Mechanization of Train Seat Reservation System (Report 1)

Yutaka OHNO †

1. Introduction

There is a limitation to the human control of a sprawling system. This has motivated recent developments in mechanization and automation of such a system. Seat reservation system for passenger traffic represents one of the earliest developments toward mechanization. Such development being based on the electronic computer technique, a system analysis of reservation operation and a technico-economic investigation of the functions to be performed by the mechanized system are pre-requisite to realization of such mechanization; particularly, full attention should be paid to the reliability of the system.

The present paper deals with the investigations and analysis involved with mechanization of seat reservation system with reference to the practical performance of MARS-1 realized as the result thereof.

2. Peculiarities of seat reservation process

Apparently the seat reservation process is a kind of inventory control. Upon request from a client the number of a seat fitting certain conditions (train, date of boarding, travel section, class) is picked up from the file of seats to be reserved for him. The kinds and quantities of the seats classified according to different conditions amount to such an enormous total that it would need a great deal of man-hours to pick up a certain seat out of the file and correct the file accordingly. Meanwhile, clients are scattered and they will visit the ticket windows at random for seat reservation.

Each window will phone to the Ticket Reservation Centre which possesses such a file for inquiry about the availability of the requested seat. Occurrence of client requests tends to be concentrated in a certain time zone (Fig. 1) and the number of such requests varies by the day. Thus, the situation of reserved seats on the file differs from moment to moment; and in the crowded time zone the time

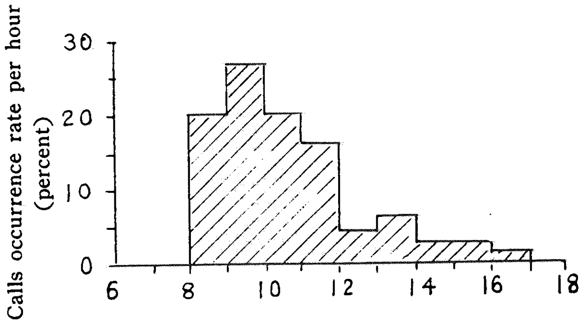


Fig. 1 Time distribution of calls occurrence.
100 % : 8~17 hr

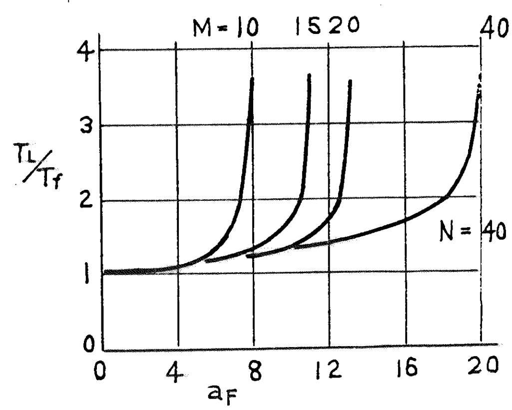


Fig. 2 Relation between T_L/T_f and a_F at N=40
where M = No. of clerks
N = No. of files
T_L = Mean holding time of a telephone channel, min.
T_f = Mean processing time of a seats file, min.
$a_F = c.\ T_f$
c = No. of calls for all N files per minute

† Automation Lab.

required for dealing with one request must be as short as possible, otherwise the line-holding or waiting time for the telephone will be extended with the result that a long queue is formed at the window, leading to deterioration of service to the clients. In order to satisfy the client's inquiry with promptness, the contents of the seat file, as it changes from moment to moment, must be grasped with accuracy and alacrity ; and the file must be so maintained all the time as to reflect an accurate situation of seat reservation at a given moment. While the quantity of seats is small, the operation may be humanly controlled satisfactorily by utilizing the communication means like the telephone; but when their quantitity exceeds a certain limit, it will already be beyond human control and some extent of improvement or expansion of facilities will never be able to cope with the situation (Fig. 2).

3. Different methods of mechanization and their problems

Many methods of mechanization of the seat reservation system are conceivable. Their problems are to be investigated as follows.

3.1 Requirements for mechanized system

As evident from the statement under 2, the ticket reservation window must be able to give an answer to the client's inquiry about the availability of a certain seat the client wants in sufficiently short time through simple process; the answer should, if necessary, be produced in the form of an automatically printed ticket. To make this possible, the automatic, high-speed reading and writing of the informations on the reservation situation from or into the seat file must be realized. The file, if the process of reserving seats satisfying certain conditions is finished rapidly and its rate varies, should be corrected for every reservation. Moreover, such a file is located far from each window, so establishment of an information channel between the two is indispensable. The client's requests to be carried over such a channel being in the nature of varied and casual occurrence, the file-handling device will essentially be a multi-input-output one.

In some cases it will suffice to so control the number of reserved seats that it may not exceed the rated capacity (Rated capacity-reservation system) ; in others, individual seats must be specified and reserved (Individual seat-reservation system). In either case each section of line between stations is treated independently for seat reservation. Under such treatment the rate of seat occupancy will be improved and such order of reservation as leaves as few seats unsold as possible can be devised.

Under a system involving many widely scattered windows, a failure of the control center dealing with the file will have grave consequence, sometimes resulting in a false action remaining uncorrected. Thus, great importance should be attached to the problem of reliability.

Meanwhile, the system and quantities of seat reservation being subject to changes with progress of the times, some flexibility should be reserved for any method.

3.2 System of mechanization

3.2.1 Input-output system

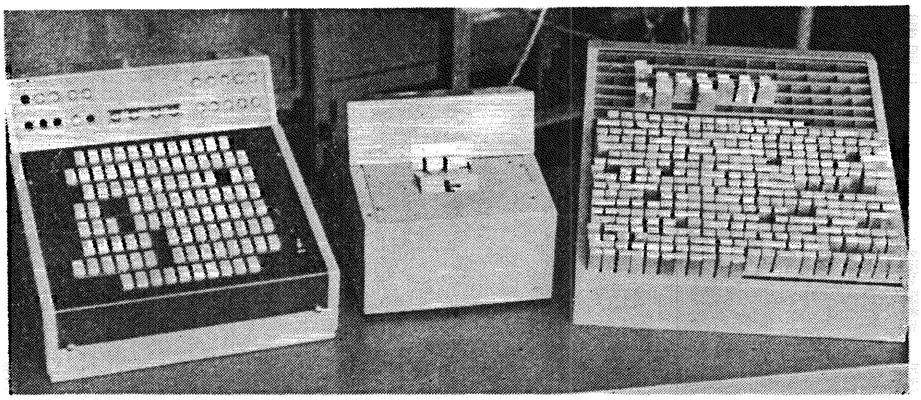
The input-output device, which is set with the requirement of seat put by a client as an input and displays an answer to this input, is commonly called Agent Set. There are several types of Agent Sets as used in the aircraft seat reservation system: One using a teleprinter (or an electrical typewriter) ; one using push buttons and keys for setting of informations (with coupled use of coding plates and optical means for designating the air-port and flight number) ; or mark-sensing system, in which a card is marked with a pencil and the marking is automatically read.

Even these will not satisfactorily serve the purpose, if the answer is to be produced in the form of a printed ticket. The names of trains and stations in Japanese letters may be printed as tickets by the Teleprinter system, if "Kana" letters are adopted instead of Chinese characters; but even this will not be sufficiently adequate for the purpose, because of the skill required in the operation of the system.

Thereupon, the Japanese National Railways has decided to adopt for its projected seat reservation system such a type of Agent Set, in which all the conditions that can be expressed numerically are set by the push-button device; and as for the names of stations and trains, their coded stamps are prepared, and they are picked and inserted in the printing machine and after the numerical answer from the Control Center is printed out, they are additionally printed; this will facilitate the operation, and cut the time of operation, with increased flexibility and reduced cost. The prototype of this system, as illustrated in Fig. 3, has been proved satisfactory in its tentative exploitation.

3.2.2 File processing system

The manner of processing the file will, as stated before, depend on the number and progress of reservations and cancellations; the processing speed will not be elevated unless the file is mechanized. If, as is practiced by the Pennsylvania Railroad, U.S.A., industrial television is utilized for simultaneous processing at many windows of the same file—so-called Parallel Processing, the efficiency of file



Button set Ticket printer Coded stamps of trains & station names

Fig. 3 New type agent set for JNR reservation system

processing will be substantially enhanced, but so long as it relies on manual labor, no drastic improvement can be expected therefrom.

A large-capacity memory unit of the electronic computer can be made available for the file of train seats. It can fully satisfy the requirement for rapidity of file processing. However, in order that the file may be checked for every reservation and its content may accordingly be corrected, real time or "instant" operation of the file processing device should be realized. Realization of such operation has involved a number of technical difficulties, the solution of which is not always easy. When there is no great daily variance in the situation of seat reservation and the balance of seats not yet reserved can be predicted with fair accuracy, there will be no need for the checking and correction of the file on every reservation; the results of reservations and cancellations may be carried out later for a certain appropriate interval of time. Then, not being a "real time" processing, the system can be realized with comparative ease by the present level of technique. At present, as one of such delayed processing system there is one coupled with use of availability panel (Fig. 4), which contains only the informations on the availability of seats meeting certain conditions. This will give a very much simplified seat file and at each window this panel will be consulted upon the client's inquiry. A certain quantity of reservations and cancellations are reported to the Center Processor, where the detailed Seat Information File will be corrected for the corresponding time interval. When the non-reserved seats meeting certain conditions are sold out, the corresponding information at the availability panel will be changed.

Uuder such delayed processing system, a precise information on the situation of seat reservation cannot be gained at each window, therefore the mechanization under such system cannot be called perfect. Particularly, when there is a wide variance in the daily situation of seat reservation meeting respective conditions, or when the individual seat reservation system is adopted, no reservation can be

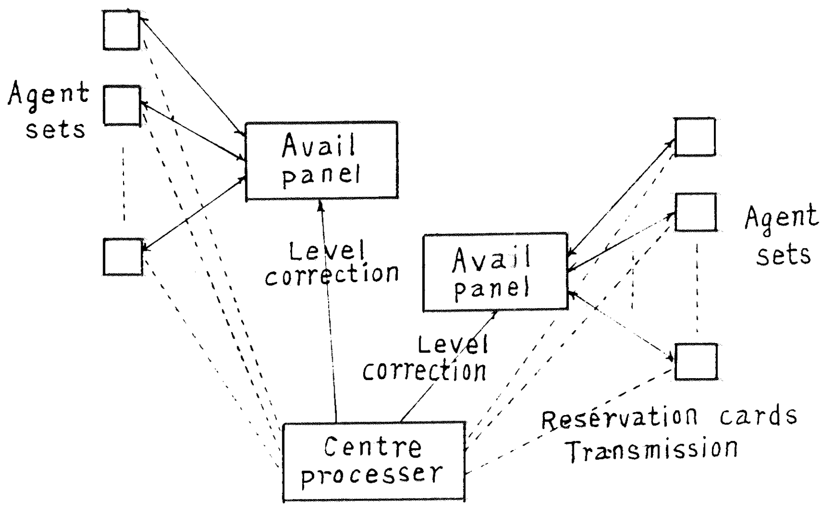


Fig. 4 Availability panel system

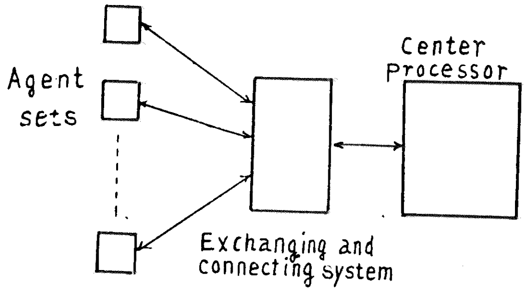


Fig. 5 Real time system

made without detailed information on the situation of reservation. Thus, it would be imperative to adopt the "real time" processing system. Fig. 5 illustrates the conception of such system as applied to train seat reservation. The technical problems to be solved for this system to materialize include:

(1) Realization of a high-access memory unit with more than several million bits to be used for the seat file.

(2) Matching of the number of requests occurring at random in each Agent Set with processing capacity of the Center Processor (central device for processing the file).

(3) Transmission of information between Agent Sets and Center Processor.

(4) Monitoring and checking of various operations.

(5) Improvement of reliability and availability and establishment of emergency countermeasures.

The Center Processor, being a sort of specialized electronic computer of multi-input output type, will not be constructed on the basis of a conventional general purpose machine. Meanwhile, whether its programming shall be a fixed or a stored one will be decided depending on how much flexibility shall be imparted to the system. However, in view of the latest development in computer technology, there is a prospect of such special processing becoming realizable on a general purpose stored program type machine.

4. Problems in "real time" processing

Two important problems in "real time" processing are to be discussed here: the balance between input and processing, and the reliability of system.

4.1 Request occurrence and its processing

The Centre Processor handles numerous requests taking place at random on the "time sharing" basis. If the processing speed is too slow, the requests at the peak time will not be able to be disposed of in time with the result that a queue is formed at the window; if it is too fast, the system will be economically unprofitable.

(1) Processing speed and construction of system

The time required for the Center Processor to handle one request will be negligibly short, but it will take considerable time to carry out an exchange of request and answer informations. This will require installation of several sets of tranceiver registers at the Control Center, where the informations of requests as received at the register will be successively fed into the Processor and the answers thereto will be set at the same register; the register, at which the answer has been set, will instantly begin to transmit the answer to the Agent Set.

Putting T_p=Processing time for each request (sec),

C_{Max} =Frequency of request occurrence in unit of peak time (sec^{-1}),

T_R=Holding time of tranceiver register (total of time for sending and receiving the information, the processing time and the process-waiting time) (sec), and

M=Quantity of tranceiver registers,

the following relationship must hold.

$$1/T_P \geq M/T_R \geq C_{Max}$$

There is a limit to the value of T_P and the value of C_{Max} may become larger. Therefore, if, under the condition $1/T_P < C_{Max}$, the seat file is divided into N sets, each of them to be processed in parallel, or N sets of processors are provided, or the processing work within the processor is so designed that T_P may become essentially equal to $1/N$, there will be need to have the condition $N/T_P \geq M/T_R \geq C_{Max}$. T_P will mainly depend on the construction of the file; when a large-capacity magnetic drum is used, T_P will take a value of several tens of m sec, being restricted by its access time; usually the figure will run up to the order of 100 m sec. Putting T_P=0.1 sec, C_{Max}=8 sec^{-1}, T_R=4 sec, the condition will be $40 \geq M \geq 32$; accordingly the value of M will be fairly large.

At $N \neq 1$ it would be conceivable to concentrate all of N sets of file at the Center or scatter the installation of N sets of processors in different places; which system to adopt should be decided according to the distribution of request occurrence of the condition of information-transmitting system.

4.1.2 Problem of waiting in the system

The Center equipped with N sets of file is provided with M sets of tranceiver registers; each register is connected to the necessary set of files according to the request information which has been set in it. Meanwhile, each Agent Set can be connected to any of idle registers. Uuder such arrangement, waiting will be inevitable with the scheme of the system benig as illustrated in Fig. 6. Analysis of this waiting problem in generalized form will be considerably difficult; but if some special conditions are assumed, a simple, approximate analysis will be possible. Namely, the problem is divided into two; one concerning the phase, in which the Agent Sets are connected through transmission channels to the tranceiver registers; and one concerning the next phase, in which the informations, after having been set at the tranceiver register, are connected to the necessary file for processing, that is, the waiting within the Center Processor.

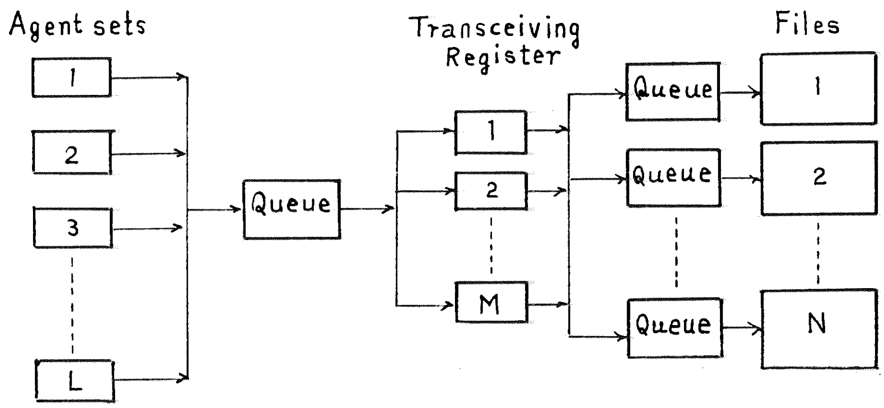


Fig. 6 Queueing in the real time system

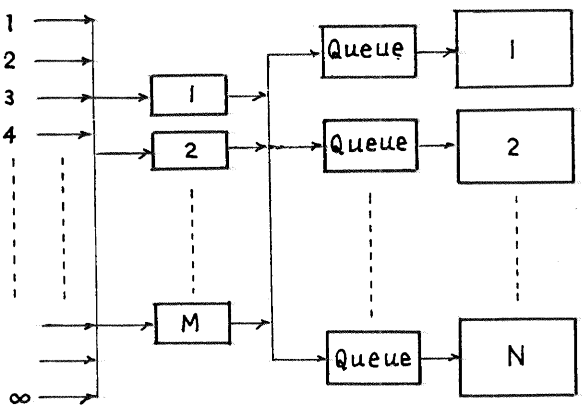


Fig. 7 Queueing in the centre processor

(1) Waiting within the Center Processor

Fig. 7 represents the state of the system under this waiting. In this connection the following is assumed.

The distributions of "calls" on the receiving register follow the Poisson's law, while the file-processing time has an exponential distribution. Moreover, the "calls" are handled in the order of their occurrence. The problem will be simplified, if the following condition is added to the above assumptions: The "calls" on the tranceiver register shall be so distributed that each file in N sets may get uniform load; this is the condition to minimize the expense of file processing per call and one that should be taken into consideration for practical designing.

Putting λ = Apparent frequency of "calls" per unit time (sec^{-1})
ν = Reciprocal of mean processing time of file (sec^{-1})

$$Q_{M.N}(a_0) = \sum_{h=0}^{\mu} {}_{N+h-1}C_h \cdot a_0^h, \quad \text{and} \quad a_0 = \lambda/\nu$$

We get the following results from the equation of probability balance.

The average waiting time for the register-connected "call" to be file-processed, T_W sec, will be given by

$$T_W = \frac{a_0}{\nu[1-P(M)]} \cdot \frac{Q_{M-1.N+1}(a_0)}{Q_{M.N}(a_0)}$$

where

$$P(M) = \frac{Q_{M.N-1}(1)}{Q_{M.N}(a_0)} \cdot a_0^M$$

The probability of necessity for waiting, P_W will be given by

$$P_W = 1 - \frac{1}{1-P(M)} \cdot \frac{Q_{M-1.N-1}(a_0)}{Q_{M.N}(a_0)}$$

The true average number of "calls" on this system, λ_R, will be given by

$$\lambda_R = \lambda \cdot [1 - P(M)] \ \text{sec}^{-1}$$

Fig. 8 plots the values of T_W at N=2. In a system designing, λ_R is provided as one of the conditions and λ is an apparent value used for calculation.

(2) Waiting in transmission system

The number of Agent Sets is sufficiently large, so that if the number of tranceiver registers is put as M, the problem will be to determine the waiting time before the request call, occurring at a certain Agent Set, can be connected to any of these registers when a request for connection to the Center is issued. Fig. 9 shows the arrangement of this system. The average holding time for a register, T_R, is given by

$$T_R = T_T + T_W + T_P \ \text{(sec)}$$

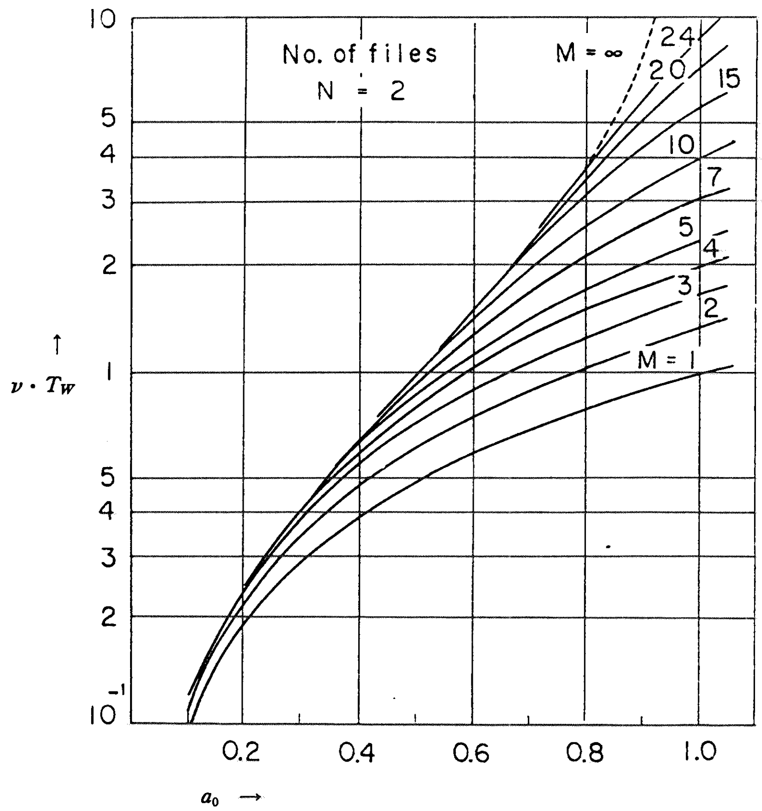


Fig. 8 Mean waiting time in the centre processor at N=2

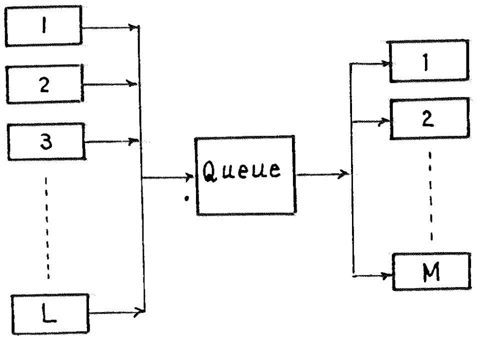


Fig. 9 Queueing in the transmission system

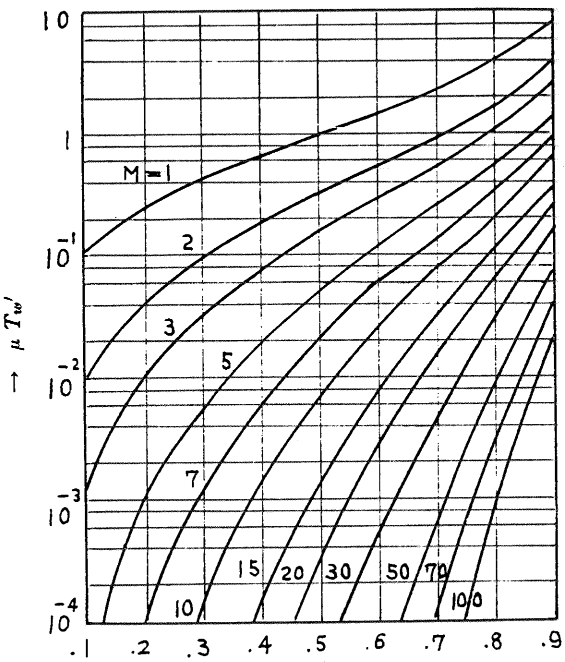


Fig. 10

where T_T is the time required for information transmission.

Assuming that the register-holding time follows a distribution of $e^{-\mu t}$ (where $\mu=1/T_R$); the calls on the agent sets do a Poisson distribution of average λ_R; and the connections to registers take place in the order of "call" occurrence, the Erlang's waiting formula can be directly applied.

Putting T'_W=connection-waiting time (sec)

$A_R=\lambda_R/\mu$ and M=number of registers,

the relationship as illustrated by Fig. 10 can be obtained.

Based on the above-mentioned relationships, the number, M, of tranceiver registers, which is optimum with no waste, can be determined.

4. 2 Enhancement of reliability

When the seat reservation system is to be composed as a "real time" processing one, effective measures for guarantee of the reliability considering the following character of the system should be taken;

(1) Seat reservation operation shall be performed throughout the day with hardly any interruption.

(2) The file-processing at peak time shall be finished within an exceedingly limited interval (say, 50 m sec~100 m sec).

(3) Mistaken file-processing is hard to be corrected, unless revealed and corrected on the spot, because the client leaves the window.

(4) Destruction of the seat file will entail a wide-spread chaos with great damage, because the contents of the file represent the results of all reservation processings executed in the reservation period and will be impossible to reproduce.

4. 2. 1 Duplicated equipment system

Duplication of equipment in a system will be one of the useful means to improve its reliability. It is theoretically proved very effective in elevating the reliability of system to split a system into as small circuits as possible and arrange several sets of same circuits in parallel for the sake of duplicity. Such arrangement, involving many problems for realization, is found impracticable. But to provide duplicate sets of same or equivalent devices in a system will be practical and such arrangement is finding frequent applications.

For this arrangement, the system can be composed variously. Here it is supposed that N sets of equipment (actually the case of $N=2$ is most frequent) are provided and the following methods of operation will be taken up for discussion;

(1) Standby system — One set is normally in operation and another set is standing by ready to replace it instantly when it fails.

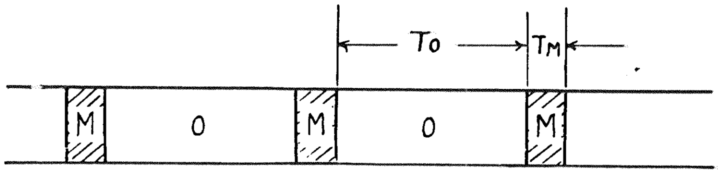


Fig. 11 Periodic-maintenance operation
M : Maintenance, O : Operation

Table 1 Comparison of standby and parallel systems of operation
($N=2$, $(\lambda+\mu)\ T_0 \gg 1$, $\mu \gg \lambda$)

Itemes	Standby		Parallel
Reliability $R(t)$	$e^{-\lambda t}(1+\lambda t)$	>	$e^{-\lambda t}(2-e^{-\lambda t})$
Availability $\bar{A_2}$	$1-1/2\ (\lambda/\mu)^2$	>	$1-(\lambda/\mu)^2$
Probability of single operation P_s	1.00	>	$2\lambda/\mu$
Operations check process time per request T_0	self-check		mutual check
Guarantee of file	imperfect	>	nearly perfect

(2) Parallel system — All of N sets are simultaneously placed in service to perform the same function, and they mutually check their performance.

Analysis will be made to see which of these two can better contribute to the reliability of the system.

4.2.2 Comparison of standby and parallel systems

Whichever system may be adopted, there will be normal schedule time of operation, T_0; and after that comes the time T_M for maintenance and idling (often this is entirely lacking), followed by the operation time T_0 again and so on (Fig. 11). In the beginning of T_0 all of the N sets are in the state of being able to function. If these assumptions are made, the occurrence of a failure in one set will follow Poisson's distribution. therefore the probability (reliability) of a set actually in operation at $t=0$ finding itself capable of operation at $t=$over t will be

$R(t)=e^{-\lambda t}$, where $1/\lambda=$ average life.

The probability (maintainability) of a possible fault being removed in less than the time t is given by

$M(t)=1-e^{-\mu t}$, where $1/\mu=$ average repair time

The results of calculations executed on similar consideration to the queueing theory on the basis of these relationships are shown in Table 1; the reliability in this Table was for the sake of simplification computed on the assumption that a failure in the operating machine will not be attended to for repair until after a certain period (that is, time for maintenance T_M) has elapsed.

Availability is defined as follows.

The probability of a system being in the operable state at an arbitrary moment t within the planned period of normal operation T_0 is put as $A(t)$. $\bar{A}_N$ represents a mean value for the time interval T_0 in a system with parallel arrangement of N sets and $\bar{A}'_N$ is the corresponding value in one of stand-by arrangement of N sets. In regard to both reliability and availability, the latter system is superior to the former, but with a small margin.

If it is considered that in the common electronic computer $1/\mu=1$ hr or thereabout and $1/\lambda=$several 10 hr or more, where as in the seat reservation system $T_0=$more than 20 hr, we have $(\lambda+\mu)\ T_0\gg1$, $\mu\gg\lambda$, so in Table 1 this was adopted for simplification of equations.

If $1/\lambda=40$ hr, $1/\mu=1$ hr, we get $\bar{A}_1=\bar{A}_1'=0.975$, $\bar{A}_2\fallingdotseq0.99937$, $\bar{A}_2\fallingdotseq0.99963$; accordingly great improvement is obtained by putting $N=2$.

The probability of single operation is $P_s=1$ for the standby system in which one set is all the time in single operation; this probability is extremely low for the parallel system. This fact is related to the checking of the operation described in the following.

Under real time processing system, false action must be detected on the spot. Therefore the checking of actions should receive full attention in the interest of better over-all reliability of the system. In the standby system which necessarily relies on single operation, the actions must be checked constantly. This checking will be performed by a set of processors using its own program; there will be some cases where considerable difficulty is felt in checking the logical actions other than arithmetic operations and in such cases a large number of program steps will be involved, prolonging the processing time.

By contrast, in the parallel system it suffices that the results of operations in each operating device are mutually checked by coincidence and such checking takes short time with nearly 100% rate of false action discovery; accordingly there is no great prolongation of processing time on account of the checking.

Another problem in the real time processor of information is how to cope with a possible destruction of the master file. This file for the seat reservation system holds the recordings of all operational results covering the whole period of reservation. Therefore its loss by destruction will mean a great chaos which will not be corrected by a mere repairing of the memory unit and will cost considerable money for correction. Under stand-by operation the safe custody of the "real time" process file is not guaranteed, there being only one set of it, whereas under parallel operation it is seldom that there is only one set of such file, hence total destruction of informations is practically unthinkable. From these considerations it may be concluded that the parallel system rather than the standby will be suitable for seat reservation.

The parallel system is admittedly disadvantageous as compared with the standby in that the former is less adaptable for jobs other than the handling of inquires from Agent Sets. If, however, the latest feature of "program interruption" in the electronic computer is exploited, the difficulty can be liquidated.

4.2.3 Determination of the number of sets to be duplicated

It is doubtless that from the standpoint of reliability, the larger the number of duplications N, the better; but this should also be studied from the standpoint of economics.

One way of economic evaluation of the system is to weigh the total operational cost of the system against the possible damage to the enterprise caused by a failure as well as the routine operational cost and the depreciation of primary cost.

Putting K = Routine operational cost of 1 set of devices per unit time (¥/hr),

Q = Loss per unit time due to stoppage of operation (¥/hr), and

C_N = Average cost of parallel operation with N sets per unit time (¥/hr),

we may get

$$C_N = N.K+Q\cdot(1-\bar{A}_N).$$

Thereupon, if the value of N that can satisfy $C_N-C_{N-1}=0$ is r, an integral that is smaller than this value and nearest to it will be the number of sets sought for the system:

$$r=\frac{\log\dfrac{Q\cdot\mu}{K(\lambda+\mu)}}{\log\left(\dfrac{\lambda+\mu}{\lambda}\right)}$$

The values of K and Q, particularly that of Q, are often fairly difficult to calculate; in the seat reservation system the values will widely differ depending on the factors to be considered.

In most cases it would be impractical to put $N>2$; it must be decided then to take $N=2$ or $N=1$. The condition necessitating $N=2$ or more is approximated by

$$\frac{\mu}{\lambda} \leq \frac{Q}{K}$$

If $Q=$ ¥ 25 million/day and $K=$ ¥ 0.3 million/day, we shall have $N=2$ for a system with $1/\lambda=$ 40 hr, $\mu=1$ hr.

5. Summary

In the present Report 1, several controversial points to be considered in realizing the mechanization of seat reservation system were discussed; particularly analysis was made of the problems involved with the "real time" processing of information. In the ensuing report MARS-1, the prototype of seat reservation system designed on the results of studies made in Report 1 will be described with reference to its reliability in practical use and other matters.

(Pre. Report No. 69, 1961)