Satellite Imaging Service Analysis Using Queueing Theory

Yan Mai & P. L. Palmer

Surrey Space Centre
University of Surrey, Guildford, GU2 7XH, UK
Tel +44 1483 259278, Fax +44 1483 259503
p.palmer@eim.surrey.ac.uk

Abstract

Earth observation using small satellites is leaving demonstration status and being proposed for more and more commercial applications. By analysing such a service from both end-user and satellite operator point-of-view, it is hoped to provide the information such as service performance, on-board resource status and key parameters for system optimisation before the spacecraft is designed, launched and put into service.

In this paper, queueing theory, traditionally used to perform traffic and efficiency analysis in tele-communication and other queueing system, is applied in this new area - a commercial imaging service using small satellites. The introduction of queueing theory in our application will eliminate the main difficulty of using the traditional solution - operation simulation which is huge computational complexity that arises when the operation spans a long period. In this paper, only the imaging download process, which is usually the system bottleneck for Low Earth Observation spacecraft, will be analyzed. Unlike traditional queueing systems where the service is continuous, our application suffers regular idle periods when the satellite is not visible from the ground-station for image download. The distribution and duration of such idle periods are the subject of orbital-dynamics. Therefore available queueing theory is not applicable directly and needs some extension to handle this specific problem of satellite imaging services. In this paper, three queueing models are discussed: M/G/1, M/G^a/1 and GI/G/1 together with the analysis of their suitability to our application. An extension to using M/G^a/1 is outlined which can provide a better approximation of the service than traditional queueing models. Some basic service parameters, such as queue length distribution, mean service occupation and mean service waiting time, can thereby be calculated. All results presented are compared with that from operation simulation. Limitation and constraints of using queueing theory in this application are also discussed. As a conclusion of this research work, it is shown that queueing theory will be appropriate for the early stage performance analysis in a quick but gross manner which can provide some basic performance parameters, while operation simulation can be treated as a refinement and a method capable of providing more complete solutions that will certainly take much longer time.

1 Introduction

More and more space applications are emerging whilst satellites tend to be smaller in size, faster in manufacturing and more enhanced in functionalities. Among these applications, Earth Imaging and Observation[1][2] is one of the most important. This application is leaving experimental and demonstrating status and entering commercial usage phase. Several proposals, such as, GANDA[3], Disaster Monitoring system[5][4], are at implementation phase and have attracted growing interest from all over the world. Because of this advent of the era of Earth observation using small satellite for commercial service, it is desirable to analyse the imaging service as an individual topic. This kind of analysis will hopefully provide the system designer with information such as Service performance, On-board resources management and furthermore, key parameters for system optimisation.

For a typical imaging system in commercial applications, the data flow is described in figures 1. Firstly, user submits his imaging request including satellite information, ground target latitude, longitude, illumination requirement and preferred imaging time window. The request or correspondent schedule file will be uploaded to the satellite. According to the schedule, satel-
Satellite will take image of the specific ground target. This is the first phase of satellite imaging service. Considering the small size of such request, even with limited system uplink capacity, it won’t lead to significant delay for satellite to obtain the request. Meanwhile camera on-board can be operated continuously and thus the performance of this phase is largely independent of the occurrence rate of user requests. The second phase starts from the end point of the first phase, that is, the time an image being stored in the onboard RAM disk. After this, image will be downloaded when the satellite is seen by a ground station as earlier as possible. The downloading discipline is generally according to their sequence of hitting the RAM disk. Since an image could contain huge amount of data and small satellite only has low-band width downlink, the second phase will be the bottleneck of the system and therefore will be the main focus of our analysis.

1.1 Simulate to analyse the imaging service

There are two methods to analyse the satellite imaging service. First one is the traditional method that we can directly simulate the system status by using either traditional satellite orbit propagator SGP4[6] or the satellite orbit estimation programme FPSCA[9][10], in which the latter is much faster than the previous one[8].

To carry out the simulation, we first generate a set of random numbers corresponding to the time when users put their requests. These random numbers follows Poisson distribution. The reason why we can make this assumption is based on the nature of user request and properties of Poisson distribution[11]. Then a set of random ground targets will be created, corresponding to every user’s request respectively. Without losing universality, we assume these targets follow uniform distribution.

After this, we select a satellite and a NORAD file[7], which provides the orbit parameters of the satellite. By using FPSCA, we can predict the image taken and downloaded time. When the program is running, we record system status in very small time interval. Finally, by post-processing these system states logging information, we can work out on-board image queue length distribution and service waiting time distribution, etc. The following figures show the changes of image queue length, that is, the number of images in RAM Disk, onboard satellite (POSAT) as time going. The x axis indicates the time elapsed after service started, while the y axis represents the image queue length. Figure 2 shows the system status when user request rate is 20 images/day, figure 3 shows user request rate in-
creases to 35 images/day. From these two figures, we can see that at this stage, the onboard system is still stable.

Figure 4: Satellite Onboard Image Queue Length: image arriving rate=40 images/day

The following figure 4 shows the queue length status when user request rate further increases to 40 images/day, the onboard image number starts to build up: after 50 days, the number of onboard images is up to 200. This continuously build-up curve indicates service is out of capacity and system is no long stable.

The figure 5 shows an even worse case - when user request rate is as high as 100 images/day, the number of onboard images is building up over 3000 after 50 days!

Figure 5: Satellite Onboard System Status: image arriving rate=100 images/day

From the above simulation results, we are able to get some idea about onboard system status and so a vivid estimation of the system capacity or limitation. However, such kind of simulation does take long time even by using our fast prediction method FPSCA.

1.2 Background of Queueing Theory

The second method is based on statistic mathematics, or more specifically, to use queueing theory to analyse the service.

A queueing system can be described as customers arriving for service, waiting for service if it is not immediately available, and if having waited for service, leaving the system after being served (see figure 6). Such a system can be subscribers' calls arriving at a telephone exchange, patients waiting in a doctor’s reception room, machines waiting to be serviced by repairmen, and cars waiting at a traffic intersection. This theory is based on statistic mathematics and was first brought forward almost a century ago. It has a couple of traditional applications including telephony traffic analysis, which was the first case of using queueing theory in solving real world problem[14]. Recently queueing theory has been widely used in computer performance analysis and network traffic optimisation[15]. Other less significant applications include machine repair efficiency analysis, taxi stands control and so on[16][17].

Figure 6: Description of a queueing system.

The behaviour of queues is principally affected by the following parameters: Arrival time distributions, e.g. memoryless(M), deterministic(D), Erlang k(Ek), general(G); Service time distribution, similar probabilities to the above; Service discipline, i.e. how waiting jobs are chosen for service and examples include first-in-first-out (FIFO), last-in-first-out (LIFO), service in random order (SIRO); Number of servers: single server or multiple servers; Waiting space: how many customers can be accommodated in the system (including those being served). In order to describe a queue model according to above specification, we can use groups of symbols and separate each group by a slash. The first group is Arrival time distribution, the second is service time distribution, the third is the number of server, the fourth is the waiting room size and so

Yan Mai
Although queueing theory is employed in wide spread applications, it is the first time to apply it in our case, satellite imaging service analysis.

2 Case Study of Directly Using Queueing Theory

2.1 Imaging service contains two sub-queues

In the satellite imaging service, which we propose to use queueing theory to analyze, the whole process is composed of two sub-queues which actually correspond to the two service phases respectively. (see figure 7).

![Figure 7: Satellite Imaging Service Data Flow - Two Subsequent Queues](image)

The first one, as mentioned previously, is the image taking process. It starts from user request and ends when image is taken and buffered in the RAM disk, see figure 7. This process is not the key part of our analysis using queueing theory because it is not the bottleneck of the system. In this process, request uploading is very simple and only consumes very little resources. Then image will be taken when the spacecraft is above the target. The only waiting time comes from that satellite need to rotate to be above the target, and in general, no request is at conflict with any other. So such waiting time can be precisely predicted using FPSCA program. On the other hand, since the service discipline in the process is not very explicit, a first come request may be serviced later, and vice versa. It is purely dependent on the satellite orbit and target location. The queue discipline is neither FIFO nor LIFO, which makes the application of queueing theory here more difficult. Based on all of these reasons, we will not discuss the details of the first queue in this paper.

The second queue in the whole service starts from image hitting the RAM disk and ends when it is finished being downloaded. This is the bottleneck of the system and it can not be predicted precisely. The waiting time and system length behaves randomly. User requests may be at conflict with each other. The queue discipline here is simply FIFO, that is, an image stored in ramdisk earlier will be downloaded to the ground before an image stored later.

2.2 User arrival in second sub-queue

As explained in last section, we will mainly study and discuss the second sub-queue. Firstly, we will concentrate on user arrival distribution. The user arrival in second queue is the image hitting the RAM disk. Then the first question is what kind of distribution it follows. Since user request in previous queue is Poisson or Memoryless and targets are assumed to distribute uniformly, the image hitting RAM disk should be independent of each other, and the probability of such an event should be the same in any fixed time intervals as long as the interval is the same, which is the characteristic of Poisson distribution. So it looks like the user arrival in second queue should still be Poisson distribution. Actually, simulation also proved so.

![Figure 8: Image Arrival RAM Disk Distribution: user request rate=20request/day](image)
happens following Poisson distribution, the interval of such occurrence is exponential distribution, which can be clearly seen here. Figure 8 shows the user arrival rate is 20 request/day, and figure 9 is for arrival rate at 100 request/day. We also tried other satellites and all of these simulations gave us consistent results.

Figure 9: Image Arrival RAM Disk Distribution: user request rate=100 request/day

2.3 Service time distribution of the second sub-queue

Now we can think about the service time distribution of the second queue. First of all, we try to approximate this to typical single service case, that is, in every service period only one request will be processed, while in a multiple service or batch service case, more than one requests will be processed in a service instance. Then we calculate the service time of each image as the difference between its arriving RAM disk and its finish point of downloading using the orbit simulation program FPSCA. After collecting such values for a lot of events, we can draw the service time distribution.

From figure 10, we can see that the service time distribution is a general distribution which doesn’t have a simple closed form expression. We can denote it as \( \frac{dB}{dt} \). Then we use M/G/1 model to describe this queue, which is Poisson arrival, general single service pattern and 1 server. Following the theory of M/G/1 model, although we no longer have a Markov process because of the relaxation of the exponential assumption on service times, it is imbedded within this non-Markov stochastic process a Markov chain (referred to as an imbedded Markov chain) at every service departure epoch.

This allows the utilization of Markov-chain theory in the analysis of the M/G/1 model. We define \( P_{ij} \) to be the transition probability, which represents the system one-step transition probability from \( i \) state to \( j \) state. We denote the transition probability matrix by [12]:

\[
P = [P_{ij}],
\]

where

\[
P_{ij} = Pr\{\text{system size immediately after a departure point is } j \text{ if system size after previous departure was } i\}
\]

\[
P_{ij} = \int_0^{\infty} \frac{e^{-\lambda t} \lambda^n t^i}{(j-i+1)!} dB(t), (j \geq i - 1, i \geq 1) \quad (1)
\]

Simplification results by defining

\[
k_n = Pr\{n \text{ arrivals during a service time } S=t\}
\]

\[
k_n = \int_0^{\infty} \frac{e^{-\lambda t} \lambda^n}{n!} dB(t) \quad (2)
\]

So that \( P_{ij} \) can be seen to equal to \( k_{j-i+1} \) and [12]

\[
P = [P_{ij}] = \begin{bmatrix}
    k_0 & k_1 & k_2 & \cdots \\
    k_0 & k_1 & k_2 & \cdots \\
    0 & k_0 & k_1 & \cdots \\
    0 & 0 & k_0 & \cdots \\
    \cdots & \cdots & \cdots & \cdots
\end{bmatrix} \quad (3)
\]

Assuming steady state is achievable, which means the user arrival rate is less than service rate, the steady-state probability vector, \( \pi = \{\pi_n\} \), can be found as the
solution to the stationary equation (see [12] Appendix 4).

\[ \pi_j = \sum_{0}^{\infty} \pi_i P_{ij} \Rightarrow \vec{\pi}P = \vec{\pi} \]

This yields

\[ \pi_i = \pi_0 k_i + \sum_{j=1}^{i+1} \pi_j k_{i-j+1} \quad (4) \]

In order to compute equation (2) and then solve the final steady state solution in closed form, we use Laguerre Polynomials[13] to fit the showed curve in figure 10. The advantage of Laguerre Polynomials is that it is a polynomial multiplied by an exponential term, which makes the calculation for equation (2) easier by using step-integration.

\[ y_m(t) = L_m(\beta t)e^{\frac{\beta}{m!}} \quad (5) \]

\[ y_n(t) = L_n(\beta t)e^{\frac{\beta}{n!}} \quad (6) \]

where \( L_m \) and \( L_n \) are Laguerre Polynomials.

By using the orthogonality of Laguerre Polynomials, we can have:

\[ \int_{0}^{\infty} y_m(t)y_n(t)dt = \frac{1}{m!n!} \int_{0}^{\infty} L_m(\beta t)L_n(\beta t)e^{-\beta t} d\beta t \quad (7) \]

let \( x = \beta y \)

\[ \int_{0}^{\infty} y_m(t)y_n(t) = \sigma_{mn} \quad (9) \]

If let

\[ \frac{dB}{dt} = \sum_{m=0}^{\infty} a_m y_m(t) \quad (10) \]

We then have

\[ a_m = \sqrt{\beta} \int_{0}^{\infty} \frac{dB}{dt} L_m(\beta t)dt \quad (11) \]

By integrating \( \frac{dB}{dt} \) and using the recurrence formula of Laguerre Polynomials, we calculated the coefficients \( a_m \) for \( \frac{dB}{dt} \). The fitting result presented by the diamonds is shown in figure 11. They match the solid line well and Figure 12 shows the coefficients converge well. Then we worked out the transition matrix and calculated the stationary distribution of the system length (details see [12]pp.223-239). The result is shown in figure 13. The x axis is the image queue length, y axis is the probability. “*” represents the result of queueing theory, while “+” shows the simulation result. Obviously, they don’t agree to each other. So we concluded that directly using queueing theory M/G/1 model doesn’t suit for our satellite imaging service second sub-queue.
3 Adaptation and Extension of Queueing Theory

3.1 Re-defined service time of the second sub-queue

Since our previous analysis using queueing theory doesn’t match our service model well, we have to extend the available models and make some adaptation on it. First of all, we re-defined the service time in the second sub-queue as the interval of successive image downloading finish time. On the other hand, we also simplified our satellite orbit model in **FPSCA**, in which case only considers coarse search. Without considering refinement in our satellite visible opportunity prediction, it will miss some satellite passes to a specific ground station. However, our coarse search predicts the main passes during each day which have good communication quality while all of the missing passes have short period and low elevation angle which actually make it quite difficult for normal sequential data downloading. So this simplification is not far away from the real world application. Figure 14, 15 and 16 show the new service time distribution for ground station with latitude and longitude as (0,180), (30,180) and (80,180) repectively. The x axis is the service time in day and the y axis is the counts of occurance which can be further calculated as probability.

From these figures, we noticed that the new service time distribution is although still general distribution, turns to be more like a deterministic distribution.

3.2 Result of adapted M/G/1 model

Generally, an imaging satellite in LEO will be visible from medium latitude ground stations for several times each day with a good communication link (figures 14 - 16). Each time the satellite is capable of downloading several images. In order to show the validity of the new service time definition in queueing theory we use M/G/1 to test it. We assume the satellite downloads only one image every pass and so re-simulate the situation accordingly. Figure 17, 18 and 19 show the results of adapted M/G/1 model composed of re-defined service time and simplified satellite orbital model. For these three figures, the x axis is the queuing length of the system, in other words is number of images in the system; the y axis is the system steady state. **87** represents the calculated probability from
queueing theory, while “+” displays the one from simulation. Three figures show the results for different ground station (0,180), (30,180) and (80,180) respectively. All of them show us a good agreement between simulation and queueing theory. Therefore the newly defined service time makes sense in our application using queueing theory.

![Figure 16: New Service Time Distribution for Simplified Model: ground station location (80,180)](image1)

![Figure 17: Steady State Solution for Simplified Model: ground station location (0,180)](image2)

![Figure 18: Steady State Solution for Simplified Model: ground station location (30,180)](image3)

![Figure 19: Steady State Solution for Simplified Model: ground station location (80,180)](image4)

3.3 $M/G^X/1$ bulk service and its result

3.3.1 Constant bulk size

In our previous analysis related to queueing theory, we treated our imaging queue as a single service. That is, every time the satellite passes over the ground station, only one image will be downloaded. But in reality, during each pass the satellite is capable of downloading a couple of images. So we can no longer treat the model as a single service $M/G/1$, and more appropriately, it becomes $M/G^X/1$ which can process bulk requests in each service. Unlike [22], in this section we treat the bulk size, $x$ as constant.

As our case is different from the general bulk service model, we derived a new transition probability matrix:

$$P_{ij} = \begin{cases} 
\frac{P_{ij}}{b-i} \sum_{n=0}^{k_n} k_n & 0 < i < b + 1, j = 0 \\
\frac{k_{j-i+b}}{b} & 1 \leq i \leq b, j \neq 0 \\
0 & \text{otherwise}
\end{cases}$$

(12)

and $P_{0j} = P_{ij}$.
The new matrix is as following:

\[
\begin{bmatrix}
\sum_{n=0}^{x-1} k_n & k_x & k_{x+1} & k_{x+2} & \cdots \\
\sum_{n=0}^{x-2} k_n & k_x & k_{x+1} & k_{x+2} & \cdots \\
\sum_{n=0}^{x-3} k_n & k_{x-1} & k_x & k_{x+1} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots \\
\end{bmatrix}
\] (13)

where \( x \) - bulk size is a constant.

Figure 20: Steady State Solution accuracy for Simplified Model when bulk service size is a constant: ground station location (30,180)

Using the above formula, we have calculated the error of the queueing theory steady-state probability compared with the simulation from:

\[
error = \pi_{sim} P_{ij} - \pi_{sim}
\] (14)

where \( P_{ij} \) is the queueing theory transition probability and \( \pi_{sim} \) is the steady-state probability from simulation. Figure 20 shows the error result. We can see that the maximum error is around 0.03.

Following figure 20 we can show for the steady-state probability vector \( \pi \) in queueing theory and show the comparison of this result with \( \pi_{sim} \) in figure 21. The RMS error is 0.0014. So both figures show a good agreement between queueing theory and simulation.

3.3.2 Bulk service size is Gaussianly distributed

In the real world, however, how many images can be downloaded during each satellite pass is not a constant. It depends upon the satellite orbit and the ground station location. To analyze it in a stochastic way, if the bulk size distribution is independent of the service time, we can have a universal expression for this kind of bulk model based on the transition matrix outlined in last subsection. The final transition matrix therefore is a weighted sum of transition matrices with different bulk size constants, and the weights are the probabilities of different bulk sizes. In order to test this, we varied the bulk size following a Gaussian distribution. The results are shown in figure 22, presenting the steady-state probability error of queueing theory from equation 14, and figure 23, which compares the difference between steady-state probability vectors \( \pi \) from queueing theory and simulation. This shows an RMS error of 8.7e-04. In this case, again, queueing theory gives us a good agreement with the simulation result.

3.3.3 Other system parameters can be achieved

Once we have found the system steady-state probability vector \( \pi \), then other parameters, such as the mean queue length and the mean waiting time, can be found directly from Little’s formula[12].
4 Discussions and conclusion

We have used queueing theory to describe the image capture and download queue on an Earth Observation satellite. We have assumed that the image arrival rate on the satellite is Poisson distributed and considered various distributions for the download rate. Comparison of the statistical results from queueing theory with numerical simulation show that these models are accurate representations of the process, and can predict the image storage requirement on the satellite.

The most significant advantages of using queueing theory in this application over operation simulation is that the former needs much less computational power. The service time modelling needs to be done only once for a specific space mission. Other calculations such as numerical integration and eigen vector calculations are also much simpler compared to astro-dynamics computation even for fast algorithms such as FPSCA. However, queueing theory also has some limitations.

In our previous analysis, we have assumed the distribution of ground targets of users follows a uniform distribution which leads to a Poisson arrival in the second sub-queue. However, in real applications, the target distribution could be different, such as imaging follows a sea ship route or only for a specific region. So we changed the ground target distribution to be Gaussian, as shown in figure 24.

We wish to know the distribution of the interval between successive images arriving in the RAM disk: whether it is still in exponential form implying Poisson arrival or not. Figure 25 shows the result, it tells us that if the ground target distribution is not uniform, image arrival at the RAM disk, is the user arrival in the second queue, is no longer Poisson distributed. In this case, for the second queue, it is no longer M/G/1 or M/Gx/1 queueing theory. We would use the GI/G/1 model for this case. However, such a complicated mathematical derivation is at odds with the motivation in this application, that is, a more computationally inexpensive way for imaging service analysis.

On the other hand, we have only considered one single satellite (1 server) throughout our analysis. In Earth Imaging and Observation applications, however, constellations of satellites turn out to be a better strategy. So if we want to apply queueing theory to analyze this kind of application, we need to think about multiple servers instead of single server queues which requires more complicated calculations.

References

Figure 24: We set the ground target is Gaussian distribution, centre is (0,180).


Figure 25: This figure shows the image arrival RAM disk distribution when the ground target is Gaussian distribution, centre is (0,180).


