Priority Repair Schemes in the Deep Space Network

I. Eisenberger
Communications Systems Research Section
G. Lorden
California Institute of Technology

A method is given to increase the cost-effectiveness of spares pools by performing repairs of each type of module on a priority basis whenever the number of available spares falls below a critical level. For a system already operating with established spares pools, the problem of choosing the critical levels is solved by an algorithm which attains the largest possible uptime ratio (UTR) for any specified total amount of priority repairing. Provisioning of new spares in conjunction with priority repairing is also optimized so as to achieve any UTR goal with minimum cost. Examples are given to show that even a small amount of priority repairing can yield a substantial reduction of sparing costs.

I. Introduction

The idea of a priority scheme for repair facilities in the Deep Space Network (DSN) is based on simple considerations. Experience has shown that the average time for repair of failed modules is about two weeks, assuming that the repair is performed locally at the complex. Most of that time is spent waiting for the actual repair. So long as an adequate number of spares is available, the two-week turnaround time is not critical. But when spares are temporarily unavailable because they have been used for recent replacements of failed modules, the two-week wait incurs a substantial risk of downtime. If these critical situations are treated on a priority basis and the turnaround time is shortened, say, to one week, then there is a dramatic improvement in the system uptime ratio (UTR), the fraction of time the entire system is operable. For example, if only 5 to 10% of all repairs are performed on a priority basis, the improvement in UTR is nearly as large as if all repairs had a turnaround time of one week. This is illustrated by the examples in Section III.

In Section IV, the idea of priority repairing is considered in the context of cost-effective spares provisioning (Ref. 1), the problem of providing a spares package for a system. A method of combining the choice of a priority scheme and a spares package is given which optimizes the trade-offs between total cost, uptime ratio, and total priority workload, i.e., the frequency of priority repairs. An example in Section V illustrates the fact that substantial savings in the cost of spares needed to meet uptime ratio goals are realizable through the introduction of priority schemes, even when the priority workload is held at less than 10%.

II. Description of Optimal Priority Schemes

The crucial thing to determine is when to give priority to repairs of a particular type of module. Our method chooses a critical level r for the number of available spares, depending upon the type of module and the spares complement. Whenever the number of spares available falls below r, priority is given to that type. Priority status ends when there are once again at least r spares. At any time, several module types may have priority status. It is assumed that these are repaired in the order that they went into priority.

The different choices of r for different types of modules are based on a trade-off between the priority workload, i.e., frequency of priority repairs, and the system uptime ratio. Let $i = 1,\cdots,k$ denote the types of modules in the system n_i the number of type i operating in the system, N_i the number of spares, and r_i the critical level to be chosen. Assume given the mean time between failures (MTBFs) for the k types, the mean time to repair without priority μ (same for all types), and the mean time to repair under priority μR (an $R < 1$ is specified). The Markov method of Ref. 2 can be used to calculate for each $i = 1,\cdots,k$ the uptime ratio UTR_i of the ith type and the frequency of priority repairs F_i (e.g., 4.2 priority repairs per year). The system uptime ratio is defined by

$$UTR = (UTR_1) \times (UTR_2) \times \cdots \times (UTR_k) \qquad (1)$$

and the system frequency of priority repairs is

$$F = F_1 + \cdots + F_k \qquad (2)$$

The trade-off is made by choosing a value of $d > 0$ and maximizing

$$\log(UTR) - d \cdot F = \sum_{i=1}^{k} (\log UTR_i - d \cdot F_i) \qquad (3)$$

over all priority schemes, i.e., sets of critical levels $\{r_1,\cdots,r_k\}$. This is done by maximizing the right-hand side of Eq. (3) term-wise, i.e., for each i, the value of r_i is chosen among the possible values $0,1,\cdots,N_i+1$ to maximize $\log UTR_i - d \cdot F_i$. It should be noted that as d gets larger, the r_i's get smaller and so do the F_i's and UTR_i's.

Every priority scheme $\{r_1,\cdots,r_k\}$ obtained in this way for some d is optimal in the sense that neither its UTR nor its F can be improved upon without sacrificing the other. To see this, suppose that for a particular $d = d^*$, the values UTR^* and F^* are obtained by the algorithm described above. Let UTR^{**} and F^{**} denote the values for some other priority scheme. Since Eq. (3) was maximized for $d = d^*$ by using the algorithm, it is clear that

$$UTR^* - d^* \cdot F^* \geqq UTR^{**} - d^* \cdot F^{**} \qquad (4)$$

By Inequality (4), if $F^{**} < F^*$, then $UTR^{**} < UTR^*$, i.e., the UTR is sacrificed. Similarly, if $UTR^{**} > UTR^*$, then $F^{**} > F^*$, i.e., F is sacrificed.

It can also be shown that every optimal priority scheme is obtainable by using the algorithm for some $d > 0$; in other words, if a priority scheme is not obtainable by the algorithm for some d, then either its UTR and F can both be improved upon simultaneously or one of them can be improved without sacrificing the other.

III. Priority Schemes for Existing Spares Packages

A variety of examples were computed to determine the kind of uptime ratio improvement one can expect from the use of optimal priority schemes as described in Section II. For these examples, it was assumed that the numbers of spares $N_1,N_2,\cdots$, of each module type are given, along with the numbers operating, $n_1,n_2,\cdots$, and the failure rates and repair rate. In each example, several choices of the parameter d of Section II were made, and in each case the uptime ratio and a total rate of priority repairs were computed. The latter rate was used to determine the fraction of all repairs performed under priority. This fraction seems more meaningful intuitively than the rate of priority repairs itself.

These results yielded plots of uptime ratio against the priority repair fraction, such as the one in Fig. 1, which is typical. The value $R = 0.5$ was used for Fig. 1, whereas $R = 0.25$ (quadruple rate for priority repairs) was used for the same package in Fig. 2. This package consisted of a total of 55 spares for 30 module types.

Notice that in both Figs. 1 and 2 the improvement in uptime ratio is dramatic, even with a small fraction of priority repairs. The improvement levels off rapidly once the priority fraction reaches 10–15%. In other words, 10–15% priority fraction yields very nearly as high an UTR as 100%, the latter amounting to performing all repairs at the high priority repair rate.

The examples computed revealed that the steep improvement in UTR for small priority fractions is even more pronounced when the existing system uptime ratio is high and also when R is low. For example, a larger spares package for the same system as in Figs. 1 and 2 showed an initial UTR of 0.955, which improved to 0.998 with a 4% priority fraction and $R = 0.25$.

IV. Method of Spares Provisioning Using Priority

The problem is to choose the number of spares N_i and the priority level r_i for each module type i so that the overall package achieves the greatest UTR for the lowest total cost and total priority fraction. To formulate this mathematically, use relations (1) and (2) defining the system uptime ratio and system frequency of priority repair, together with the relation

$$C = C_1 + \cdots + C_k \tag{5}$$

where C denotes the total cost of spares, and the C_i's are the cost of spares for the different types of modules. The determination of optimal trade-offs is made by extending the method of Section II. Replacing the fundamental relation, (3) of that section is the relation

$$\log(UTR) - d \cdot F - e \cdot C = \sum_{i=1}^{k} (\log UTR_i - d \cdot F_i - e \cdot C_i) \tag{6}$$

which expresses the quantity to be maximized for given d and $e > 0$. The reason for seeking to maximize Eq. (6) is the same as in Section II: A total package of N_i's and r_i's maximizing Eq. (6) for some d and $e > 0$ is optimal in the sense that neither its UTR, nor its F, nor its C can be improved without sacrificing one or both of the others. As in Section II, the fact that the quantity to be maximized, in this case Eq. (6), is expressed as a sum over i means that it can be maximized term-wise. So, for each i, it is necessary to choose an N_i and an r_i to maximize

$$\log UTR_i - d \cdot F_i - e \cdot C_i = \log UTR_i(N_i,r_i) - d \cdot F_i(N_i,r_i) - e \cdot C_i(N_i) \tag{7}$$

where the right-hand side shows the dependence on N_i and r_i. Since the cost of spares C_i depends only on N_i and not on r_i, the problem is immediately solved by fixing N_i and choosing the r_i which maximizes

$$\log UTR_i(N_i,r_i) - d \cdot F_i(N_i,r_i)$$

among the possible values of $r_i = 0,1,\cdots,N_i+1$.

Once the optimal choice of r_i for each N_i has been determined in this way, the dependence on r_i in Eq. (7) is removed, and one has to choose N_i to maximize

$$\log UTR_i(N_i) - d \cdot F_i(N_i) - C_i(N_i) = V(N_i) - PN_i$$

using a briefer notation, where P denotes e times the unit cost of spares of type i. The expression $V(N_i) - PN_i$ is exactly of the form considered in Ref. 1 (the value-cost lemma) in connection with the optimal choice of spares packages without consideration of priority. The algorithm given in that paper for recursively generating *all* optimal packages applies in this case because the required monotonicity condition has been found to be satisfied over the range of interest. In the present context, this means that *for a given* $d > 0$ the algorithm generates all the optimal packages obtainable over an arbitrary interval of e values.

In practice, one wants a listing of the specifications (uptime ratio, cost, and priority fraction) for a variety of optimal packages, keeping the specifications within prescribed ranges. This is most effectively accomplished by fixing a d value, listing the specifications of all optimal packages whose cost (or uptime ratio) is less than a prescribed limit, and then repeating the process for other d values (higher d values give lower priority fractions). Because of the discrete nature of the packages, small changes in d will often produce the same (or many of the same) sets of specifications. It is not difficult, however, by judicious choices of d, to keep this repetition to a minimum and still obtain virtually all the optimal packages over the range of interest.

V. A Spares Provisioning Example

The spares provisioning algorithm described in Section IV was applied to several examples of system configurations in an attempt to determine the range of typical cost reductions achievable using priority schemes. These cost reductions are measured from the cost of optimal spares provisioning for the same system uptime ratio without priority (based on Ref. 1).

The example illustrated below was chosen with the stipulation that the fraction of priority repairs be 9–10%, which was felt to be within the range of reasonable levels. Figure 3 shows the cost in thousands of dollars of the spares packages as a function of the system uptime ratio.

The bottom curve does this for the no-priority case with mean repair time of one week, half the normal two-week time. The middle curve shows the case of 9–10% priority repairing, with one-week repair time under priority, two weeks otherwise. The top curve depicts the case where no priority is used and the repair time is 1.82 weeks—a little less than two weeks, this value being chosen to yield the same *average* rate of repair as the 9–10% priority scheme, namely, 0.55 per week. This value is used for the top curve to provide a truer comparison with the middle curve representing the priority schemes: the difference between the two is not due to any improvement in average repair time, but is only due to the effectiveness of the priority schemes in allocating repair work according to immediate needs.

Note that the cost reduction of the middle (priority) curve over the top curve is about 20% in the range of uptime ratios from 0.9 to 0.995. Even if the high-priority repair time were always in effect, so that the bottom curve would apply, the additional cost reduction would be only about 8% in the uptime ratio range below 0.95, dropping steadily to about 3% at the 0.99 level. Thus, the priority approach achieves nearly all of the cost reduction possible through speedier repairs, even though the speedier repairs are required less than 10% of the time.

References

1. Eisenberger, I., Lorden, G., and Maiocco, F., "Cost Effective Spares Provisioning for the Deep Space Network," in *The Deep Space Network Progress Report 42-20,* pp. 128–134, Jet Propulsion Laboratory, Pasadena, Calif., Apr. 15, 1974.
2. Eisenberger, I., Lorden, G., and Maiocco, F., "A Preliminary Study of Spares Provisioning for the Deep Space Network," in *The Deep Space Network Progress Report,* Technical Report 32-1526, Vol. XVIII, pp. 102–110, Jet Propulsion Laboratory, Pasadena, Calif., Dec. 15, 1973.

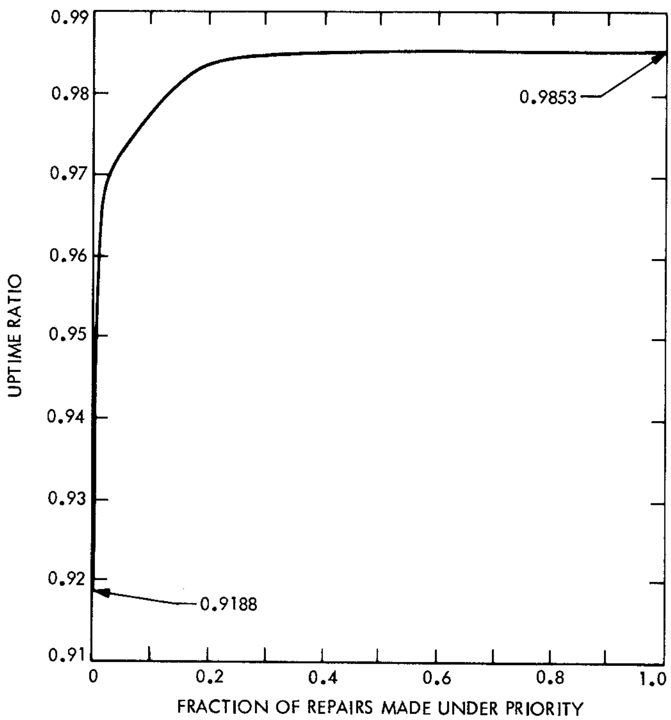

Fig. 1. Improvement in system UTR as the fraction of repairs made under priority increases, μ = 336 h, R = 0.5

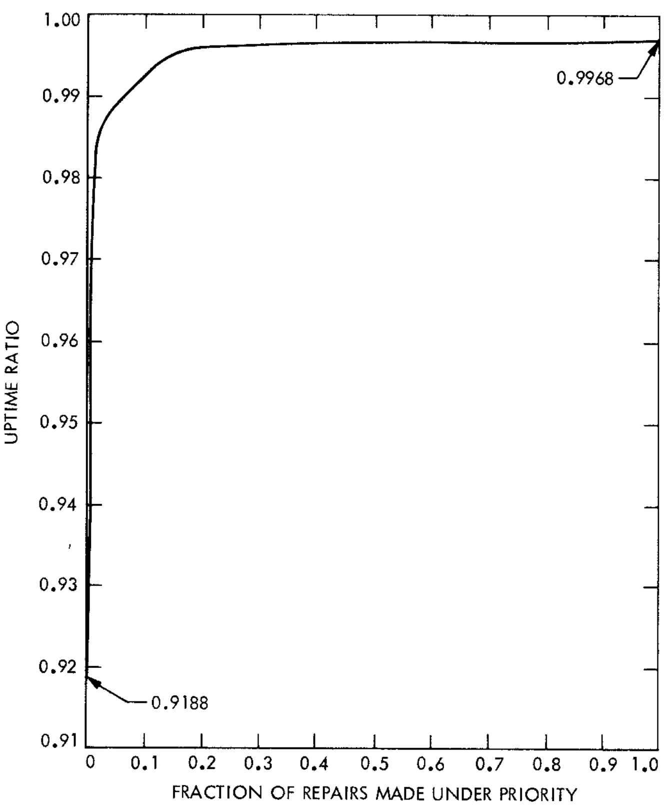

Fig. 2. Improvement in system UTR as the fraction of repairs made under priority increases, μ = 336 h, R = 0.25

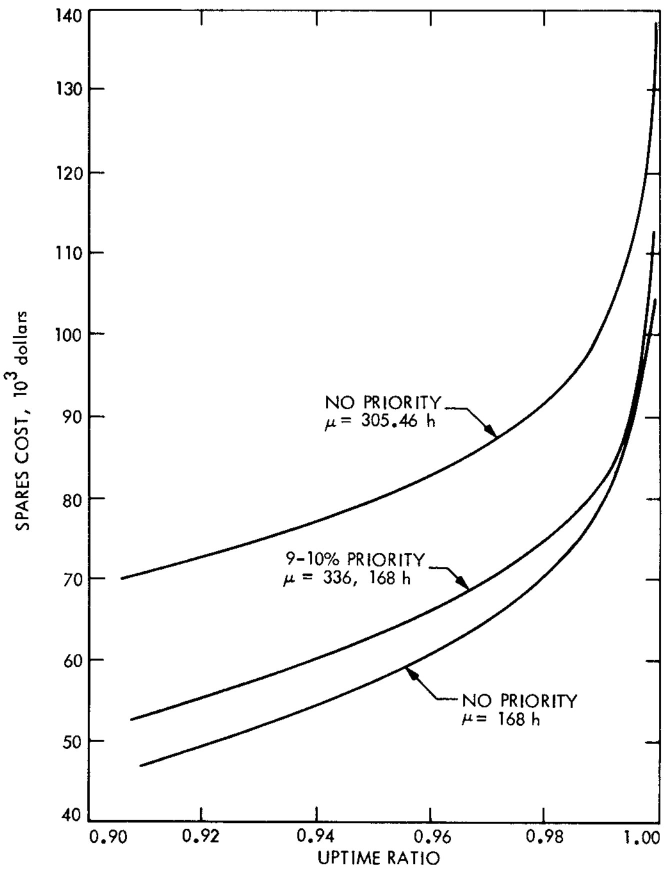

Fig. 3. Cost of spares vs UTR for repairs with no priority and 9–10% repairs made under priority