

NASA/TM—2018-219927



Space-Based Communications Buffer Sizing

Louis M. Handler
Glenn Research Center, Cleveland, Ohio

August 2018

NASA STI Program . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) Program plays a key part in helping NASA maintain this important role.

The NASA STI Program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI Program provides access to the NASA Technical Report Server—Registered (NTRS Reg) and NASA Technical Report Server—Public (NTRS) thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- **TECHNICAL PUBLICATION.** Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counter-part of peer-reviewed formal professional papers, but has less stringent limitations on manuscript length and extent of graphic presentations.
- **TECHNICAL MEMORANDUM.** Scientific and technical findings that are preliminary or of specialized interest, e.g., “quick-release” reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.
- **CONTRACTOR REPORT.** Scientific and technical findings by NASA-sponsored contractors and grantees.
- **CONFERENCE PUBLICATION.** Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.
- **SPECIAL PUBLICATION.** Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.
- **TECHNICAL TRANSLATION.** English-language translations of foreign scientific and technical material pertinent to NASA's mission.

For more information about the NASA STI program, see the following:

- Access the NASA STI program home page at <http://www.sti.nasa.gov>
- E-mail your question to help@sti.nasa.gov
- Fax your question to the NASA STI Information Desk at 757-864-6500
- Telephone the NASA STI Information Desk at 757-864-9658
- Write to:
NASA STI Program
Mail Stop 148
NASA Langley Research Center
Hampton, VA 23681-2199



Space-Based Communications Buffer Sizing

Louis M. Handler
Glenn Research Center, Cleveland, Ohio

National Aeronautics and
Space Administration

Glenn Research Center
Cleveland, Ohio 44135

Level of Review: This material has been technically reviewed by technical management.

Available from

NASA STI Program
Mail Stop 148
NASA Langley Research Center
Hampton, VA 23681-2199

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
703-605-6000

This report is available in electronic form at <http://www.sti.nasa.gov/> and <http://ntrs.nasa.gov/>

Space-Based Communications Buffer Sizing

Louis M. Handler
National Aeronautics and Space Administration
Glenn Research Center
Cleveland, Ohio 44135

1.0 Introduction

This paper considers the transmission interruption timing for space-based communications due to weather. The length of the weather interruption is one of the main influences determining the size of the storage needed to save the data to be sent at a later time. This analysis uses the Markov chain probabilities for weather to determine how many cloudy or rainy time steps in a row will not allow data transmission. The mean and standard deviation are computed, which along with the average data rate to be sent can be used to help size storage buffers in which to save data between periods of good weather.

Nomenclature

RF radiofrequency

Symbols

a	availability ratio
A	event A
B	event B
n	number of time steps
N	number of time steps before success or failure can be determined
p	speed-up factor
P	probability function of a specified event
$P(A)$	probability of occurrence of event A
$P(A \cap B)$	probability of occurrence of both event A and event B at same time; that is, intersection of the two events
$P(A \cup B)$	probability of occurrence of either event A or event B at same time; that is, union of the two events
$P(B)$	probability of event B
$\text{Pr}[x,y]$	probability of weather y (0 or 1) at the current time step given weather x (0 or 1) at the previous time step
R_d	average data rate per time step
X	total minutes in an average duty cycle
Y	available minutes in an average duty cycle
σ	standard deviation of number of time steps in a row of bad weather
σ^2	variance of number of time steps in a row of bad weather
Σ	summation from lower to upper limit
ΔT	size of each time step
μ	average number of time steps in a row of bad weather
μ_{good}	average number of time steps in a row of good weather

2.0 Buffer Sizing

2.1 Markov Probabilities

This paper uses the Markov chain probabilities for weather, where the weather at the next time step depends only on the weather at the current time step. This paper will determine the probabilities for how many time steps in a row will not allow data transmission. Using the values for how long data cannot be transmitted due to weather, buffer size will be computed to hold the data not transmitted.

The transmission frequency determines whether cloudy weather or rainy weather can block transmission. For a radiofrequency (RF) transmitter and/or receiver combination, rain would block the signals but mere cloudiness would not. For an optical transmitter and/or receiver combination, cloudiness would block the signal. Optically, rain is not treated separately since it rarely appears without clouds. Although it is possible to have various levels of cloudy or rainy in the Markov chain model, only two levels were used to simplify the model. The levels are clear to allow signals to pass and either cloudy or rainy that blocks the signals.

The Markov chain uses state transition probabilities. In this case, the probabilities are determined for the change from clear at one time step to cloudy or rainy at the following time step and vice versa. These states are denoted by 0 and 1, with 0 for rainy or cloudy, and 1 for clear.

The definitions of the transition probabilities are explained as follows:

1. $\text{Pr}[0,0]$ is the probability of cloudy or rainy given cloudy or rainy at the previous time step.
2. $\text{Pr}[0,1]$ is the probability of clear given cloudy or rainy at the previous time step.
3. $\text{Pr}[1,0]$ is the probability of cloudy or rainy given clear at the previous time step.
4. $\text{Pr}[1,1]$ is the probability of clear given clear at the previous time step.

It is assumed that there are no intermediate states, so it must either be clear or be cloudy or rainy at each time step, regardless of the value at the previous time step. Thus, the probabilities add up to 1:

$$\text{Pr}[0,0] + \text{Pr}[0,1] = 1 \quad (1)$$

$$\text{Pr}[1,0] + \text{Pr}[1,1] = 1 \quad (2)$$

The mean and standard deviation of the number of bad weather time steps in a row can be used to help size storage buffers to save data between periods of clear weather.

The buffer sizing may be estimated by Monte Carlo simulation or computation. For simulation, the initial weather state is computed at each Earth terminal randomly. It should not matter which type of weather is at the start because if there was a higher likelihood of one type of weather over the other, the weather would soon change according to the Markov probabilities specified. To eliminate any bias introduced by this assumption at this initial point, the initial state is ignored as far as any data transmission is concerned.

For the next time step, the probability of a cloudy or rainy time period is computed, given the previous time period's weather state at each Earth terminal. This probability is used to determine the weather state using a random variable. That step is repeated by storing the current weather state so that it can be tested by the next step as the previous weather state.

Statistics are computed for the number of each type of weather period, the amount of data lost or received during each period, the sequential number cloudy or rainy time steps, and percentages.

2.2 Probabilities of Length of Bad Weather

The rest of this paper uses the Markov probabilities for weather to determine how many cloudy or rainy time steps are likely in a row. The probabilities allow the weather to change gradually from the initial conditions, using intermediate values. The transition probabilities are assumed to remain constant long enough for the computations to be reasonably valid. The fact that some locations have a rainy or dry season during which the probabilities change is not taken into account here.

The mean and standard deviation of the number of time steps in a row of bad weather can be used to help size storage buffers to save data between periods of good weather. In Equation (3), probabilities are computed for each number of bad weather periods in a sequence.

The probability of n steps in a row of type 0 is as follows. For the probability of 1 exactly; that is, probability of 0 then 1 is

$$P(1) = \Pr[0,1] \quad (3a)$$

For the probability of 2 exactly; that is, probability of 0 then 0 then 1 is

$$P(2) = \Pr[0,0]\Pr[0,1] \quad (3b)$$

For the probability of 3 exactly; that is, probability of 0 then 0 then 0 then 1 is

$$P(3) = \Pr[0,0]\Pr[0,0]\Pr[0,1] \quad (3c)$$

For the probability of 4 exactly; that is, probability of 0 then 3 more 0's then 1 is

$$P(4) = \Pr[0,0]^3 \Pr[0,1] \quad (3d)$$

Or more generally, for the probability of n exactly; that is, probability of 0 then $(n-1)$ more 0's then 1 is

$$P(n) = \Pr[0,0]^{n-1} \Pr[0,1] \quad (3e)$$

The probability for each number of bad weather time steps in a row is used to compute the mean and standard deviation of the number of bad weather time steps in a row, which are used to compute the amount of data lost during each period of bad weather.

2.3 Average

The expected value of the number of bad weather time steps in a row (0's in a row), denoted by μ , is obtained by summing the number of 0's in a row multiplied by its probability, $P(n)$:

$$\mu = \sum_{n=1}^{\infty} n \Pr[0,0]^{n-1} \Pr[0,1] \quad (4)$$

The summation will converge when the probabilities are less than 1. The successive terms start decreasing when

$$n \Pr[0,0]^{n-1} \Pr[0,1] > (n+1) \Pr[0,0]^n \Pr[0,1] \quad (5)$$

Since the probabilities are positive, divide both sides by $\{\text{Pr}[0,0]^{n-1} \text{Pr}[0,1]\}$ to obtain

$$n > (n+1) \text{Pr}[0,0] \quad (6)$$

Subtracting $\{n \text{Pr}[0,0]\}$ from both sides and solving for n results in

$$n > \frac{\text{Pr}[0,0]}{(1 - \text{Pr}[0,0])} \quad (7)$$

Since this value of n may be large when the terms start decreasing and convergence may actually happen much later, it is better to solve Equation (4) in closed form. This may be done similar to the proof for the sum of a geometric series.

Equation (4) can be written as

$$\mu = 1 \text{Pr}[0,1] + 2 \text{Pr}[0,0] \text{Pr}[0,1] + 3 \text{Pr}[0,0]^2 \text{Pr}[0,1] + \dots + (n+1) \text{Pr}[0,0]^n \text{Pr}[0,1] + \dots \quad (8)$$

Multiply Equation (8) by $\text{Pr}[0,0]$ and line up the exponents:

$$\mu \text{Pr}[0,0] = \text{Pr}[0,0] \text{Pr}[0,1] + 2 \text{Pr}[0,0]^2 \text{Pr}[0,1] + \dots + n \text{Pr}[0,0]^n \text{Pr}[0,1] + \dots \quad (9)$$

Then subtract the two equations:

$$\mu - \mu \text{Pr}[0,0] = \text{Pr}[0,1] + \text{Pr}[0,0] \text{Pr}[0,1] + \text{Pr}[0,0]^2 \text{Pr}[0,1] + \dots + \text{Pr}[0,0]^n \text{Pr}[0,1] + \dots \quad (10)$$

Changing notation, Equation (10) may be expressed as

$$\mu - \mu \text{Pr}[0,0] = \sum_{n=1}^{\infty} \text{Pr}[0,0]^{n-1} \text{Pr}[0,1] \quad (11)$$

But the right-hand side can be shown to be equal to 1, either by summing the pieces of Equation (3) because the probabilities have to add up to 1 or by going back to the sum of a geometric series and using $\text{Pr}[0,0] = 1 - \text{Pr}[0,1]$; that is

$$\sum_{n=1}^{\infty} \text{Pr}[0,0]^{n-1} \text{Pr}[0,1] = \frac{\text{Pr}[0,1]}{1 - \text{Pr}[0,0]} = \frac{\text{Pr}[0,1]}{\text{Pr}[0,1]} = 1 \quad (12)$$

Solving Equation (11) for μ after substituting Equation (12):

$$\mu = \frac{1}{1 - \text{Pr}[0,0]} = \frac{1}{\text{Pr}[0,1]} \quad (13)$$

2.4 Standard Deviation

The standard deviation is done similarly. The standard deviation squared of the number of time steps in a row of bad weather (0's in a row), denoted by σ^2 , is obtained by summing:

$$\sigma^2 = \sum_{n=1}^{\infty} (n - \mu)^2 \text{Pr}[0,0]^{n-1} \text{Pr}[0,1] \quad (14)$$

Expanding:

$$(n - \mu)^2 = n^2 - 2n\mu + \mu^2 \quad (15)$$

Substituting Equation (15) into Equation (14), applying the distributive law, and using Equations (4) and (12) results in

$$\sigma^2 = \left\{ \sum_{n=1}^{\infty} n^2 \Pr[0,0]^{n-1} \Pr[0,1] \right\} - 2\mu^2 + \mu^2 \quad (16)$$

Simplifying the terms on the right:

$$\sigma^2 = \left\{ \sum_{n=1}^{\infty} n^2 \Pr[0,0]^{n-1} \Pr[0,1] \right\} - \mu^2 \quad (17)$$

Multiplying both sides by $\Pr[0,0]$:

$$\Pr[0,0]\sigma^2 = \left\{ \sum_{n=1}^{\infty} n^2 \Pr[0,0]^n \Pr[0,1] \right\} - \mu^2 \Pr[0,0] \quad (18)$$

To match up the exponents, set $n = m-1$ in the previous equation:

$$\Pr[0,0]\sigma^2 = \left\{ \sum_{m=2}^{\infty} (m-1)^2 \Pr[0,0]^{m-1} \Pr[0,1] \right\} - \mu^2 \Pr[0,0] \quad (19)$$

The lower limit of summation can be extended to $m = 1$ to match the summation in Equation (17). Rename back to n . Then subtract Equation (19) from Equation (17):

$$\sigma^2 - \sigma^2 \Pr[0,0] = \left\{ \sum_{n=1}^{\infty} (n^2 - (n-1)^2) \Pr[0,0]^{n-1} \Pr[0,1] \right\} - \mu^2 (1 - \Pr[0,0]) \quad (20)$$

Recalling

$$n^2 - (n-1)^2 = n^2 - (n^2 - 2n + 1) = n^2 - n^2 + 2n - 1 = 2n - 1 \quad (21)$$

Substituting Equation (21) into Equation (20):

$$\sigma^2 - \sigma^2 \Pr[0,0] = \left\{ \sum_{n=1}^{\infty} (2n - 1) \Pr[0,0]^{n-1} \Pr[0,1] \right\} - \mu^2 (1 - \Pr[0,0]) \quad (22)$$

Using the distributive law and substituting Equations (4) and (12) into Equation (22):

$$\sigma^2 (1 - \Pr[0,0]) = \{2\mu - 1\} - \mu^2 (1 - \Pr[0,0]) \quad (23)$$

From Equation (13), substitute for $(1 - \Pr[0,0]) = \frac{1}{\mu}$:

$$\frac{\sigma^2}{\mu} = 2\mu - 1 - \frac{\mu^2}{\mu} \quad (24)$$

Simplify as

$$\frac{\sigma^2}{\mu} = \mu - 1 \quad (25)$$

or finally as

$$\sigma^2 = \mu(\mu - 1) \quad (26)$$

2.5 Implications

2.5.1 Buffer Size Computation

To compute the size of a buffer, multiply the average outage in the number of bad weather steps, μ , times the size of each time step, ΔT , times the average data rate per time step, R_d :

$$\text{Size}_\mu = \mu \Delta T R_d \quad (27)$$

For safety, the 3σ value would be

$$\text{Size}_{3\sigma} = (\mu + 3\sigma) \Delta T R_d \quad (28)$$

For a normal distribution of values, the 3σ size would include 99.7 percent of the data. This size may be improved by compression, depending on the type of data possibly with some tradeoffs in timing.

2.5.2 Buffer Size Adequacy

This buffer size may not be adequate if the good weather period is not long enough to send the data accumulated during the bad weather period. Equation (13) specifies μ for bad weather, but there is a corresponding μ_{good} for good weather. Similar to Equation (4), the expected value of the number of time steps in a row of good weather (1's in a row), denoted by μ_{good} , is obtained by summing the number of 1's in a row by its probability:

$$\mu_{\text{good}} = \sum_{n=1}^{\infty} n \Pr[1,1]^{n-1} \Pr[1,0] \quad (29)$$

As in the analysis above, it can be found to be

$$\mu_{\text{good}} = \frac{1}{1 - \Pr[1,1]} = \frac{1}{\Pr[1,0]} \quad (30)$$

The data keep accumulating during periods of bad weather, even if periods of good weather intervene, unless the bad weather outlasts the good by enough that the transmission of both the new and stored data cannot be accomplished. The transmission of the stored data does not have to be accomplished at the same rate as the data is transmitted in real time due to preformatting, precompression, data access, and so on. The usual data rate may not be the maximum data rate possible when sending the saved data. Let p be the speed-up factor, if any, applied to speed up transmission of saved data during good weather. A value of 1 indicates no speed up. A value of less than 1 is actually a slowdown. There is a likely to be a problem when the amount of data that can be sent in the good weather at the speed-up rate is insufficient to transmit the stored data as shown by the inequality in data size units:

$$\mu_{\text{good}} \Delta T (p R_d) < \mu \Delta T R_d \quad (31)$$

Dividing both sides of the previous equation by $(\Delta T R_d)$, there is likely to be a problem when

$$p \mu_{\text{good}} < \mu \quad (32)$$

Substituting Equations (13) and (30) into the previous equation results in

$$\frac{p}{\Pr[1,0]} < \frac{1}{\Pr[0,1]} \quad (33)$$

Multiplying both sides by $\Pr[0,1] \Pr[1,0]$ is

$$p \Pr[0,1] < \Pr[1,0] \quad (34)$$

Therefore, there is a likely problem when p times the probability of good weather given bad weather at the previous time step is less than the probability of bad weather given good weather at the previous time step. The higher the value of p , the greater the speed up, the less likely for there to be a problem with not having enough good weather.

Furthermore, the good weather may have to outlast the bad on average by enough to send both new and old data. The above assumes that sending can be done in parallel with no additional time taken.

2.5.3 Cyclic Schedule Delays

There may also be scheduling delays due to other priorities, a daily duty cycle, power constraints, heat dissipation, orbital dynamics, and so on. These are not considered here in detail but could add to the time before the data can be sent, increasing the buffer size beyond the 3σ value due to weather. For example, if the orbit and daily duty cycle are such that the signal can only be scheduled every Y available minutes out of every X minutes, it can only send for Y/X of the time and saves data $(X - Y)/X$ of the time, regardless of the weather.

Instead of just the weather, the probabilities would have to reflect both the good weather and the availability of both the transmitter and receiver. Since the schedule is usually set up in advance, the weather may be considered independent of the schedule. Using the theorem for the joint probability of independent events,

$$P(A \cap B) = P(A)P(B) \quad (35)$$

where $P(A)$ is the probability of occurrence of event A, $P(B)$ is the probability of event B, and $P(A \cap B)$ is the probability of occurrence of both event A and event B at the same time; that is, the intersection of the two events, then the probability that the signal can get through is equal to the product of the probability that the transmitter and/or receiver is available times the probability that the weather is good. Let $a = Y/X$, the availability ratio. The probability that the transmitter and/or receiver is available is a . The probabilities that the weather is good when the transmitter and/or receiver is available are

$$\Pr'[0,1] = a \Pr[0,1] \quad (36)$$

and

$$\Pr'[1,1] = a \Pr[1,1] \quad (37)$$

The other equations are computed using those values so that as before

$$\mu' = \frac{1}{\Pr'[0,1]} \quad (38)$$

and

$$\sigma^2 = \mu'(\mu' - 1) \quad (39)$$

Since μ' is inversely proportional to the probability, the buffer size will need to be larger when there is less availability for the same weather statistics.

2.5.4 Multiple Destinations

There may be buffer sizing improvements as well as timing and schedule improvements, due to alternate paths for the signal to travel to multiple destinations. This technique might be used to find a destination whose weather allows for successful communication or even possibly an alternate site in space so weather is not an immediate problem. The signal might be sent out on all paths at once or to a path that is recognized as available and clear. The probability that multiple destinations have bad weather is smaller than for any single destination. The buffer size is decreased when any one path can be successful at any time and contribute to the whole, consequently stopping the process for the unsuccessful destination. The probabilities and buffer size depend on the exact relationships that lead to successful transmission.

If success depends on the entire signal being received, then the buffer size will usually be determined by the location with the better weather for the longest period of time and bad weather for the shortest period of time. The buffer size can then be computed for the location with the smallest μ so that

$$\mu' = \min(\mu_1, \mu_2) \quad (40)$$

Alternately, if success is determined at each time step, the joint probability for multiple destinations is used. A joint probability for two locations may be defined as $\Pr'[0,1]$ = probability of clear at either place given cloudy or rainy at the previous time step at the same place. Separating the two locations and applying a logical distributive law, this may be rewritten in a more computable form as $\Pr'[0,1]$ = probability of clear at location 1 given cloudy or rainy at the previous time step at location 1 or clear at location 2 given cloudy or rainy at the previous time step at location 2 (refer to definition 2, for $\Pr[0,1]$).

Using the theorem for the joint probability of independent events,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (41)$$

where $P(A \cup B)$ is the probability of the occurrence of either event A or event B at the same time; that is, the union of the two events, and assuming that the weather at the two locations are independent, the resulting equation is

$$\Pr'[0,1] = \Pr_1[0,1] + \Pr_2[0,1] - \Pr_1[0,1]\Pr_2[0,1] \quad (42)$$

Similarly,

$$\Pr'[1,1] = \Pr_1[1,1] + \Pr_2[1,1] - \Pr_1[1,1]\Pr_2[1,1] \quad (43)$$

For example, if there are two destinations, both with $\Pr[0,1] = 0.5$, then substituting in Equation (42), $\Pr'[0,1] = 0.5 + 0.5 - 0.25 = 0.75$; that is, the probability of finding a clear location increases from 50 to 75 percent with a proportional decrease in buffer size required. This computation assumed that what happened at each time step was independent of the adjacent time steps; however, that does not have to be the case.

However, if success requires N number of time steps before success or failure can be determined and a change can occur, the probability of N clear time steps in a row is

$$\Pr''[1,0] = \Pr[1,1]^{N-1} \Pr[1,0] \quad (44)$$

and for two destinations, using Equation (41),

$$\Pr''[1,0] = \Pr_1[1,1]^{N-1} \Pr_1[1,0] + \Pr_2[1,1]^{N-1} \Pr_2[1,0] - \Pr_1[1,1]^{N-1} \Pr_1[1,0] \Pr_2[1,1]^{N-1} \Pr_2[1,0] \quad (45)$$

Then if both locations have the same probabilities; that is, $\Pr_1[1,0] = \Pr_2[1,0] = \Pr[1,0]$ and $\Pr_1[1,1] = \Pr_2[1,1] = \Pr[1,1]$, then

$$\Pr''[1,0] = 2 \Pr[1,1]^{N-1} \Pr[1,0] - \Pr[1,1]^{2N-2} \Pr[1,0]^2 \quad (46)$$

and the ratio of Equation (46) divided by Equation (44) is

$$\text{ratio} = 2 - \Pr[1,1]^{N-1} \Pr[1,0] \quad (47)$$

Then if all the probabilities were 0.5, $\Pr''[1,0] = 2(0.5)^N - (0.5)^{2N}$, which is still much improved over $(0.5)^N$ for the same time period with a single destination.

Table I shows how quickly the ratio of probabilities approaches 2.0 as the number of time steps included in the computation increases.

TABLE I.—MULTIPLE DESTINATION PROBABILITIES

N	One destination (Eq. (44)), percent	Two destinations (Eq. (46)), percent	Ratio (Eq. (47))
1	50.00	75.00	1.500
2	25.00	43.75	1.750
3	12.50	23.44	1.875
4	6.25	12.11	1.938
5	3.13	6.15	1.969
6	1.56	3.10	1.984
7	0.78	1.56	1.992
8	.39	0.78	1.996
9	.20	.39	1.998
10	.10	.20	1.999

Bibliography

Ehnberg, Jimmy S.G.; and Bollen, Math H.J.: Simulation of Global Solar Radiation Based on Cloud Observations. *Solar Energy*, vol. 78, 2005, pp. 157–162.

Shambayati, Shervin: On Selecting the Size of a Spacecraft's Buffer. *The Interplanetary Network Progress Report*, vol. 42–174, 2008.

