Weighted Earliest Deadline Scheduling and Its Analytical Solution for Admission Control in a Wireless Emergency Network*

Jiazhen Zhou and Cory Beard
Department of Computer Science/Electrical Engineering
University of Missouri-Kansas City
5100 Rockhill Road, Kansas City, MO 64110, USA
Contact: beardc@umkc.edu

Abstract: For wireless networks operating during emergency conditions, emergency, handoff, and new calls are the main kinds of demand. Different operators will have different policies about which should have higher priority, and how to implement that kind of priority. This paper focuses on applying queuing methods for admission control for the above three types of calls. A Weighted Earliest Deadline scheduling method is introduced and evaluated that provides flexible priority for emergency and handoff traffic. Then an analytical framework for this scheduling method is given and two new computation methods are provided. These allow computation of expiration probability and average waiting time, and are simple to use both for priority queue scheduling and our new Weighted Earliest Deadline scheduling method.
Keywords: Emergency traffic, handoff, aging process, scheduling, priority queuing.

1. INTRODUCTION

Network congestion can happen due to a lot of reasons. In this paper, we mainly focus on disaster events as they cause congestion in wireless networks. After disaster events happen, tremendous stress is placed on networks due to the rise in traffic demand, including demand from general calls and emergency calls. As pointed out in [1,2,3], network demand can be up to 5 times of normal. Among the traffic demands, emergency traffic should be given higher priority for saving life and property.

In the wireless cellular network, an ongoing call can handoff to another cell. The goal of a cellular operator is to try to avoid calls being terminated due to lack of resources when moving into a new cell. So, the problem to be studied here is how to simultaneously and effectively support emergency users and handoff calls in a wireless network when congestion happens. The approaches we can use include preemption, delay-based, or resource conservation policies, and each has its advantages and disadvantages [4,5]. In this paper, we mainly focus on the delay-based approach which has been least studied. Our study focuses on a single cell and assumes the following: (1) all call durations are independently, identically, and exponentially distributed, (2) after service is completed, each call is terminated or leaves this cell, and (3) there is no handoff for emergency calls since we assume most emergency users will be stationary within a disaster area. However, for assumption (3), the model given here can be easily extended to a more

*This work is supported by the United States National Science Foundation under CAREER Award ANI-0133605.

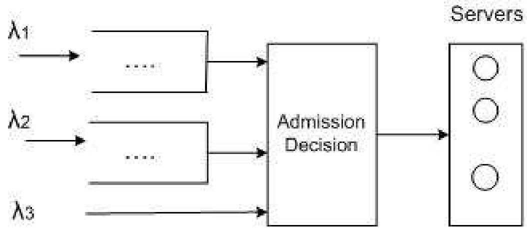


Figure 1. A Queuing Based Admission Control Policy in a Wireless Network

general situation.

A delay-based policy can also be called a queuing-based policy. The idea is that when congestion happens, new arrivals of high priority calls will be put into queues, and low priority calls are simply rejected. Figure 1 shows how queuing-based admission control can be implemented for 3 classes of traffic.

When resources become free, the calls waiting in queues will be served first. By using queues, the resource utility will be improved, and high priority calls will more likely be admitted than low priority ones, thus decreasing dropping probabilities for high priority calls. The cost of using this approach is the delay for the admission of high priority calls (because they need to wait in queues), and the increase of dropping rates for low priority calls. In contrast, if queues are not used when congestion happens, all classes of calls will have the same probability to be dropped regardless of the arrival rate of each class.

In [4], a delay-based policy is used for emergency calls without handoff calls considered, while in [6] queuing for handoff calls and new calls is introduced based on use of priority queues. In [7], two queues are used for different handoff calls, priority queuing is used as the basic tool, and calls are allowed to switch priorities. In this paper, we study the case when emergency calls, handoff calls, and general new calls coexist. We are not confined to using only priority queues; instead a flexible queue scheduling method is introduced.

The main contributions of this paper include:

1) Flexible ways to control the admission of emergency traffic and handoff traffic - For our study case, both emergency calls and handoff calls have priority over new calls. But which one has the highest priority will be interpreted differently by various operators. And even when one kind has higher priority over another, it's hard to say that one class should have absolute priority over another like priority queuing would enforce. So, based on the above facts, we bring out the Weighted Earliest Deadline scheduling method. Weighted Earliest Deadline chooses the next calls to be served based on a configurable weighting between the two classes.

2) Analytical representations for the new scheduling method - We provide 2-dimensional Markov chain representations of Weighted Earliest Deadline scheduling, and show how priority queuing is a special case.

3) Low complexity methods for solving the expiration probability and average waiting time in a 2-dimensional Markov aging process - In [6,7] priority queues are used, and they suggest using Mason's rule to compute the expiration probability for low priority calls. But for our case, using Mason's rule would be too complex and would have unacceptable computational

complexity, especially when queue length is long. We provide two simpler methods.

The rest of this paper is thus organized: In section 2, the Weighted Earliest Deadline scheduling method is brought out; in section 3, an analytical representation and new simple computation methods are provided. Section 4 provides evaluation of the effectiveness of the Weighted Earliest Deadline scheduling method; in section 5, an implementation example is given; and in section 6, we conclude this paper with a discussion of how these new analytical tools can be used in new areas of promising research.

2. WEIGHTED EARLIEST DEADLINE SCHEDULING

Since both emergency calls and handoff calls are important, both should be given priority compared with general new calls. In this paper, priority is implemented through using queues for emergency and handoff traffic in the following way: when there is no resource available upon arrival, new calls will be dropped, while a handoff or emergency call will be put into a queue for that type of call, unless the corresponding queue is full.

Those calls put into queues will not stay there for an unlimited time until they can gain access to a channel. Instead, it is assumed that users will be impatient and terminate the waiting call after some time. In other words, we consider the call to be "expired" when it is terminated due to impatience.

When there are two or more queues, a simple approach is to use priority queuing, which will give one queue strict priority over the other queue and a good analytical result can be obtained. But this absolute priority may create unfair resource allocation. To remove this kind of unfairness and to have flexibility in admitting two classes traffic, we introduce the *Weighted Earliest Deadline* scheduling method, which is modified from the Earliest Due Date (EDD) scheduling policy [9].

With this method, we consider the remaining lifetime (until expiration) and make a weighted comparison. For two classes, these weights can be simplified by normalizing one weight to 1, so all that is needed is one parameter, which is called the balance parameter and denoted as b_p. The scheduling rule is:

```
If b_p(EndureTime[1] - ElapsWaitTime[1]) < (EndureTime[2] - ElapsWaitTime[2])
   Choose customer in queue 1 to serve
Else
   Choose queue 2 to serve
```

The waiting calls to be compared are always taken from the heads of queues and all calls in a queue are assumed to have the same average endurance times. EndureTime means the maximum time this call can wait in the queue and ElapsWaitTime means the time already elapsed while waiting.

When $b_p = 1$, the result is EDD scheduling. If $b_p = 0$ or $b_p = \infty$, the result is priority queues. As b_p increases, then the likelihood of class 2 calls being chosen will also increase, and there will be one balance point where class 1 and class 2 calls will be chosen with equal likelihood. This is discussed in detail in section 3.

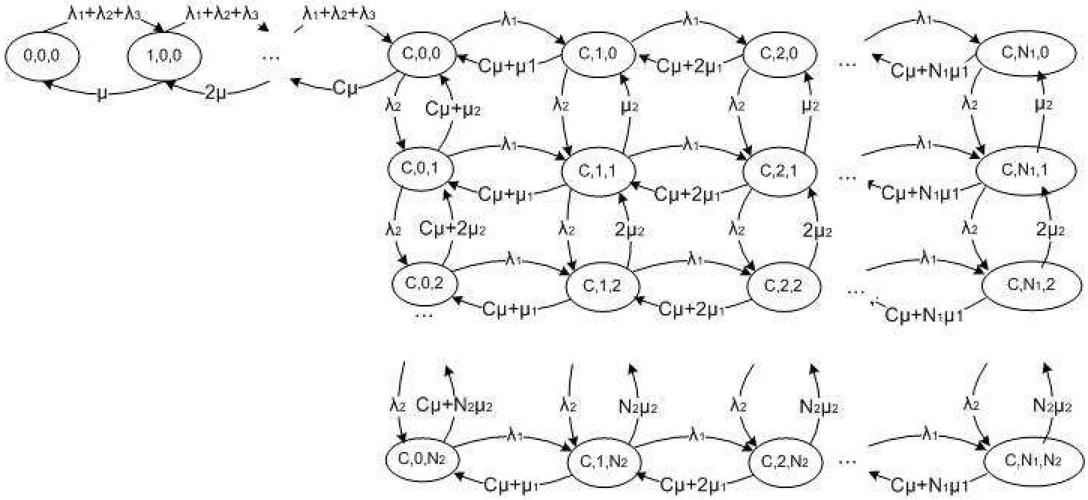


Figure 2. State Diagram for 3 Classes of Traffic using Priority Queues

3. ANALYTICAL MODEL FOR WEIGHTED EARLIEST DEADLINE SCHEDULING

We assume the arrival rate of each class of calls to be independently and exponentially distributed with the rate λ_1, λ_2, and λ_3 for emergency, handoff, and new calls respectively. Service times are also exponentially distributed and all classes have the same service rate μ. The expiration processes of emergency and handoff calls are also independently exponentially distributed with rate μ_1 and μ_2.

As already stated, for Weighted Earliest Deadline scheduling if we set $b_p = 0$ or $b_p = \infty$, then the result is a priority queue. This case can be easily described using a 2-dimensional Markov chain as shown in Figure 2 where the state (C, i, j) would correspond to C calls in progress, i and j calls waiting in their respective queues. But when $b_p \neq 0$ and $b_p \neq \infty$, the system can not be represented and analyzed using a priority queue anymore. The remainder of this section will discuss the representation and analysis for this general case.

3.1. Analytical framework for Weighted Earliest Deadline scheduling

For scheduling, the servers can be viewed as being split between two queues. Following is the analysis of how this split is represented.

We denote T_{l1}, T_{l2} as the remaining lifetime for class 1 and 2 calls. The probability for class 1 calls to be served is: Pr($b_p T_{l1} < T_{l2}$). Define $X_1 = b_p T_{l1}$, $X_2 = T_{l2}$. The remaining lifetime is exponentially distributed and thus has the memoryless property. It can be shown that X_1 is exponentially distributed with rate $\frac{\mu_1}{b_p}$ and X_2 is exponentially distributed with rate μ_2. So:

$$Pr(X_1 < X_2) = \frac{\frac{\mu_1}{b_p}}{\mu_2 + \frac{\mu_1}{b_p}} = \frac{\mu_1}{\mu_1 + b_p \mu_2} \qquad (1)$$

Let $A = \frac{b_p \mu_2}{\mu_1 + b_p \mu_2}$ ($A \geq 0$ and $A \leq 1$). We can conclude that the actual service rate for class 1 calls leaving the queue and entering into service is $C\mu(1 - A)$. This produces the state diagram for Weighted Earliest Deadline scheduling as shown in Figure 3. Note, for example, the differences in Figures 2 and 3 for the departures from state $(C, 1, 1)$.

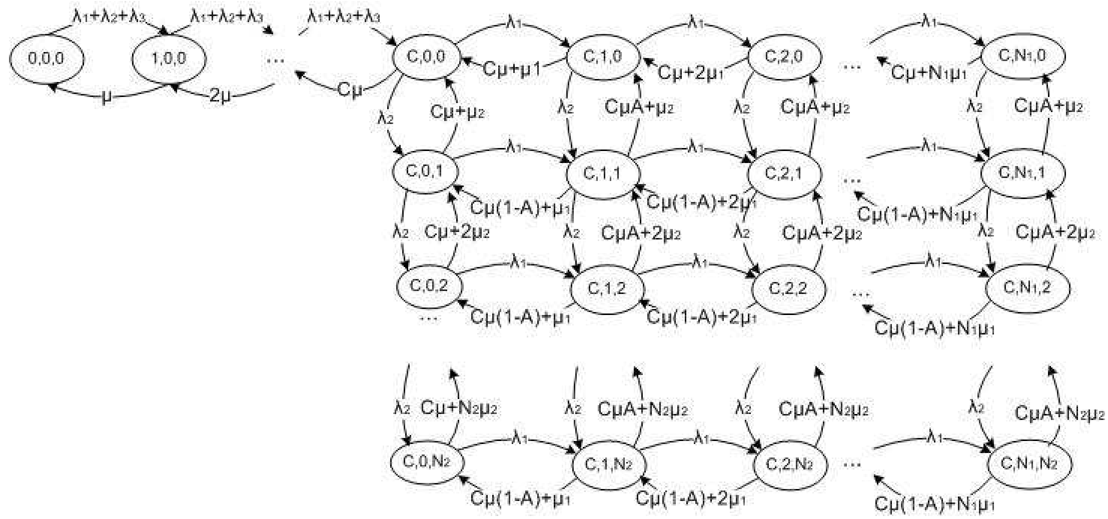


Figure 3. State Diagram for Weighted Earliest Deadline Scheduling

3.2. Computation of expiration probability and average waiting time

Due to limited queue length and limited endurance time for waiting, the calls can fail in two ways: immediately rejected (blocked) due to a full queue, or dropped after waiting some time in the queue (expired).

We denote the steady state probability of being in state (C, i, j) as $P(C, i, j)$. These can be computed as mentioned in [6]. Then, the blocking probabilities of the 3 classes are simple to obtain:

$$P_B^1 = \sum_{i=0}^{N_2} P(C, N_1, i) \qquad P_B^2 = \sum_{i=0}^{N_1} P(C, i, N_2) \tag{2}$$

$$P_B^3 = \sum_{i=0}^{N_1} \sum_{j=0}^{N_2} P(C, i, j) = 1 - \sum_{k=0}^{C-1} P(k, 0, 0) \tag{3}$$

For the computation of expiration probability for emergency and handoff calls, we provide two methods. The method given in section 3.2.1 is very simple, but the other is more complex but also more useful because it can also be used to determine waiting times, using a method similar to that in [6]:

$$P_{Exp}^1 = \sum_{i=0}^{N_1-1} \sum_{j=0}^{N_2} P(C, i, j) R_1(i+1, j) \qquad P_{Exp}^2 = \sum_{i=0}^{N_1} \sum_{j=0}^{N_2-1} P(C, i, j) R_2(i, j+1) \tag{4}$$

$R_1(i+1,j)$ means the expiration probability for the class 1 calls when they enter the system at state (C,i,j), which means this is for calls entered as the $(i+1)^{th}$ ones waiting in queue 1. Similarly, $R_2(i,j+1)$ means the expiration probability for the class 2 calls when they enter the system at state (C,i,j). With $R_1(i,j)$ and $R_2(i,j)$, we can also compute the average waiting time in queue, which is described in section 3.2.3.

In [6,7], the authors use priority queues and suggest a solution to $R_2(i,j)$ using the general gain formula (Mason's Rule). This requires formidable work especially when queue length is medium or long, which is difficult to build upon to develop our model for general scheduling. Instead, we provide a simpler approach to solve $R_1(i,j)$ and $R_2(i,j)$ for Weighted Earliest Deadline scheduling in section 3.2.2.

3.2.1. Probability Flow Computational Method for Expiration Probability

First the simpler approach. At a state (C,i,j), the arrival rate for each class is λ_1 or λ_2, and the expiration rate for each class is $i\mu_1$ and $j\mu_2$ independently. The average total number of departures per unit time from a state for expirations by class 1 is $i\mu_1$. Then, the probability of expiration is the fraction of expirations per unit time over arrivals per unit time, hence $\frac{i\mu_1}{\lambda_1}$. For class 2 calls this is $\frac{j\mu_2}{\lambda_2}$. Thus, instead of using the equations in (4) which as we will see require rather complex deduction, we can find the overall expiration probability just based on the steady state probability, the expiration rate, and the arrival rate at each state as follows:

$$P^1_{Exp} = \sum_{i=1}^{N_1}\sum_{j=0}^{N_2} P(C,i,j)\frac{i\mu_1}{\lambda_1} \qquad P^2_{Exp} = \sum_{i=0}^{N_1}\sum_{j=1}^{N_2} P(C,i,j)\frac{j\mu_2}{\lambda_2} \tag{5}$$

Clearly this is a simple approach, but it does not facilitate computation of average waiting times, so we introduce the other computational method in the next section.

3.2.2. State Transfer Based Computational Method for $R_1(i,j)$ and $R_2(i,j)$

Given there are i and j calls in queue 1 and queue 2 respectively, we denote $S_1(i,j) = 1 - R_1(i,j)$ as the probability for the calls at the tail of queue 1 to be *successfully* served before expiration, where i=1...N_1. Similar definition is given for $S_2(i,j) = 1 - R_2(i,j)$, and j=1...N_2.

Now we show how to solve $S_1(i,j)$ and $S_2(i,j)$. Assume that after a new class 1 call arrives, the system enters into a middle state (C,i,j) where both i and $j > 0$ and queues are not full. After this, the system can transfer to states $(C,i-1,j)$, $(C,i,j-1)$, $(C,i+1,j)$, or $(C,i,j+1)$ with probability $\frac{C\mu(1-A)+i\mu_1}{C\mu+i\mu_1+j\mu_2+\lambda_1+\lambda_2}$, $\frac{C\mu*A+j\mu_2}{C\mu+i\mu_1+j\mu_2+\lambda_1+\lambda_2}$, $\frac{\lambda_1}{C\mu+i\mu_1+j\mu_2+\lambda_1+\lambda_2}$, $\frac{\lambda_2}{C\mu+i\mu_1+j\mu_2+\lambda_1+\lambda_2}$ respectively.

When transferred to state $(C,i-1,j)$, the probability for the i^{th} call to expire before any channel is released and any other $i-1$ calls expire is $\frac{\mu_1}{C\mu+i\mu_1+j\mu_2+\lambda_1+\lambda_2}$. Thus the probability for the i^{th} call not to expire when transferred into $(C,i-1,j)$ is $\frac{C\mu(1-A)+(i-1)\mu_1}{C\mu+i\mu_1+j\mu_2+\lambda_1+\lambda_2}$

It is important to note that if the system is transferred into state $(C,i+1,j)$ due to a later arrival, the i^{th} call in queue 1 which we care about is still in position i, and the calls behind it will not affect its expiration probability at all. So, the successful probability for the call is still $S_1(i,j)$. Then, we can get:

$$
\begin{aligned}
S_1(i,j) &= \frac{C\mu(1-A)+(i-1)\mu 1}{C\mu+i\mu_1+j\mu_2+\lambda_1+\lambda_2}S_1(i-1,j)+\frac{C\mu*A+j\mu_2}{C\mu+i\mu_1+j\mu_2+\lambda_1+\lambda_2}S_1(i,j-1)\\
&+ \frac{\lambda_2}{C\mu+i\mu_1+j\mu_2+\lambda_1+\lambda_2}S_1(i,j+1)+\frac{\lambda_1}{C\mu+i\mu_1+j\mu_2+\lambda_1+\lambda_2}S_1(i,j) \quad (6)
\end{aligned}
$$

After simplification this becomes:

$$
\begin{aligned}
(C\mu+i\mu_1+j\mu_2+\lambda_2)S_1(i,j) &= (C\mu*A+j\mu_2)S_1(i,j-1)+\lambda_2 S_1(i,j+1)\\
&+(C\mu(1-A)+(i-1)\mu 1)S_1(i-1,j) \quad (7)
\end{aligned}
$$

When considering edge states, for the $i=1$ case we have:

(1) $j=0$

$$(C\mu+\mu_1+\lambda_2)S_1(1,0)=\lambda_2 S_1(1,1)+C\mu \quad (8)$$

(2) $j=1...N_2$-1

$$(C\mu+\mu_1+j\mu_2+\lambda_2)S_1(1,j)=\lambda_2 S_1(1,j+1)+C\mu(1-A)+(C\mu A+j\mu 2)S_1(1,j-1) \quad (9)$$

(3) $j=N_2$

$$(C\mu+\mu_1+N_2\mu_2)S_1(1,N_2)=C\mu(1-A)+(C\mu A+N_2\mu 2)S_1(1,N_2-1) \quad (10)$$

Thus we have N_2+1 equations, N_2+1 variables, and a right side not equal to 0. Obviously a unique solution can be obtained for each $S_1(1,j)$.

Then, for $i=2....N_1$, we have:

(1) $j=0$

$$(C\mu+i\mu_1+\lambda_2)S_1(i,0)=\lambda_2 S_1(i,1)+(C\mu+(i-1)\mu 1)S_1(i-1,0) \quad (11)$$

(2) $j=1...N_2-1$

$$
\begin{aligned}
(C\mu+i\mu_1+j\mu_2+\lambda_2)S_1(i,j) &= (C\mu*A+j\mu_2)S_1(i,j-1)+\lambda_2 S_1(i,j+1)\\
&+(C\mu(1-A)+(i-1)\mu 1)S_1(i-1,j) \quad (12)
\end{aligned}
$$

(3) $j=N_2$

$$
\begin{aligned}
(C\mu+i\mu_1+N_2\mu_2)S_1(i,N_2) &= (C\mu*A+N_2\mu_2)S_1(i,N_2-1)\\
&+(C\mu(1-A)+(i-1)\mu 1)S_1(i-1,N_1) \quad (13)
\end{aligned}
$$

Since each $S_1(i,j)$ only depends on $S_1(i-1,j)$, $S_1(i,j-1)$, and $S_1(i,j+1)$, these equations can be solved from $i=2$ to N_1 step by step. For each step of the computation, the matrix we need to construct is just $(N_2+1)*(N_2+1)$ in dimension.

Similar to the deduction for computing $S_1(i,j)$, for $S_2(i,j)$ we have:

$$
\begin{aligned}
(C\mu+i\mu_1+j\mu_2+\lambda_1)S_2(i,j) &= \lambda_1 S_2(i+1,j)+j\mu_2 S_2(i,j-1)\\
&+(C\mu+(i-1)\mu_1)S_2(i-1,j) \quad (14)
\end{aligned}
$$

The detailed equations for boundary cases are similar to those for computing $S_1(i,j)$, so these have been omitted.

With $S_1(i,j)$ and $S_2(i,j)$ solved using the above equations, we can apply equations in (4) to get the expiration probability. It has been verified that both this method and probability flow computational method produce the same results for blocking and expiration probabilities. This verification, however, was performed experimentally, since analytical verification methods are yet to be found.

3.2.3. Computation of Average Waiting time

Following a method similar to that in [6], based on the successful service probability at each state ($S_1(i,j)$ and $S_2(i,j)$) computed as in above subsection, it is possible to obtain the average waiting time for those calls which are successfully served.

Denote $W_1(i,j)$ as the expected waiting time taken by those class 1 customers that enter the queue at the i_{th} position and are later successfully served. It is obvious that

$$S_1(i,j) = Prob(t^1_{exp} > W_1(i,j)) \tag{15}$$

where t^1_{exp} is the actual expiration time for a particular class 1 call, which is exponentially distributed with the rate μ_1. So,

$$S_1(i,j) = Prob(t^1_{exp} > W_1(i,j)) = e^{-\mu_1 W_1(i,j)} \tag{16}$$

Thus we can obtain:

$$W_1(i,j) = \frac{-1}{\mu_1}\ln(S_1(i,j) \qquad W_2(i,j) = \frac{-1}{\mu_2}\ln(S_2(i,j))) \tag{17}$$

The average waiting times for class 1 and class 2 calls are calculated as:

$$\overline{W}_1 = \sum_{i=0}^{N_1-1}\sum_{j=0}^{N_2} P(C,i,j)W_1(i+1,j) \qquad \overline{W}_2 = \sum_{i=0}^{N_1}\sum_{j=0}^{N_2-1} P(C,i,j)W_2(i,j+1) \tag{18}$$

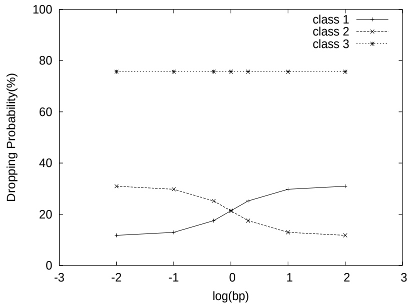
4. Evaluation of Weighted Earliest Deadline Using the Analytical Solution

Based on the analytical solutions provided above, the Weighted Earliest Deadline scheduling policy can be evaluated. We will see how good is this policy works and how the balance parameter b_p and different arrival rates affect the dropping rate and average waiting time.

We set the arrival rates for the three classes of traffic to 0.6, 0.6, and 0.2 respectively. The number of channels is 100; both queues' lengths are 10. The average call holding time for all 3 classes ise 100 seconds, and the average expiration time for emergency and handoff calls are both 10 seconds (in the examples later in this paper, these conditions hold if not specially specified). Then we change the value of b_p. From figure 4 we can see that when $b_p = 0$, the dropping probability for class 1 is the lowest. With b_p increases, the dropping probability for class 1 also increases, and for class 2 it decreases, while for class 3 there is almost no change. When $b_p = 1$, class 1 and class 2's dropping probabilities are equal, which shows the "balance" achieved. In contrast, if no queuing were used, all 3 classes would have a dropping probability of 30.12%.

In Figure 5, we can see that the effect of b_p on average waiting time has the same trend as its effect on dropping probability.

Figure 6 shows the effect as the arrival rate of class 3 increases, and it can be seen how dropping rates change for the 3 classes of calls. Here the average expiration time of class 2 calls is set at 5 seconds, and b_p is fixed at 1, so as expected the performance of emergency and handoff calls is different. The dropping probability of all 3 classes increases as the arrival rate of class 3 increase. And after some investigation we find that dropping probabilities all increase at the same percent, but the average waiting time of class 1 and class 2 (not shown here) didn't change with different λ_3 values.



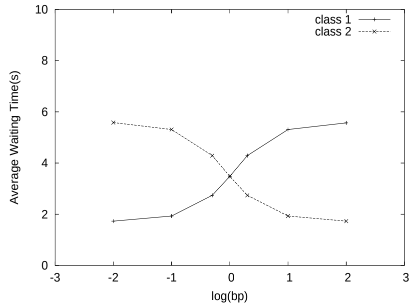


Figure 4. Dropping Probability vs. b_p Figure 5. Avg. Waiting Time vs. b_p

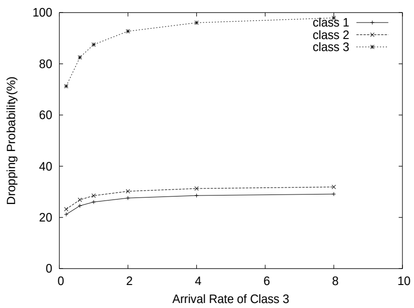
In Figure 7, we show the effect as the arrival rate of class 1 calls increases. The total number of channels is 60, and b_p is fixed at 1. We find that the average waiting time of class 1 and class 2 also increases as the arrival rate of class 1 increases. And the average waiting time for class 1 can even be more than the average expiration time of 10 seconds. This is because in a really congested case only those calls whose expiration time is very long can wait long enough until there is a free channel. Figures 6 and 7 show that good further research is possible to explain these phenomena analytically through exact or approximate calculations.

Through the above the experiments the following conclusions can be made about Weighted Earliest Deadline scheduling:

(1) The queuing method can help decrease high priority calls' dropping rates and ensure their priority over low priority calls.

(2) Through the balance parameter, the dropping probability and average waiting time of the two higher priority classes can be adjusted to get more ideal performance for a particular class, and the dropping probability of low priority calls is not affected. The balance parameter just adjusts the balance between the two priority classes.

(3) An increase in arrival rates of low priority calls will make all 3 classes' dropping probabilities increase in the same proportion, but will not affect the average waiting time of the 2 high priority classes.



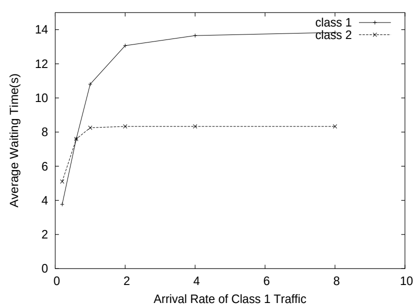


Figure 6. Dropping Probability vs. λ_3 Figure 7. Avg. Waiting Time vs. λ_1

(4) By using queuing itself, we cannot ensure that high priority calls will not be affected by changes in low priority traffic demand.

5. EXAMPLE OF APPLYING WEIGHTED EARLIEST DEADLINE SCHEDULING

With Weighted Earliest Deadline scheduling and the analytical solution that is provided, we can decide the appropriate balance parameter to satisfy a given system's requirements.

Example: Given are arrival rates for emergency, handoff, and new calls that are 0.2, 0.6, and 0.4 calls/sec respectively. The total number of channels is 100, and average call duration in each cell is 100 seconds. The average endurance times for waiting in queues are 10 seconds for both emergency and handoff calls. The requirement is to find the b_p that makes the dropping probability for emergency calls less than or equal to 5%, while also giving the best possible performance for handoff calls.

Solution: Set emergency calls as class 1 traffic, and handoff calls as class 2. Then given the characteristic illustrated in Figure 4, we find the value of b_p that gives a blocking of 5% for emergency calls. Then this will also give the best possible performance for the handoff class after meeting that constraint. The value of 0.1875 for b_p is the best value to satisfy these requirements, which results in blocking probabilities of 0.0497 and 0.0863.

6. CONCLUSION

This paper introduced Weighted Earliest Deadline scheduling to provide a good balance between different classes of wireless priority calls that are queued if they cannot first find a channel. The analytical solution for the expiration probability and average waiting time of this kind of scheduling is provided. The main novel contribution of this work is to provide an analytical framework for more flexible queue scheduling when using admission control.

We also find that in some cases the dropping probabilities of emergency and handoff calls are still not ideal by just using the queuing method alone. We have also not yet considered the relationships between different classes of calls and possible retrying after blocking. Promising possibilities for future work include: combining the queuing method with one or two other polices (e.g., using an Upper Limit policy [5] to set a limit on channels used for new calls), system cost optimization by considering both dropping rate and waiting time, and exploiting the relationship between handoff calls and new calls, blocking and retrying, etc. The model presented here, therefore, is a useful tool for answering many of the important questions about how to allocate resources to wireless calls during emergencies.

REFERENCES

1. Federal Emergency Management Agency, "Technology Applications by the Federal Emergency Management Agency in Response, Recovery, and Mitigation Operations." Paper Presented to the 27th Joint Meeting of the U.S./Japanese Panel on Wind and Seismic Effects, Tokyo/Osaka, Japan, May 16-27, 1995.
2. G. Philip and R. Hodge, "Disaster Area Architecture," *Proceedings of IEEE MILCOM*, 1995, pp. 833-837.
3. S. Adamson and S. Gordon, "Analysis of Two Trunk Congestion Relief Schemes," *Proceedings of IEEE MILCOM '93*, pp. 902-906, 1993.

4. C. Beard, "Preemptive and Delay-Based Mechanisms to Provide Preference to Emergency Traffic." *Computer Networks Journal*, Vol. 47:6, pp. 801-824, 22 April 2005.
5. C. Beard and V. Frost, "Prioritized Resource Allocation for Stressed Networks," *IEEE/ACM Transactions on Networking*, Vol. 6, no. 5,October 2001, pp. 618-633.
6. C.-J. Chang, T.-T. Su, and Y.-Y. Chiang "Analysis of a cutoff priority cellular radio system with finite queueing and reneging/dropping," *IEEE/ACM Transactions on Networking*. vol. 2, no. 2, pp.166-175, Apr. 1994.
7. A.E. Xhafa and O.K. Tonguz, "Dynamic priority queuing of handover calls in wireless networks: a analytical framework," *IEEE Journal on Selected Areas in Communications*, Vol.22, Iss.5, June 2004
8. D. Hong and S. S. Rappaport, "Traffic Model and Performance Analysis for Cellular Mobile Radio Tele- phone Systems with Prioritized and Non-prioritized Handoff Procediires," *IEEE Trans. on Vehicular Technology*, vol. VT-35, no. 3, pp. 77-92, Aug. 1986.
9. J. R. Jackson, "Scheduling a production line to minimize maximum tardiness," Res. Rep. 43, Management Sci. Rep., Univ. of Calif., Los Angeles, 1955.