Journal of Research of the National Bureau of Standards—B. Mathematics and Mathematical Physics
Vol. 63B, No. 2, October–December 1959

Capacity Requirement of a Mail Sorting Device: II[1]

A. J. Goldman

(July 14, 1959)

The combinatorial analysis of a mathematical model of a sorting device suggested by S. Henig is completed. The relevant parameters are r, the number of destinations for the mail, and k, the number of letters entering the device during each cycle of operation. The capacity of the device, if it is never to jam, should be between rk and $rk-(r-1)$ inclusive; arguments indicating that the latter value is preferable are given.

1. Introduction

We deal with a highly idealized mathematical model of a mail sorting device suggested by S. Henig (NBS Electronic Instrumentation Section). For our purposes the operation of the device can be described[2] as follows. Mail (to any of r destinations) enters the system. After k letters have entered, the device "asks itself" to which destination it contains the *most* letters.[3] All letters to this *predominant destination* are then *dropped out* of the device, another k letters enter, and the process continues. The device is said to *jam* if, after a dropout, it is still so "full" that entrance of the next k letters would cause an "overflow."

Investigation of the most appropriate capacity for such a device leads to mathematically interesting problems of two general types. If the capacity is to be so chosen that jamming *never* occurs, then the problems are essentially combinatorial in nature and can be treated mainly by "counting" methods. If, however, the object is the more modest one of keeping the frequency of jamming down to some specified "tolerable level," then rather difficult probabilistic problems arise; this is due to the fact[4] that the behavior of the device constitutes a Markov chain with a great many states, governed by the probabilistic distribution of mail by destinations.

The present paper completes an earlier analysis[5] of the *combinatorial* questions, which showed that

$$(r-1)(k-1)+k=rk-(r-1) \qquad (1)$$

is a possible capacity for the device, and (if the device begins operation empty) is in fact the *minimum* capacity. This left open the possibility of gaining some advantage by using a *greater* capacity than that given in (1) and beginning operation with the device partly or entirely filled. It will be proved below that no capacity in excess of rk should be employed, and that any advantage arising from a capacity greater than $rk-(r-1)$ would be only transient.

The author acknowledges several improvements in the exposition suggested by J. R. Rosenblatt. Unfortunately a somewhat more formal treatment than that of the previous paper seemed required by the greater complexity of the analysis below.

2. Statement of Results

In order to describe our results, it is convenient to define

$$x(t)=\text{number of letters in the device just before the } t\text{th dropout}, \qquad (2)$$

and also to give a precise definition of "capacity."

A non-negative integer C will be called a *capacity* for the device if, should the device begin operation containing no more than C letters, it can never subsequently contain more than C letters under the dropout rule described above; symbolically,

$$x(t_0)\leq C \quad \text{implies} \quad x(t)\leq C \qquad \text{for all} \quad t\geq t_0. \qquad (3)$$

This condition is clearly necessary to insure that the device never jams.

THEOREM 1. *C is a capacity if and only if*

$$rk-(r-1)\leq C.$$

A capacity C will be called *efficient* if there is at least one set of initial contents for the device, with no more than C letters, such that the device might *possibly* contain C letters again at some later time; symbolically, for *some* set of initial contents,

$$x(t_0)\leq C \quad \text{and} \quad x(t)_{\text{max poss}}=C \quad \text{for} \quad t>t_0. \qquad (4)$$

If this condition is not satisfied, then the full capacity C of the device will not actually be needed after the start of operation, a wasteful situation.

THEOREM 2. *C is an efficient capacity if and only if*

$$rk-(r-1)\leq C\leq rk.$$

A capacity C will be called *essentially efficient*[6] if there is at least one set of initial contents with no more than C letters such that the device with nonzero probability will contain C letters infinitely

[1] Part of a project sponsored by the Post Office Department, Office of Research and Development.
[2] The physical device can also operate under dropout rules other than the one described below.
[3] A rule for breaking "ties" between destinations is also required.
[4] Pointed out (unpublished memorandum, July 1956) by J. R. Rosenblatt (NBS Statistical Engineering Laboratory).
[5] B. K. Bender and A. J. Goldman, Capacity requirement of a mail sorting device, J. Research NBS **62**, 171 (1959) RP2948.
[6] Use of the term "essential" here, and of the subscript "ess" in (5), was suggested by a roughly similar usage in measure theory.

often [7]; symbolically, for some set of initial contents,

$$x(t_0) \leq C \quad \text{and} \quad x(t)_{\max\ \text{ess}} = C \quad \text{for} \quad t > t_0. \tag{5}$$

This definition is strengthened to that of a *uniformly essentially efficient capacity* by requiring that the stated condition hold for *all* initial contents of no more than C letters; i.e., that

$$x(t_0) \leq C \quad \text{implies} \quad x(t)_{\max\ \text{ess}} = C \quad \text{for} \quad t > t_0. \tag{6}$$

THEOREM 3. *The only essentially efficient capacity is* $rk-(r-1)=C$; *it is also uniformly essentially efficient.*

The proofs of these theorems are given in section 3. Section 4 contains a very crude probabilistic estimate related to theorem 3.

3. Proofs

3.0. Preliminaries

It is helpful to define

$$M(t) = \max_i x_i(t)$$

where

$x_i(t)$ = number of letters to the ith destination in the device just before the tth dropout.

LEMMA 1. $M(t) > k$ *implies* $x(t+1) < x(t)$.
LEMMA 2. $x(t) > rk$ *implies* $x(t+1) < x(t)$.
LEMMA 3. $M(t) = k$ *implies* $x(t+1) = x(t)$.
LEMMA 4. $M(t) \geq k$ *implies* $x(t+1) \leq x(t)$.
LEMMA 5. $x(t) \geq rk-(r-1)$ *implies* $x(t+1) \leq x(t)$.
LEMMA 6. $x(t_0) \leq rk-(r-1)$ *implies* $x(t) \leq rk-(r-1)$ *for all* $t \geq t_0$.
LEMMA 7. $x(t_0) \geq rk-(r-1)$ *implies* $x(t) \leq x(t_0)$ *for all* $t \geq t_0$.

Lemma 1 is proved by observing that if $M(t) > k$, then more than k letters leave the device in the tth dropout, whereas only k letters enter before the $(t+1)$-st dropout. Since there are only r destinations, $x(t) > rk$ implies $M(t) > k$, and so lemma 2 follows from lemma 1. Lemma 3 is an obvious consequence of the dropout rule, and lemma 4 is obtained by combining lemmas 1 and 3.

To prove lemma 5, note that since there are only r destinations,

$$x(t) \geq rk-(r-1) > r(k-1) \quad \text{implies} \quad M(t) > k-1,$$

so that lemma 4 can be applied. Lemma 6 merely states the fact [5] that $rk-(r-1)$ is a capacity. In proving lemma 7, two possibilities must be considered. If, on the one hand,

$$x(t) \geq rk-(r-1) \qquad \text{for all} \quad t \geq t_0,$$

then by lemma 5, $x(t) \leq x(t_0)$ for all $t \geq t_0$. If, on the other hand, there is a $t_1 > t_0$ such that

$$x(t) \geq rk-(r-1) \qquad \text{for} \quad t_0 \leq t < t_1$$

but

$$x(t_1) < rk-(r-1),$$

then by lemma 5, $x(t) \leq x(t_0)$ for $t_0 \leq t < t_1$, and by lemma 6

$$x(t) \leq rk-(r-1) \leq x(t_0) \qquad \text{for all} \quad t \geq t_1,$$

so that again

$$x(t) \leq x(t_0) \qquad \text{for all} \quad t \geq t_0.$$

3.1. Proof of Theorem 1

To prove the "only if" statement, suppose $C < rk-(r-1)$ and $x(t_0) \leq C$. If, on the one hand, $x(t_0) = 0$, then (see *op. cit.* footnote 5, p. 79) there is a sequence of possible events leading to a situation in which

$$x(t) = rk-(r-1) > C; \qquad t \leq t_0 + (r-1)(k-1) + 2.$$

This violates condition (3) and so C is not a capacity. If, on the other hand, $x(t_0) > 0$, then a sequence of possible events can be constructed leading to a situation in which

$$x(t_1) = 0; \qquad t_1 \leq t_0 + C. \tag{7}$$

To do this, choose the k letters entering the device between the tth and $(t+1)$-st dropouts, for $t_0 \leq t < t_1$, to consist entirely of letters to some one destination for which $x_i(t) > 0$. Then $M(t) > k$ for $t_0 \leq t < t_1$, and so by lemma 1 the number of letters in the device steadily decreases until condition (7) is satisfied. From this point we may argue as in the case $x(t_0) = 0$ (now using t_1, instead of t_0), again obtaining the conclusion that C is not a capacity.

To prove the "if" statement, suppose

$$C \geq rk-(r-1) \quad \text{and} \quad x(t_0) \leq C.$$

If $x(t_0) \leq rk-(r-1)$, then by lemma 6

$$x(t) \leq rk-(r-1) \leq C \qquad \text{for all} \quad t \geq t_0,$$

while if $x(t_0) \geq rk-(r-1)$, then by lemma 7

$$x(t) \leq x(t_0) \leq C \qquad \text{for all} \quad t \geq t_0;$$

in either case condition (3) is satisfied, and so C is a capacity.

[7] If this condition is not satisfied, then with probability one the device will, from some point on, fail to use its full capacity.

3.2. Proof of Theorem 2

To prove the "if" statement, suppose

$$rk-(r-1) \leq C \leq rk.$$

The initial contents of the device can clearly be chosen so that $x(t_0)=C$ and $M(t_0)=k$. Among the destinations with $x_i(t_0)=k$, select the one "preferred" by the tie-breaking rule at the t_0-th dropout; let k letters to this destination enter the device between the t_0-th and (t_0+1)-st dropouts. Then $M(t_0+1)=k$; among the destinations with $x_i(t_0+1)=k$, select the one preferred by the tie-breaking rule at the (t_0+1)-st dropout, and let k letters to *this* destination enter the device between the (t_0+1)-st and (t_0+2)-nd dropouts, etc. This construction results in $M(t)=k$ for all $t \geq t_0$, so that, by lemma 3,

$$x(t)=x(t_0)=C \qquad \text{for all} \quad t \geq t_0.$$

Hence C is an efficient capacity.

The "only if" statement will be proved next. By theorem 1, no $C<rk-(r-1)$ can be a capacity; thus no such C can be an efficient capacity. Suppose $C>rk$ and $x(t_0) \leq C$. There are three possibilities to be considered. If $x(t_0) \leq rk-(r-1)$, then by lemma 6

$$x(t) \leq rk-(r-1) \leq rk < C \qquad \text{for all} \quad t \geq t_0$$

so that C is not an efficient capacity. If

$$rk-(r-1) < x(t_0) \leq rk,$$

then by lemma 7

$$x(t) \leq x(t_0) \leq rk < C \qquad \text{for all} \quad t \geq t_0,$$

so that C is not an efficient capacity. Finally, if $x(t_0)>rk$, then by lemma 2 there is a $t_1>t_0$ (with $t_1 \leq t_0+C-rk$) such that

$$rk < x(t) < x(t_0) \leq C \qquad \text{for} \quad t_0 < t < t_1$$

and

$$x(t_1) \leq rk;$$

the arguments for the first two possibilities can be applied to t_1 to yield

$$x(t) \leq x(t_1) \leq rk < C \qquad \text{for all} \quad t \geq t_1$$

so that $x(t)<C$ for all $t>t_0$ and again C is not an efficient capacity.

3.3. Proof of Theorem 3

First, $rk-(r-1)$ *is* a uniformly essentially efficient capacity. To see this, note that the construction used above to prove the "only if" part of theorem 1 shows that if

$$x(t_0) \leq rk-(r-1)$$

then with probability one [8] there is a $t_1>t_0$ for which

$$x(t_1)=rk-(r-1);$$

by the very same argument, with probability one there is a $t_2>t_1$ for which

$$x(t_2)=rk-(r-1),$$

and so on. In fact, for $rk-(r-1)$ the phrase "with nonzero probability" in the definition of "uniformly essentially efficient capacity" could actually have been strengthened to "with probability one."

Second, no $C>rk-(r-1)$ is an essentially efficient capacity. To prove this, observe that by theorem 2 attention can be confined to the case

$$rk-(r-1) < C \leq rk.$$

Suppose $x(t_0) \leq C$. If

$$x(t_1) \leq rk-(r-1) \qquad \text{for some} \quad t_1 \geq t_0,$$

then by lemma 6

$$x(t) \leq rk-(r-1) < C \qquad \text{for all} \quad t \geq t_1,$$

so that the device cannot contain C letters infinitely often. Thus we may assume that

$$x(t) > rk-(r-1) \qquad \text{for all} \quad t \geq t_0, \tag{8}$$

so that, by lemma 5,

$$x(t+1) \leq x(t) \qquad \text{for all} \quad t \geq t_0. \tag{9}$$

If $x(t_0)<C$, then by (9)

$$x(t) \leq x(t_0) < C \qquad \text{for all} \quad t \geq t_0,$$

so that the device can never contain C letters. Therefore assume that $x(t)_0=C$. If

$$M(t_1) > k \qquad \text{for some} \quad t_1 \geq t_0,$$

then by lemma 1 and (9)

$$x(t) \leq x(t_1+1) < x(t_1) \leq x(t_0) = C \qquad \text{for all} \quad t > t_1,$$

so that the device cannot contain C letters infinitely often. Therefore assume

$$M(t) \leq k \qquad \text{for all} \quad t \geq t_0.$$

From (8) and the fact that there are only r destinations it follows that $M(t) \geq k$, and so $M(t)=k$. Thus the only possibility for the device to contain C letters infinitely often is given by

$$x(t)=C, \qquad M(t)=k \qquad \text{for all} \quad t \geq t_0;$$

this clearly has probability zero.

[8] The construction shows that if $x(t_0) \leq rk-(r-1)$, then for any block of $[(r-1)(k-1)+2+rk-(r-1)]$ values of $t \geq t_0$, there is a nonzero probability (independent of the particular block) that $x(t)=rk-(r-1)$ at least once.

4. A Probabilistic Estimate

Suppose C is an efficient capacity which is *not* essentially efficient; i.e. (according to theorems 2 and 3),

$$rk-(r-1)<C\leq rk.$$

We know that with probability one, there is a $t_1\geq t_0$ such that

$$x(t)<C \qquad \text{for all} \quad t>t_1.$$

It is of interest to estimate how early the first such t_1 (the last moment at which the full capacity of the device is used) might occur. In this primarily combinatorial paper, we will be content with an exceedingly crude probabilistic estimate.

Denote by E the event

$$x(t_0)=C \qquad \text{and} \qquad M(t_0)=k.$$

By the proof of theorem 3, $t_1=t_0$ unless E occurs. Therefore the function of t_1 to be estimated is

$$P(t_1)=\text{Prob}\{x(t)=0 \quad \text{for some} \quad t>t_1|E\}, \tag{10}$$

where Prob$\{A|B\}$denotes the conditional probability of A given B.

Let the relative frequencies (or probabilities) of letters to the different destinations be denoted by

$$0<f_1\leq f_2\leq \ . \ . \ . \ \leq f_r,$$

let

$$C=r(k-1)+s+1; \qquad 0<s<r,$$

and let

$$F_s=f_1+f_2+ \ . \ . \ . \ +f_s.$$

It will first be shown that, for any $t>t_0$,

$$\text{Prob}\{x(t)=C, \quad M(t)=k|x(t-1)=C, \quad M(t-1)=k\} \leq (1-F_s)^k \tag{11}$$

Suppose

$$x(t-1)=C, \qquad M(t-1)=k.$$

Since

$$x(t-1)=C>r(k-1)+s=(r-s)(k-1)+sk,$$

it follows that $x_i(t-1)=k$ for at least $s+1$ destinations. The letters to some one of these destinations leave the device in the $(t-1)$-th dropout, and $M(t)>k$ if any of the k letters which next enter the device go to any of the remaining s destinations with k letters each. Since the sum of any s frequencies f_i is at least F_s, the inequality (11) holds.

By the proof of theorem 3, for any one $t>t_0$, both $x(t-1)=C$ and $M(t-1)=k$ hold if and only if

$$x(t')=C \qquad \text{and} \qquad M(t')=k \qquad \text{for} \quad t_0\leq t'<t.$$

Thus the probability on the left side of (11) can be rewritten

$$\text{Prob}\{x(t)=C, \ M(t)=k|x(t')=C, \ M(t')=k \text{ for } t_0\leq t'<t\},$$

so that the inequality (11) yields

$$\Pi \ \text{Prob}\{x(t)=C, \ M(t)=k|x(t')=C, \ M(t')=k t_0<t<t_1 \text{ for } t_0\leq \ t'<t\}\leq (1-F_s)^{k(t_1-t_0)}.$$

This last product is equal to

$$\text{Prob}\{x(t)=C, \quad M(t)=k \quad \text{for} \quad t_0<t\leq t_1|E\},$$

which, according to the proof of theorem 3, is an upper bound for $P(t_1)$ and so we have proved that

$$P(t_1)\leq (1-F_s)^{k(t_1-t_0)}. \tag{12}$$

If for example $k=2$ and $F_s=0.001$, then after 500 dropouts the probability that the device will make any further use of its full capacity is less than 0.50.

WASHINGTON, D.C. (Paper 63B2–10)