can be expressed in an analog way. After a few manipulations two equations arise in terms of $p_1(0, z_2)$, $p_2(z_1, 0)$ and $p_1(0, 0)$. The general idea behind solving this is to find combinations (z_1, z_2) that make this function zero. These (z_1, z_2) form a closed curve in the R^2 plane. One of the functions is smooth on the inside of the curve and the other function on the exterior of the curve, but the function has to be smooth on the curve itself as well. This gives a boundary value problem that can be solved.[5]

4.2 k-Limited, exhaustive

In this model one of the two queues is served in an exhaustive manner and the other is served k-limited. To analyse this model, a Markov chain with states $(Q_{1,n}, Q_{1,n}, C_n)$ is used. $Q_{i,n}$ is the same as in Section 4.1 and C_n is zero when the last departure was from the first queue and $j = 1, .., k$ when the last departure was the jth successive departure from the second queue. $\pi(q_1, q_2, c) = \lim_{n\to\infty} P[(Q_{1,n}, Q_{2,n}, C_n) = (q_1, q_2, c)]$ is the probability the chain is in a certain state after a departure and $p_j(z_1, z_2) = \sum_{q_1} \sum_{q_2} \pi(q_i, q_2, c) z_1^{q_1} z_2^{q_2}, j \in \{0, 1, ..., k\}$ is its generating function, just like in the previous section.
Because the system is in equilibrium we can deduct $k + 1$ balance equations for the $k + 1$ unknown functions. By some heavy calculus these $k+1$ functions can be reduced to expressions in c_l, the probability that there are $l \in \{1, ..., k - 1\}$ customers waiting in the second queue at the start of a serving of queue two. These c_l's can be found by using *Rouché's theorem*, resulting in the PGF's.[8]

4.3 k-limited, k-limited

The previous two examples don't directly make clear why the k-limited, k-limited case is very difficult to solve exactly. Putting the problem in a Markov model with states $(Q_{1,n}, Q_{2,n}, C_n)$, where C_n is $j = 1, ...k_1$ if the nth departure was the jth in a row departing from queue one and is $j = k_1 + 1, ..., k_1 + k_2$ if the nth departure was the $j - k_1$th customer in a row departing from queue two.
Now the balance equations look like

$$p_1(z_1, z_2) = \frac{B_1(\lambda_1(1 - z_1) + \lambda_2(1 - z_2))}{z_1}(p_{k_1+k_2}(z_1, z_2) - p_{k_1+k_2}(0, z_2) + \sum_{l=k_1}^{k_1+k_2-1}(p_l(z_1, 0) - p_l(0, 0)) + p_0 \frac{\lambda_1}{\lambda_1 + \lambda_2} z_1) \qquad (12)$$

$$p_j(z_1, z_2) = \frac{B_1(\lambda_1(1 - z_1) + \lambda_2(1 - z_2))}{z_1}(p_{j-1}(z_1, z_2) - p_{j-1}(0, z_2)), \qquad (13)$$

with $j \in \{2, ..., k_1\}$. And analog for departures from the other queue.
These equations do not look much more difficult than the equations in the previous example. The difference is that we can not find nice expressions for the c_l's from Section 4.2. This is because in the previous case we knew the first queue was empty at the beginning of a service turn at the second queue, because it was served with an exhaustive service policy, but in this case there can be any number of customers waiting in the other queue. Therefore there are infinitely many possibilities and one cannot find the steady-state queue length distributions.

5 Approximations

A technique that can be used for approximating behavior in queueing systems is LT-HT interpolation. LT stands for light traffic and says something about how the system will behave when the arrival intensity of customers is very low. HT stands for heavy traffic and explains how the system behaves when the traffic intensity is very high, actually at the moment the system becomes instable.
Determining the waiting time in light traffic, there will be no problem, but since the waiting

time at the unstable queue goes to infinity when λ_i becomes big, trying to find an equation for the heavy-traffic limit of just the waiting time is not enough.

The utilization of a queue i is defined as $u_i = \rho + \frac{\lambda_i}{k_i}S$. If for all the queues in the system it holds that $u_i < 1$, the system is stable. The utilization is a lineair function of λ_i, so multiplying the waiting time of a queue by $1 - u_i$ might be a way to tame the asymptotic behavior of the queue when u_i goes to one. To verify this, we use the PCL for k-limited service.

$$\sum_{i=1}^{n} \rho_i(1 - \frac{E(L_i)}{k_i})E(W_i) = \frac{\rho}{2(1-\rho)} \sum_{i=1}^{n} \lambda_i B_i^2 + \frac{S}{2(1-\rho)} \sum_{i=1}^{n} \rho_i(1-\rho_i) + \frac{S}{1-\rho} \sum_{i=1}^{n} \frac{\rho_i^2}{k_i} - \sum_{i=1}^{n} \frac{\rho_i(1-\rho_i)}{2} \frac{E(L_i(L_i-1))}{\lambda_i k_i}. [9] \quad (14)$$

This is a more simple version of the PCL then the one in the article of Everitt, but our case is more simple than the general case because we use deterministic service and swichover times. Looking for $E(L_i)$ we can say that on average $(1-\rho)E(L_i)$ customers will arrive during the time the server is idle, but during this time there will also arrive $\lambda_i S$ customers in queue i, so $E(L_i)$ must be $\frac{\lambda_i S}{1-\rho}$. The second factorial moment of L_i, $E(L_i(L_i-1))$ is generally unknown, but finite. Only when the utilization of a queue tends to one, this second factorial moment is known; $E(L_i(L_i-1)) \to k_i(k_i-1)$.
If we multiply both sides of the PCL with $\frac{1-\rho}{\rho}$, the equation is

$$\sum_{i=1}^{n}(1-u_i)E(W_i) = \frac{1}{2} \sum_{i=1}^{n} \lambda_i B_i^2 + \frac{S}{2\rho} \sum_{i=1}^{n} \rho_i(1-\rho_i) + \frac{S}{\rho} \sum_{i=1}^{n} \frac{\rho_i^2}{k_i} - \frac{1-\rho}{\rho} \sum_{i=1}^{n} \frac{\rho_i(1-\rho_i)}{2} \frac{E(L_i(L_i-1))}{\lambda_i k_i}. \quad (15)$$

The right hand side of this equation will be finite if $u_i \to 1$ and therefore we study the HT limit of $(1-u_i)E(W_i)$ instead of just $E(W_i)$.

5.1 A vacation model with k-limited service

We are going to approximate the mean waiting time of a single queue model with k-limited service, $k > 1$, with a polynomial of the form $E(W)_{app}(\lambda) = \frac{K_0 + K_1 u + K_2 u^2}{1-u}$. This is reasonable because it will show the same asymptotic behavior as the real waiting time as can be seen in (15). The reason a second degree polynomial is chosen as numerator in the approximation function is because we can find three conditions we would like $E(W)_{app}(\lambda)$ to satisfy.

In light traffic, the probability of having more than k customers in the system is almost equal to zero. Therefore the waiting time in light traffic in this system will be the same as the waiting time in a system with exhaustive service, which is known; $\frac{\lambda B^2}{2} + \frac{S}{2}$[2]. We also want $\frac{d}{d\lambda}E(W)_{app}(\lambda) = \frac{d}{d\lambda}(\frac{\lambda B^2}{2} + \frac{S}{2})$. Because of the fact the waiting time is strictly increasing as λ increases, we know that this derivative will add extra information about the behavior of the waiting time.
If the second derivative of the waiting time in light traffic was not equal to zero, we also could have demanded $\frac{d^2}{d\lambda^2}E(W)_{app}(\lambda)$ to equal this second derivative and we had improved the estimation even more.

In heavy traffic

$$\begin{aligned}
\lim_{u\to 1}(1-u)E(W) &= \lim_{u\to 1}[\frac{1}{2}\lambda B^2 + \frac{S}{2}(1-\rho) + \frac{S\rho}{k} - \frac{(1-\rho)^2}{2}\frac{1-k}{\lambda}] \\
&= \frac{B}{2}\frac{B}{B+\frac{S}{k}} + \frac{S}{2}\frac{\frac{S}{k}}{B+\frac{S}{k}} + \frac{\frac{S}{K}\frac{B}{2}}{B+\frac{S}{k}} - \frac{\frac{S^2}{k^2}}{B+\frac{S}{k}}\frac{k-1}{2} \\
&= \frac{1}{2}B\frac{B+\frac{S}{k}}{B+\frac{S}{k}} + \frac{1}{2}\frac{S}{k}\frac{S+B-S\frac{k-1}{k}}{B+\frac{S}{k}} \\
&= \frac{1}{2}(B+\frac{S}{k}) \qquad (16)
\end{aligned}$$

follows directly from (15). This is also the HT limit of a normal M/G/1 queue with service times of length $B+\frac{S}{k}$, like the switchover time of the cycle is cut into k pieces and every cut is placed behind a different service in the cycle. This makes sense since in heavy traffic we are almost sure k services will take place every cycle. To verify if it also converges in distribution to the distribution of the waiting time of an M/G/1 queue, namely the exponential distribution, we are going to simulate the second moment of the waiting times as well and compare it to the mean waiting time. If the coefficient of variation, cv, converges to one, this will support this presumption.
We want $\mathrm{lim_{u\to 1}E(W)_{app}(\lambda) = \frac{1}{2}(B + \frac{S}{k})}$, giving us the third and last equation to find all of the constants of $E(W)_{app}$.

5.2 Approximation by Zang and Vickson

To see how well the interpolation performs we compare it with a simulation and another approximation by Zang and Vickson [16]. This approximation is based on the fact that $E(Q) = E(Y) + \frac{\lambda}{2} + E(Q_{M/G/1})$, where Q is the number of customers in the queue, $Q_{M/G/1}$ is the same queue length but in a normal M/G/1 queue and Y is the number of customers at the beginning of a vacation. They find an approximation for the probability there are one or more customers in the system at the beginning of a vacation and the conditional probability of having a certain number of customers in the system given the fact there are more then zero customers at that moment.
After a couple of manipulations, they finally find the following approximation;

$$E(Y) \approx \frac{1-\rho}{1-\rho-\frac{\lambda S}{k}}(\frac{\lambda}{2} + E(Q_{M/G/1}))\frac{\lambda S}{k^{2-\frac{\rho^2}{2}}(1-\rho) - (k^{1-\frac{\rho^2}{2}}-1)\lambda S}. \qquad (17)$$

Since $E(W) = \frac{E(Q)}{\lambda}$, this leads to an approximation for the mean waiting time. Note that the formulas used here are simpler then the ones in the original article, because in our situation $E(S^2) = S^2$. We denote the approximation of the mean waiting time by Zang and Vickson by $E(W)_{zv}$

5.3 A two-queue model with k-limited service

A system with two queues is stable as long as for $i = 1, 2$ holds that $u_i = \rho + \frac{\lambda_i}{k_i}S < 1$. We call the instable queue the queue that becomes instable first. By first we meen the one of which the utilization reaches one first by increasing the λ_i's while keeping their ratio $\frac{\lambda_1}{\lambda_2}$ fixed. We call the other queue the stable queue. From now on we assume queue 2 to be the unstable queue, so $\frac{\lambda_1}{k_1} < \frac{\lambda_2}{k_2}$. We call λ_2^* the value for λ_2 when $u_2 = 1$, so $\lambda_2^* = \frac{1}{\frac{\lambda_1}{\lambda_2}B_1 + B_2 + \frac{S}{k_2}}$ and we call $\lambda_1^* = \frac{\lambda_1}{\lambda_2}\lambda_2^*$.

To estimate the mean waiting time in both queues we are going to try to find enough information about the mean waiting time in the queues to fit a polynomial of the form $E(W_i)_{app}(\lambda) = \frac{K_0 + K_1 u_i + K_2 u_i^2}{1-u_i}$. We find the information of the waiting time in light traffic again by looking at the equivalent polling model with exhaustive service and find

$$E(W_i)_{LT} = \frac{1}{2}(\lambda_1 B_1^2 + \lambda_2 B_2^2 + (1 + \lambda_i B_i)S).[2] \tag{18}$$

Again we want $E(W_i)_{app}(0) = E(W_i)_{LT}$ and $\frac{d}{d\lambda}E(W_i)_{app}(0) = \frac{d}{d\lambda}E(W_i)_{LT}$. Finding a third equation to fit the approximate mean waiting time happens in a different way for both queues. In the stable queue we are going to approximate the mean waiting time for $\lambda_1 = \lambda_1^*$. To do this, we treat the queue like it is a vacation queue with holidays of length $S + k_2 B_2$, the two switchover times and the expected time the server will spend at the instable queue. Now we can apply the LT-HT interpolation as in Section 5.1 and want $E(W_1)_{app}(\lambda^*)$ to equal the function value at λ_1^*.
Finding an HT limit for the unstable queue is difficult, since it can not be obtained easily from the PCL. The problem is we do not know what $E(L_1(L_1 - 1))$ and $E(W_1)$ are. One approximation found for the HT limit of the unstable queue is

$$\lim_{u_2 \to 1}(1 - u_2)E(W_2) = \frac{1}{2}(B_2 + \frac{S}{k_2}). \tag{19}$$

This feels unlogical, because especially when the traffic intensity rises the time spent at the other queue also rises and takes a significant part of the time the server is on his vacation. This term is not in the original HT-limit because it is possible that it is lost by scaling the waiting time with $1 - u_i$. We already deduced how much time we expect the server to spend at queue one per cycle and add this to the new HT limit. Now

$$\lim_{u_2 \to 1}(1 - u_2)E(W_2) = \frac{1}{2}(B_2 + \frac{S + B_1 * \frac{\lambda_1^*(S + k_2 B_2)}{1-\rho}}{k_2}). \tag{20}$$

We call the waiting times found by the first and second interpolation $E(W_2)_{app1}$ and $E(W_2)_{app2}$ respectively.

5.4 Approximation by Wang and Fuhrmann

The results from both variants of the interpolation are going to be compared with the approximation of Wang and Fuhrmann. They find for the mean waiting time in every queue approximately

$$E(W_i) = \frac{1 - \rho_i + \frac{\rho_i}{k_i}(1 - \frac{1}{1-\rho})}{1 - \frac{\lambda_i}{k_i}E(C)} E(R_i), \tag{21}$$

with $E(C) = \frac{S}{1-\rho}$ the expected cycle length and $E(R_i)$ the forward recurrence time of the cycle time at queue i at departure epochs. For $E(R_i)$ we do not have an exact expression

but they assume that $\frac{E(R_i)}{E(BC_i)} = \frac{E(R_j)}{E(BC_j)}$ for every $i, j \in \{1...n\}$, where $E(BC_i)$ is the expected length of a cycle given that there are k_i servings at queue i. Also the expression for $E(BC_i)$ is not exact, but they have found an approximation in existing literature; $\frac{k_i B_i + S}{1-\rho+\rho_i}$.
These equations, together with an approximation for the PCL, give an approximation for the mean waiting time in every queue, named $E(W_i)_{wf}$.[10]

6 Results

In this section the performance of the different approximations is analyzed. Since this project is originally about a blocked road, the situations we look at are based on different kinds of real traffic situations. It is reasonable to assume that in every situation the time in between departures of cars is about two seconds, so in every situation $B = 2$. It is also reasonable to assume that, in the two queue case, the switchover times are of the same length, since cars from both sides must pass the same blockage.

6.1 One queue

We are going to treat the vacation model as a simplification of the two-queue model. This means that during one vacation, another, imaginary, queue must be served and two switchover times take place. We are going to take a look at three situations, see Figures 4, 5 and 6.

Situation 1.1 $k = 5$, $S = 6$. The vacations in this situation are very short, this means that the blockage is short and the traffic intensity from the other queue is low.

Situation 1.2 $k = 7$, $S = 20$. In this situation k is big, what is reasonable in a situation the intensity from this side is big. The switchover time is intermediate, so either the number of services per cycle at the other queue is low or the blockage is not very long.

Situation 1.3 $k = 3$, $S = 30$. The vacations of this situation are very long and k is not so big. So in this situation the blockage is long and there probably will arrive more customers in the other queue per cycle.

In every Situation our approximation performs a lot better than the approximation of Zhang and Vickson. Where our case is off by a maximum of ten percent (in situation 1.2), the other approximation gives only a pretty good approximation in Situation 1.1 but even there it is off by 18 percent for one value of λ.
Our approximation is asymptotically correct, meaning it is exact when $u \to 1$. This is what we expect since the scaled waiting time in heavy traffic is known.
The coefficient of variation does converge to one in every situation. This could indeed mean that the distribution of the waiting time of a vacation model in heavy traffic is exponential.

6.2 Two queues

In the approximations where the waiting time of both queues is approximated, we can try to simulate the actual situation happening in case of a road block. In different situations road blocks can be long or short, the arrival of cars from both sides can be symmetric or not and the maximum number of cars allowed to pass every cycle can be different. We will take a look at four different cases. The results can be seen in Figures 7, 8, 9 and 10.

Util	$E(W)_{sim}$	$E(W)_{app}$	$\frac{E(W)_{sim}-E(W)_{app}}{0{,}01*E(W)_{sim}}$	$E(W)_{zv}$	$\frac{E(W)_{sim}-E(W)_{zv}}{0{,}01*E(W)_{sim}}$	cv
0.01	3,01	3,01	0,00	3,01	-0,07	0,58
0.1	3,07	3,08	-0,43	3,09	-0,85	0,57
0.3	3,28	3,39	-3,56	3,35	-2,26	0,58
0.5	3,84	4,11	-6,96	3,81	0,88	0,67
0.7	5,63	6,05	-7,40	4,98	11,63	0,82
0.9	15,98	16,52	-3,39	13,01	18,57	0,96
0.95	31,90	32,47	-1,81	28,18	11,66	0,99
0.98	79,52	80,44	-1,17	78,47	1,32	0,99
0.99	159,78	160,44	-0,32	165,00	-2,87	1,00

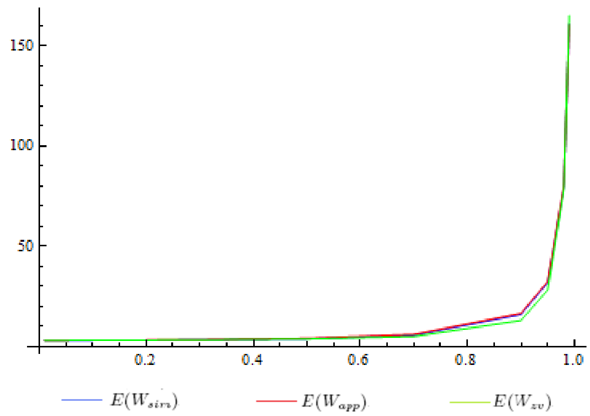

Figure 4: Situation 1.1

Util	$E(W)_{sim}$	$E(W)_{app}$	$\frac{E(W)_{sim}-E(W)_{app}}{0{,}01*E(W)_{sim}}$	$E(W)_{zv}$	$\frac{E(W)_{sim}-E(W)_{zv}}{0{,}01*E(W)_{sim}}$	cv
0.01	10,01	10,00	0,01	10,01	-0,07	0,58
0.1	10,04	10,07	-0,25	10,14	-0,99	0,58
0.3	10,17	10,44	-2,62	10,59	-4,13	0,57
0.5	10,69	11,42	-6,85	11,57	-8,26	0,58
0.7	12,98	14,25	-9,82	14,73	-13,51	0,64
0.9	28,25	30,04	-6,35	41,49	-46,88	0,84
0.95	52,32	54,23	-3,64	94,81	-81,19	0,91
0.98	124,73	127,02	-1,84	274,42	-120,01	0,97
0.99	246,69	248,43	-0,71	585,07	-137,17	0,98

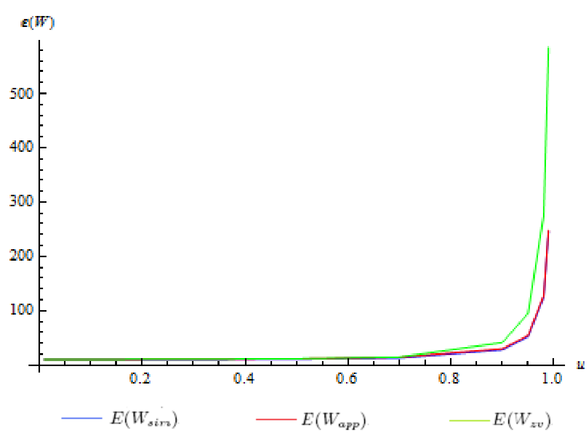

Figure 5: Situation 1.2

Util	$E(W)_{sim}$	$E(W)_{app}$	$\frac{E(W)_{sim}-E(W)_{app}}{0{,}01*E(W)_{sim}}$	$E(W)_{zv}$	$\frac{E(W)_{sim}-E(W)_{zv}}{0{,}01*E(W)_{sim}}$	cv
0.01	15,01	15,00	0,03	15,04	-0,25	0,58
0.1	15,05	15,08	-0,20	15,51	-3,02	0,58
0.3	15,77	15,82	-0,35	17,23	-9,28	0,59
0.5	18,20	18,08	0,63	21,12	-16,08	0,64
0.7	25,37	24,92	1,78	32,73	-29,04	0,74
0.9	64,69	63,75	1,45	107,64	-66,40	0,90
0.95	124,67	123,46	0,97	230,60	-84,97	0,95
0.98	304,06	303,28	0,26	607,95	-99,94	0,97
0.99	603,54	603,23	0,05	1240,31	-104,15	0,99

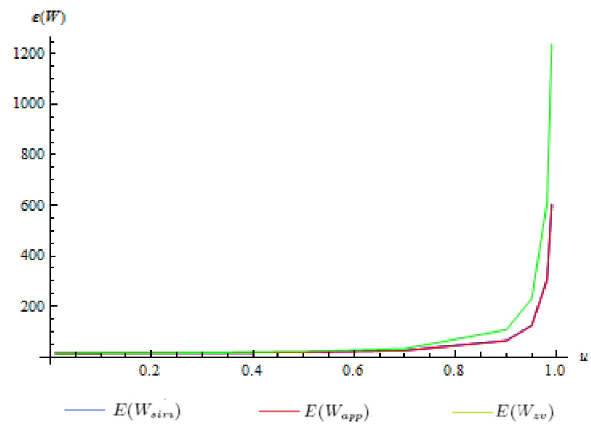

Figure 6: Situation 1.3

Situation 2.1 $\frac{\lambda_1}{\lambda_2} = \frac{1}{3}$, $S_1 = S_2 = 20$ and $k_1 = k_2 = 5$. In this situation the blockage is very long. The arrivals of customers is asymmetric, but the maximum number of cars passing the blockage per cycle is equal.

Situation 2.2 $\frac{\lambda_1}{\lambda_2} = \frac{1}{3}$, $S_1 = S_2 = 10$, $k_1 = 5$ and $k_2 = 7$. Here the blockage is about half the size as the one in the previous situation and in the more busy queue two more cars can pass every cycle.

Situation 2.3 $\frac{\lambda_1}{\lambda_2} = 1$, $S_1 = S_2 = 10$, $k_1 = 7$ and $k_2 = 6$. Now the situation is almost the same as situation 2.2, except now the cars arrive with the same intensity, what makes this case almost symmetrical.

Situation 2.4 $\frac{\lambda_1}{\lambda_2} = \frac{1}{3}$, $S_1 = S_2 = 4$ and $k_1 = k_2 = 3$. In this situation the blockage is very small, but also the maximum numbers of cars allowed to pass per cycle is smaller.

Overall we can conclude approximation 1 performs the very best out of the three approximations. For the stable queue it is not asymptotically correct because the value for $E(W_{app})(\lambda_1^*)$ is not exact, but is off by a maximum of five percent in Situation 2.2 and 2.3. In Situation 2.4 it does seem to converge to the simulated mean waiting time, in this case the utilization of queue 1 is relatively small compared to queue 2 so the value for λ_1^* is more accurate. For the unstable queue it is also not asymptotically correct because the HT limit is not exact. It beats approximation 2 in every Situation, so using the time spent in the first queue in the HT limit was not a good idea.
The approximation of Wang an Fuhrmann shows bad results in most cases, only in a few Situations for one particular value for the utilization it is off less then ten percent.

Util	$E(W)_{sim}$	$E(W)_{app1}$	$\frac{E(W)_{sim}-E(W)_{app1}}{0{,}01*E(W)_{sim}}$	$E(W)_{app2}$	$\frac{E(W)_{sim}-E(W)_{app2}}{0{,}01*E(W)_{sim}}$	$E(W)_{wf}$	$\frac{E(W)_{sim}-E(W)_{zv}}{0{,}01*E(W)_{sim}}$
Queue 1							
0.01	20,05	20,04	0,04	20,04	0,04	10,12	49,50
0.1	20,41	20,41	-0,03	20,41	-0,03	11,31	44,56
0.3	21,30	21,33	-0,11	21,33	-0,11	14,48	32,01
0.5	22,29	22,40	-0,48	22,40	-0,48	18,63	16,45
0.7	23,43	23,71	-1,19	23,71	-1,19	24,20	-3,27
0.9	24,81	25,38	-2,27	25,38	-2,27	31,98	-28,90
0.95	25,21	25,87	-2,63	25,87	-2,63	34,43	-36,56
0.98	25,46	26,19	-2,87	26,19	-2,87	36,01	-41,44
0.99	25,55	26,30	-2,95	26,30	-2,95	36,56	-43,12
Queue 2							
0.01	20,01	20,02	-0,02	20,02	-0,02	10,16	49,25
0.1	20,15	20,21	-0,27	20,21	-0,29	11,70	41,92
0.3	20,65	21,09	-2,16	21,15	-2,45	16,54	19,88
0.5	22,22	23,25	-4,64	23,48	-5,66	25,19	-13,36
0.7	27,65	29,22	-5,67	29,96	-8,36	45,25	-63,65
0.9	60,13	61,85	-2,87	65,53	-8,99	145,20	-141,49
0.95	110,41	111,68	-1,15	119,88	-8,58	295,01	-167,19
0.98	259,41	261,57	-0,83	283,40	-9,25	744,39	-186,96
0.99	513,56	511,54	0,39	556,09	-8,28	1493,33	-190,78

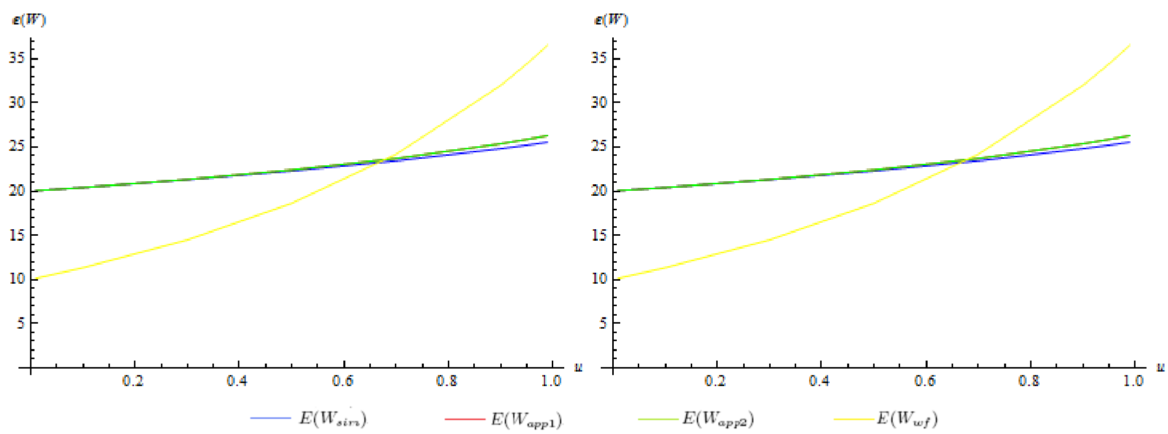

Figure 7: Situation 2.1

Approximation 1 performs the best in the first Situation, where S is very large. This is what we would expect, because if we look again at the PCL, if the S is large, the unknown terms are relatively small and therefor do not play an important role in the HT limit.
It performs the worst in situation 2.3. This is also what we would expect, since this is the most symmetric Situation. The approximation treats both of the queues very differently, while, in this case, their waiting times are very similar.
There is some room for improvement left, especially in the HT-limit of the unstable queue, but overall the results are very satisfying.

Util	$E(W)_{sim}$	$E(W)_{app1}$	$\frac{E(W)_{sim}-E(W)_{app1}}{0{,}01*E(W)_{sim}}$	$E(W)_{app2}$	$\frac{E(W)_{sim}-E(W)_{app2}}{0{,}01*E(W)_{sim}}$	$E(W)_{wf}$	$\frac{E(W)_{sim}-E(W)_{zv}}{0{,}01*E(W)_{sim}}$
Queue 1							
0.01	10,05	10,04	0,04	10,04	0,04	4,65	53,73
0.1	10,43	10,43	0,07	10,43	0,07	5,67	45,69
0.3	11,45	11,41	0,35	11,41	0,35	8,72	23,84
0.5	12,75	12,66	0,68	12,66	0,68	13,51	-6,00
0.7	14,48	14,42	0,41	14,42	0,41	21,65	-49,50
0.9	17,02	17,33	-1,78	17,33	-1,78	37,40	-119,68
0.95	17,87	18,42	-3,06	18,42	-3,06	43,77	-144,94
0.98	18,44	19,19	-4,08	19,19	-4,08	48,39	-162,46
0.99	18,64	19,47	-4,48	19,47	-4,48	50,09	-168,78
Queue 2							
0.01	10,02	10,02	0,04	10,02	0,04	5,26	47,51
0.1	10,18	10,20	-0,18	10,20	-0,23	6,34	37,74
0.3	10,61	10,82	-1,99	10,88	-2,56	9,65	9,00
0.5	11,42	12,06	-5,61	12,29	-7,65	15,47	-35,48
0.7	13,98	15,15	-8,34	15,91	-13,78	28,76	-105,64
0.9	30,09	31,19	-3,66	34,96	-16,19	94,22	-213,10
0.95	55,48	55,44	0,07	63,85	-15,09	192,09	-246,24
0.98	131,53	128,28	2,47	150,64	-14,53	485,55	-269,17
0.99	261,01	249,70	4,34	295,35	-13,15	974,60	-273,39

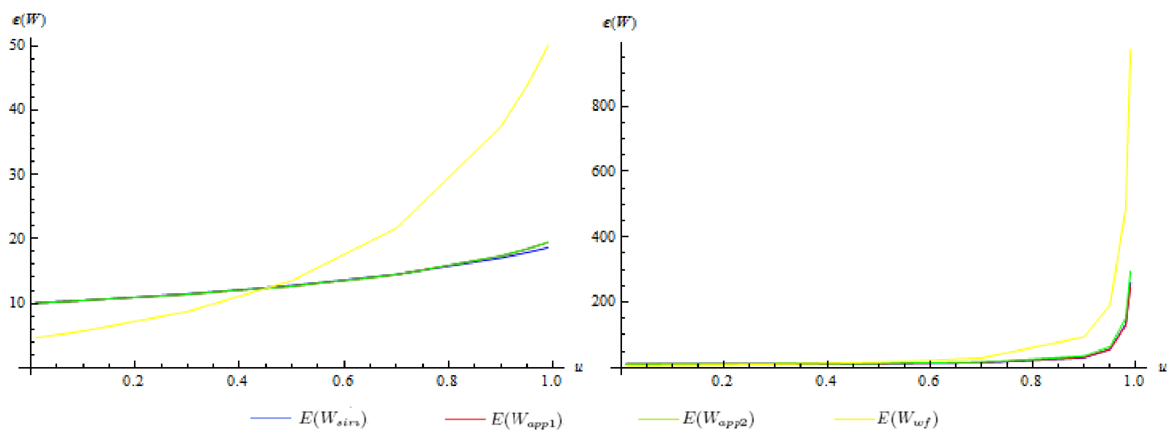

Figure 8: Situation 2.2

References

[1] I.J.B.F. Adan and J.A.C Resing. *Queueing Theory*. Eindhoven University of Technology, Eindhoven, 2001.

[2] M.A.A. Boon. *Polling models, from theory to traffic intersections*. PhD Thesis Technische Universiteit Eindhoven, Eindhoven, 2011.

[3] S.C. Borst and O.J. Boxma. Polling models with and without switchover times. *Operations Research*, (4), 1997.

[4] O.J. Boxma and W.P. Groendijk. Pseudo-conservation laws in cyclic-service systems. *Journal for Applied Probability*, (24), 1987.

[5] J.W. Cohen and O.J. Boxma. *Boundary Value Problems in Queueing System Analysis*. North-Holland Publishing Company, Amsterdam, 1982.

[6] B.T. Doshi. Queueing systems with vacations- a survey. *Queueing Systems*, (1).

Util	$E(W)_{sim}$	$E(W)_{app1}$	$\frac{E(W)_{sim}-E(W)_{app1}}{0{,}01*E(W)_{sim}}$	$E(W)_{app2}$	$\frac{E(W)_{sim}-E(W)_{app2}}{0{,}01*E(W)_{sim}}$	$E(W)_{wf}$	$\frac{E(W)_{sim}-E(W)_{zv}}{0{,}01*E(W)_{sim}}$
Queue 1							
0.01	10,03	10,03	0,00	10,03	0,00	5,26	47,55
0.1	10,34	10,35	-0,03	10,35	-0,03	6,36	38,51
0.3	11,18	11,21	-0,32	11,21	-0,32	9,72	13,07
0.5	12,35	12,50	-1,22	12,50	-1,22	15,42	-24,81
0.7	14,61	14,91	-2,05	14,91	-2,05	27,26	-86,63
0.9	22,09	22,37	-1,26	22,37	-1,26	66,97	-203,14
0.95	27,57	28,01	-1,60	28,01	-1,60	97,65	-254,23
0.98	33,33	34,38	-3,15	34,38	-3,15	132,52	-297,59
0.99	35,99	37,57	-4,38	37,57	-4,38	149,99	-316,75
Queue 2							
0.01	10,03	10,03	-0,04	10,03	-0,04	4,96	50,58
0.1	10,34	10,36	-0,12	10,39	-0,46	6,04	41,63
0.3	11,19	11,32	-1,19	11,74	-4,86	9,44	15,64
0.5	12,55	12,97	-3,35	14,57	-16,10	15,58	-24,17
0.7	15,87	16,65	-4,90	21,87	-37,83	29,94	-88,69
0.9	34,31	34,55	-0,70	60,47	-76,26	101,89	-197,00
0.95	63,41	61,24	3,43	119,00	-87,66	209,86	-230,94
0.98	153,90	141,26	8,21	294,93	-91,64	533,83	-246,88
0.99	308,23	274,60	10,91	588,23	-90,84	1073,78	-248,36

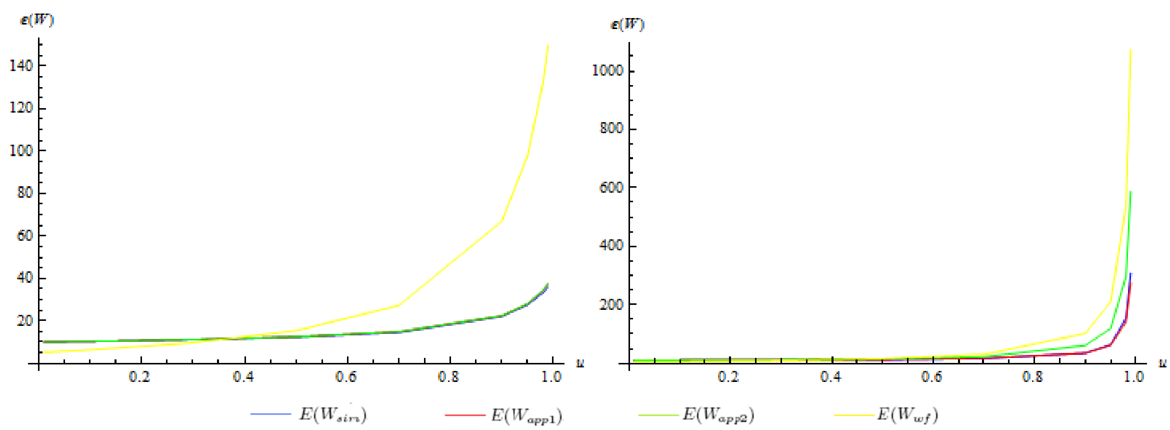

Figure 9: Situation 2.3

[7] M. Eisenberg. Queues with periodic service and changeover time. *Operations Research*, (2), 1972.

[8] I.J.B.F. Adan E.M.M. Winands and G.J. van Houtum. A state-depedent polling model with k-limited service. 2006.

[9] D. Everitt. A note on the pseudoconservation laws for cyclic service systems with limited service disciplines. *IEEE Transactions on communications*, (7), 1989.

[10] S.T. Fuhrmann and Y.T. Wang. Analysis of cyclic servie systems with limited service: Bounds and approximations. 1988.

[11] S.W. Fuhrmann. A note on the $M/G/1$ queue with server vacations. *Operations Research*, (6), 1984.

[12] S.W. Furhman. A decomposition result for a class of polling models. *IBM Research Report*, 1991.

[13] J.A.C. Resing. Polling systems and multitype branching processes. *Queueing Systems*, (13), 1993.

Util	$E(W)_{sim}$	$E(W)_{app1}$	$\frac{E(W)_{sim}-E(W)_{app1}}{0,01*E(W)_{sim}}$	$E(W)_{app2}$	$\frac{E(W)_{sim}-E(W)_{app2}}{0,01*E(W)_{sim}}$	$E(W)_{wf}$	$\frac{E(W)_{sim}-E(W)_{zv}}{0,01*E(W)_{sim}}$
Queue 1							
0.01	4,02	4,02	0,10	4,02	0,10	2,05	49,15
0.1	4,21	4,21	0,10	4,21	0,10	2,50	40,66
0.3	4,72	4,68	0,78	4,68	0,78	3,81	19,24
0.5	5,36	5,26	1,76	5,26	1,76	5,74	-7,14
0.7	6,20	6,05	2,47	6,05	2,47	8,77	-41,34
0.9	7,36	7,23	1,77	7,23	1,77	13,95	-89,50
0.95	7,72	7,64	1,07	7,64	1,07	15,85	-105,31
0.98	7,95	7,92	0,46	7,92	0,46	17,17	-115,92
0.99	8,03	8,01	0,20	8,01	0,20	17,64	-119,71
Queue 2							
0.01	4,01	4,01	-0,02	4,01	-0,02	2,05	48,96
0.1	4,11	4,13	-0,31	4,13	-0,43	2,51	39,07
0.3	4,55	4,60	-1,16	4,66	-2,48	3,93	13,58
0.5	5,62	5,67	-0,80	5,90	-4,95	6,45	-14,70
0.7	8,59	8,51	0,89	9,27	-7,98	12,25	-42,65
0.9	24,72	23,80	3,73	27,58	-11,56	41,07	-66,13
0.95	49,47	47,07	4,86	55,49	-12,16	84,26	-70,32
0.98	123,78	117,03	5,46	139,44	-12,65	213,81	-72,73
0.99	244,20	233,68	4,31	279,42	-14,42	429,72	-75,97

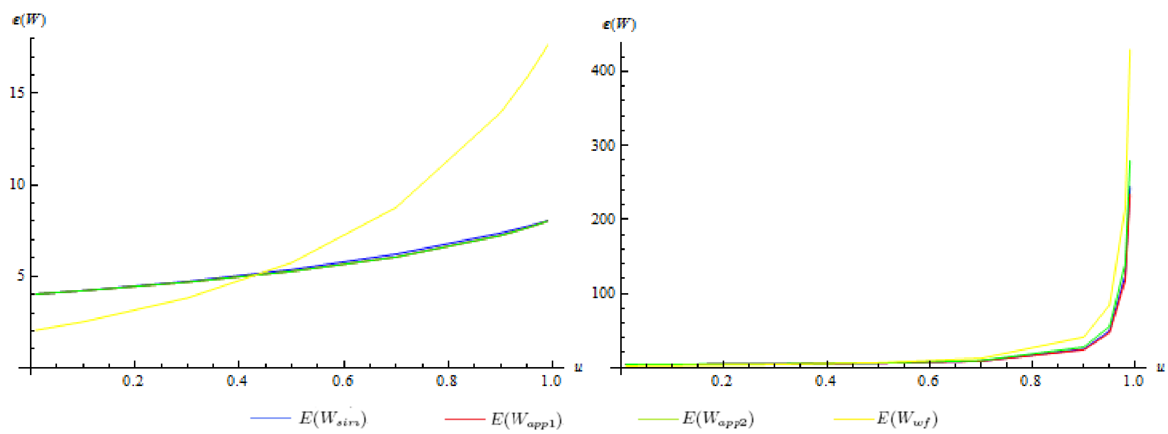

Figure 10: Situation 2.4

[14] S.M. Ross. *Probability Models.* Academic Press, Burlington, 2007.

[15] R.W. Wolff. Poisson arrivals see time averages. *Operations Research*, (2), 1982.

[16] Z. Zhang and R.G. Vickson. A simple approximation for mean waiting time in $M/G/1$ queue with vacations and limited service discipline. *Operation Research Letters*, (13), 1993.