Delay Distributions for Simple Trunk Groups with Recurrent Input and Exponential Service Times†

By LAJOS TAKÁCS

(Manuscript received June 16, 1961)

At a telephone exchange, calls appear before a simple trunk group of m lines in accordance with a recurrent process. If every line is busy, calls are delayed. The call holding times are mutually independent random variables with common exponential distribution. In this paper, methods are given for the determination of the distribution of the delay for a stationary process and various orders of service. Three orders of service are considered: (1) order of arrival, (2) random order, and (3) inverse order of arrival.

I. INTRODUCTION

In the theory of telephone traffic, the following process is of considerable interest. In the time interval $0 \leqq t < \infty$, calls appear before a simple trunk group with m lines at instants $\tau_1, \tau_2, \cdots, \tau_n, \cdots$ where the interarrival times $\tau_{n+1} - \tau_n (n = 1, 2, \cdots)$ are identically distributed, mutually independent, positive random variables with distribution function

$$\mathbf{P}\{\tau_{n+1} - \tau_n \leqq x\} = F(x) \qquad (n = 1, 2, \cdots). \tag{1}$$

We say that the call input is a recurrent process. If an incoming call finds a free line, a connection is realized instantaneously. If every line is busy, the incoming call is delayed and waits for service as long as necessary (no defections). Denote by χ_n the holding time of the nth call. It is supposed that $\{\chi_n\}$ is a sequence of identically distributed, mutually independent, positive random variables with distribution function

$$\mathbf{P}\{\chi_n \leqq x\} = H(x) \qquad (n = 1, 2, \cdots) \tag{2}$$

† An address presented on September 14, 1961, at the Troisième Congrès International de Télétrafic, Paris.

and independent of $\{\tau_n\}$. We shall consider only those systems of service which satisfy the requirements that there is no free line if there are calls waiting and that the same principle of service applies to every call (no priorities). Such a service system can be characterized by the symbol $[F(x),H(x),m]$ provided that the order of service is specified.

The ideal order of service, "*order of arrival*" or "*first come — first served*," is not always realizable, particularly at times of heavy traffic; therefore it is important to consider other orders of service also. One of these is "*service in random order*" which often describes the practical situation with high accuracy. In this case, waiting calls are chosen for service at random. Every call, independently of the others, and of its past delay, has the same probability of being chosen. Further, it is of great informative value to consider the extreme case, "*inverse order of arrival*," or "*last come — first served.*" (At present we are not concerned with "priority systems" in which "last come — first served" service is the natural order, e.g., the last information to be received may be the most important in the process of the arrival of messages.)

In what follows we shall consider the system $[F(x),H(x),m]$ in the particular case when call holding times have the exponential distribution

$$H(x) = \begin{cases} 1 - e^{-\mu x} & (x \geqq 0) \\ 0 & (x < 0) \end{cases} \tag{3}$$

and the process is stationary. We shall give methods for finding delay distributions for the three service orders mentioned.

We introduce the notation

$$\varphi(s) = \int_0^\infty e^{-sx}\, dF(x) \tag{4}$$

for the Laplace-Stieltjes transform of the distribution function of interarrival times,

$$\alpha = \int_0^\infty x\, dF(x) \tag{5}$$

for the average interarrival time, $W(x)$ for the delay distribution function, i.e., $W(x)$ is the probability that the delay is $\leqq x$, and

$$\Omega(s) = \int_0^\infty e^{-sx}\, dW(x) \tag{6}$$

for the Laplace-Stieltjes transform of $W(x)$.

If $\Re(s) \geqq 0$ then denote by $z = \gamma(s)$ the root with the smallest absolute value of the equation

$$z = \varphi(s + m\mu(1 - z)).$$

We have $|\gamma(s)| \leqq 1$ and for $r = 1, 2, \cdots$

$$[\gamma(s)]^r = r \sum_{n=r}^{\infty} \frac{(m\mu)^{n-r}}{n(n-r)!} \int_0^\infty e^{-(s+m\mu)x} x^{n-r} \, dF_n(x) \tag{7}$$

where $F_n(x)$ denotes the nth iterated convolution of $F(x)$ with itself. Let $\omega = \gamma(0)$; then

$$\omega = \sum_{n=1}^{\infty} \frac{(m\mu)^{n-1}}{n!} \int_0^\infty e^{-m\mu x} x^{n-1} \, dF_n(x). \tag{8}$$

If $m\alpha\mu \leqq 1$, then $\omega = 1$ while if $m\alpha\mu > 1$, then ω is real and $0 < \omega < 1$.

II. GENERAL THEORY

A. K. Erlang[1] was the first to consider the process $[F(x), H(x), m]$ in the particular case $F(x) = 1 - e^{-\lambda x} (x \geqq 0)$, $H(x) = 1 - e^{-\mu x}$ $(x \geqq 0)$. The case of general $F(x)$ has been treated earlier by D. G. Kendall,[2] F. Pollaczek,[3] and the author.[4]

Denote by $\xi(t)$ the number of calls in the system at the instant t; i.e., $\xi(t)$ is the total number of calls either waiting or being served. Denote by $\chi(t)$ the time difference between t and the arrival of the next call after t. Let $\xi_n = \xi(\tau_n - 0)$, i.e., the nth call finds ξ_n calls in the system, and denote by η_n the delay of the nth call. The initial state is given by $\xi(0)$ and $\chi(0)$.

The vector process $\{\xi(t), \chi(t); 0 \leqq t < \infty\}$ is a Markov process and has the same stochastic behavior for each order of service provided that there is no free line if there are calls waiting. In Ref. 4 it is proved that if $m\alpha\mu > 1$, then there exists a unique stationary process. By choosing the suitable distribution for $\{\xi(0), \chi(0)\}$ we arrive at the stationary process. For the stationary process, $\{\xi(t), \chi(t)\}$ has the same distribution for all t, and the distribution of ξ_n is independent of n. Let $\mathbf{P}\{\xi_n = k\} = P_k (k = 0, 1, \cdots)$. As shown in Ref. 4

$$P_k = \begin{cases} \sum_{r=k}^{m-1} (-1)^{r-k} \binom{r}{k} U_r & (k = 0, 1, \cdots, m-1) \\ A\omega^{k-m} & (k = m, m+1, \cdots) \end{cases} \tag{9}$$

where

$$U_r = AC_r \sum_{j=r+1}^{m} \binom{m}{j} \frac{[m(1 - \varphi(j\mu)) - j]}{C_j[1 - \varphi(j\mu)][m(1 - \omega) - j]} \tag{10}$$

$$A = \left\{\frac{1}{1 - w} + \sum_{j=1}^{m} \binom{m}{j} \frac{[m(1 - \varphi(j\mu)) - j]}{C_j[1 - \varphi(j\mu)][m(1 - \omega) - j]}\right\}^{-1} \tag{11}$$

$$C_j = \prod_{i=1}^{j} \left(\frac{\varphi(i\mu)}{1 - \varphi(i\mu)}\right) \tag{12}$$

and ω is defined by (8).

Remark 1 — In the particular case when $\{\tau_n\}$ is a Poisson process of density λ, i.e., $F(x) = 1 - e^{-\lambda x}$ if $x \geqq 0$, we have $\varphi(s) = \lambda/(\lambda + s)$ and thus

$$\omega = \frac{\lambda}{m\mu}$$

and

$$A = \frac{\frac{1}{m!}\left(\frac{\lambda}{\mu}\right)^m}{\sum_{j=0}^{m-1} \frac{1}{j!}\left(\frac{\lambda}{\mu}\right)^j + \frac{1}{m!}\left(\frac{\lambda}{\mu}\right)^m \left(1 - \frac{\lambda}{m\mu}\right)^{-1}}.$$

Also for the stationary process the distribution function of η_n is independent of n but depends on the order of service. We shall use the notation $\mathbf{P}\{\eta_n \leqq x\} = W(x)$ for all cases. It is to be noted that the expectation $\mathbf{E}\{\eta_n\}$ is independent of the order of service if the same principle of service applies to every call. We shall see later that in each case, the mean waiting time is given by

$$\int_0^\infty x\, dW(x) = \frac{A}{m\mu(1 - \omega)^2}. \tag{13}$$

III. SERVICE IN ORDER OF ARRIVAL

The following theorem has been proved earlier by D. G. Kendall,[2] F. Pollaczek[3] and the author.[4]

Theorem 1 — The delay distribution function for order of arrival service is given by

$$W(x) = 1 - \frac{A}{1 - \omega} e^{-m\mu(1-\omega)x} \qquad (x \geqq 0) \tag{14}$$

where ω is defined by (8) *and A by* (11).

Proof — We have for $x \geqq 0$ that

$$W(x) = \sum_{k=0}^{m-1} P_k + \sum_{k=0}^{\infty} P_{m+k} \int_0^x e^{-m\mu u} \frac{(m\mu u)^k}{k!} m\mu \, du \tag{15}$$

where $\{P_k\}$ is defined by (9). For, if an arriving call finds a free line, which has probability

$$\sum_{k=0}^{m-1} P_k = 1 - \sum_{k=0}^{\infty} P_{m+k} = 1 - \frac{A}{1 - \omega}$$

then its service starts without delay; if it finds every line busy and k ($k = 0, 1, \cdots$) calls waiting, which has probability $P_{m+k} = A\omega^k$, then its service starts at the $(k + 1)$st departure after the arrival. Since the departures follow a Poisson process of density $m\mu$, under this condition, the probability that the delay $\leqq x$ is

$$\int_0^x e^{-m\mu u} \frac{(m\mu u)^k}{k!} m\mu \, du.$$

Thus (15) follows from these and (14) agrees with (15).

It follows from (14) that the average delay is

$$\int_0^{\infty} x \, dW(x) = \frac{A}{m\mu(1 - \omega)^2} \tag{16}$$

and the second moment of the delay is

$$\int_0^{\infty} x^2 \, dW(x) = \frac{2A}{(m\mu)^2(1 - \omega)^3}. \tag{17}$$

IV. SERVICE IN RANDOM ORDER

In the particular case of Poisson input this process has been investiated by S. D. Mellor,[5] E. Vaulot,[6] C. Palm,[7,8]† F. Pollaczek,[9] J. Riordan,[10] R. I. Wilkinson,[11] and J. LeRoy.[12] Now we shall consider the case of recurrent input.

Denote by $W_j(x)$ ($j = 0, 1, \cdots$) the probability that the delay of a call is $\leqq x$, given that on its arrival all lines are busy and j other calls are waiting. For the stationary process the probability that an arriving call finds all m lines busy and j calls waiting is $P_{j+m} = A\omega^j$. If there is a free line when a call arrives, which has probability

$$\sum_{j=0}^{m-1} P_j = 1 - \frac{A}{1 - \omega},$$

then there is no delay.

† Ref. 8 is an English version of the material in Ref. 7.

Thus we have

$$W(x) = \sum_{j=0}^{m-1} P_j + \sum_{j=0}^{\infty} P_{m+j} W_j(x), \tag{18}$$

that is,

$$W(x) = 1 - \frac{A}{1 - \omega} + A \sum_{j=0}^{\infty} W_j(x)\omega^j. \tag{19}$$

Let us introduce the notation

$$\Omega_j(s) = \int_0^{\infty} e^{-sx}\, dW_j(x) \tag{20}$$

and

$$\Phi(s,z) = \sum_{j=0}^{\infty} \Omega_j(s) z^j \tag{21}$$

which is convergent if $\Re(s) \geqq 0$ and $|z| < 1$.

Theorem 2 — The Laplace-Stieltjes transform of the delay distribution function for random service is given by

$$\Omega(s) = 1 - \frac{A}{1 - \omega} + A\Phi(s,\omega) \tag{22}$$

where $\omega = \gamma(0)$ *is defined by* (8), A *by* (11), *and*

$$\begin{aligned} \Phi(s,z) = {} & \frac{m\mu}{s + m\mu[1 - \gamma(s)]} \exp\left\{\int_{\gamma(s)}^{z} \frac{du}{\varphi(s + m\mu(1 - u)) - u}\right\} \\ & + m\mu \int_{\gamma(s)}^{z} \frac{[1 - \varphi(s + m\mu(1 - u))]}{(1 - u)(s + m\mu(1 - u))[u - \varphi(s + m\mu(1 - u))]} \\ & \cdot \exp\left\{\int_{u}^{z} \frac{dv}{\varphi(s + m\mu(1 - v)) - v}\right\} du \end{aligned} \tag{23}$$

where $\gamma(s)$ *is the only root in* z *of the equation*

$$z = \varphi(s + m\mu(1 - z)) \tag{24}$$

in the unit circle $|z| < 1$. *The explicit form of* $\gamma(s)$ *is given by* (7) *with* $r = 1$.

Proof — For $j = 0, 1, \cdots$ we can write that

$$\begin{aligned} W_j(x) = {} & \sum_{k=0}^{j} \frac{(j + 1 - k)}{(j + 1)} \left[\int_0^x e^{-m\mu u} \frac{(m\mu u)^k}{k!}\, dF(u)\right] * W_{j+1-k}(x) \\ & + \sum_{k=0}^{j} \frac{1}{j + 1} \int_0^x e^{-m\mu u} \frac{(m\mu u)^k}{k!} [1 - F(u)] m\mu\, du \end{aligned} \tag{25}$$

where $*$ denotes convolution. If an arriving call finds every line busy and j ($j = 0, 1, \cdots$) calls waiting, then the event that the delay is $\leqq x$ can occur in the following mutually exclusive ways: either in the subsequent interarrival interval k ($k = 0, 1, \cdots, j$) services terminate and the service of the given call does not start during this time interval or, in the subsequent interarrival interval at least $k + 1$ ($k = 0, 1, \cdots, j$) services end and the service of the given call starts at the termination point of the $(k + 1)$st service.

Forming the Laplace-Stieltjes transform of (25) we get

$$(j+1)\Omega_j(s) = \sum_{k=0}^{j} (j+1-k)\Omega_{j+1-k}(s) \int_0^\infty e^{-(s+m\mu)x} \frac{(m\mu x)^k}{k!} \, dF(x) + \sum_{k=0}^{j} \int_0^\infty e^{-(s+m\mu)x} \frac{(m\mu x)^k}{k!} [1 - F(x)] m\mu \, dx \tag{26}$$

whence

$$z \sum_{j=0}^{\infty} (j+1)\Omega_j(s) z^j = \varphi(s + m\mu(1-z)) \sum_{j=0}^{\infty} j\Omega_j(s) z^j + \frac{m\mu z}{(1-z)} \frac{[1 - \varphi(s + m\mu(1-z))]}{(s + m\mu(1-z))},$$

that is,

$$[z - \varphi(s + m\mu(1-z))] \frac{\partial \Phi(s,z)}{\partial z} + \Phi(s,z) = \frac{m\mu}{(1-z)} \frac{[1 - \varphi(s + m\mu(1-z))]}{(s + m\mu(1-z))}. \tag{27}$$

If $m\mu\alpha > 1$, then $| \varphi(s + m\mu(1-z)) | \leqq \varphi(m\mu\epsilon) < 1 - \epsilon$ when $| z | = 1 - \epsilon$ and ϵ is a sufficiently small positive number. Consequently by Rouché's theorem it follows that

$$z = \varphi(s + m\mu(1-z))$$

has one and only one root $z = \gamma(s)$ in the circle $| z | < 1 - \epsilon$, where ϵ is a sufficiently small positive number. The explicit form (7) for $[\gamma(s)]^r$ can be obtained by Lagrange expansion. By definition $\Phi(s,z)$ is a regular function of z if $| z | < 1$ and $\Re(s) \geqq 0$. If we put $z = \gamma(s)$ in (27), then we get

$$\Phi(s,\gamma(s)) = \frac{m\mu}{s + m\mu[1 - \gamma(s)]}. \tag{28}$$

The solution of the differential equation (27) which satisfies (28) can be written in the form (23). Finally, (22) follows from (19).

Remark 2 — Let us introduce the notation

$$A(s,z) = z - \varphi(s + m\mu(1 - z)) \tag{29}$$

and

$$B(s,z) = \frac{m\mu}{(1 - z)} \frac{[1 - \varphi(s + m\mu(1 - z))]}{(s + m\mu(1 - z))}. \tag{30}$$

Then (23) can be written in the following equivalent form

$$\Phi(s,z) = B(s,z) - \int_{\gamma(s)}^{z} \exp\left\{-\int_{u}^{z} \frac{dv}{A(s,v)}\right\} \frac{\partial B(s,u)}{\partial u}\, du. \tag{31}$$

The function $\Phi(s,z)$ can also be expressed in the form of an infinite series as follows

$$\Phi(s,z) = \sum_{j=0}^{\infty} \Phi_j(s)[z - \gamma(s)]^j \tag{32}$$

which is convergent if $| z - \gamma(s) |$ is small enough. If

$$A(s,z) = \sum_{j=1}^{\infty} A_j(s)[z - \gamma(s)]^j \tag{33}$$

[note that $A(s,\gamma(s)) = 0$ by definition of $\gamma(s)$] and

$$B(s,z) = \sum_{j=0}^{\infty} B_j(s)[z - \gamma(s)]^j \tag{34}$$

then $\Phi_j(s)$ $(j = 0, 1, \cdots)$ can be obtained by the following recurrence formula

$$\sum_{k=0}^{j} k\Phi_k(s)A_{j+1-k}(s) + \Phi_j(s) = B_j(s) \qquad (j = 0, 1, \cdots). \tag{35}$$

This follows from (27). In particular by (35) we obtain

$$\Phi_0(s) = B_0(s), \qquad \Phi_1(s) = \frac{B_1(s)}{1 + A_1(s)},$$

$$\Phi_2(s) = \frac{B_2(s)}{1 + 2A_1(s)} - \frac{B_1(s)A_2(s)}{[1 + A_1(s)][1 + 2A_2(s)]}.$$

Formula (32) can conveniently be used to determine the moments of the distribution function $W(x)$. The rth moment

$$\int_0^{\infty} x^r dW(x)$$

can be calculated by the aid of the derivatives

$$\left(\frac{d^i\Phi_j(s)}{ds^i}\right)_{s=0} \qquad (i + j \leqq r).$$

By using the relation $\gamma(s) = \varphi(s + m\mu(1 - \gamma(s))$ we can write that

$$\gamma(s) = \omega + \frac{s\varphi'(m\mu(1-\omega))}{[1 + m\mu\varphi'(m\mu(1-\omega))]} + \frac{s^2\varphi''(m\mu(1-\omega))}{2[1 + m\mu\varphi'(m\mu(1-\omega))]^3} + \cdots. \tag{36}$$

Now in particular we have $\int_0^\infty x\,dW(x) = \dfrac{A}{m\mu(1-\omega)^2}$ (37)

and
$$\int_0^\infty x^2\,dW(x) = \frac{2A}{(m\mu)^2(1-\omega)^3}\left[\frac{2}{2 + m\mu\varphi'(m\mu(1-\omega))}\right]. \tag{38}$$

V. SERVICE IN INVERSE ORDER OF ARRIVAL

The particular case when $\{\tau_n\}$ is a Poisson process was investigated earlier by E. Vaulot,[13] J. Riordan,[14] and D. M. G. Wishart.[15] The case of recurrent input can be treated in a similar way. As noted by J. Riordan[14] the problem can be reduced to finding the distribution of the length of the busy period for the process of type $[F(x), H(x), 1]$ where $F(x)$ is defined by (1),

$$H(x) = \begin{cases} 1 - e^{-m\mu x} & (x \geqq 0) \\ 0 & (x < 0) \end{cases} \tag{39}$$

and there is only one server. In this case denote by $G(x)$ the probability that the length of the busy period is $\leqq x$. The busy period is defined as the time interval during which the server is continuously busy. Evidently $G(x)$ is independent of the order of service, provided that the server is idle if and only if there is no waiting customer in the system.

If $m\mu\alpha > 1$, then there is a unique stationary process, and for the stationary process $W(x)$ is given by

Theorem 4 — The delay distribution function for last-come, first-served service is

$$W(x) = 1 - \frac{A}{1-\omega} + \frac{A}{1-\omega}G(x) \tag{40}$$

where ω is defined by (8), *A by* (11), *and for $x \geqq 0$*

$$G(x) = m\mu \sum_{n=1}^{\infty} e^{-m\mu x}\frac{(m\mu x)^{n-1}}{n!}\int_0^x [1 - F_n(u)]\,du \tag{41}$$

with $F_n(u)$ the nth *iterated convolution of $F(u)$ with itself.*

Proof — If a call arrives and finds a free line, which has probability $1 - A(1-\omega)^{-1}$, then its service starts without delay; if on its arrival every line is busy, then we can remove all the calls waiting without affecting the distribution of the delay of the call in question. The service of this call starts when the queue size decreases to m for the first time. The waiting time of this call evidently has the same distribution as the length of the busy period for the queueing process of type $[F(x),H(x),1]$ with $H(x) = 1 - e^{-m\mu x}$ $(x \geqq 0)$. For, in both cases the arrivals have identical stochastic law and the departures follow a Poisson process of density $m\mu$. Thus we get (40). In Ref. 16 it is proved that

$$\int_0^\infty e^{-sx}\, dG(x) = \frac{m\mu[1 - \gamma(s)]}{s + m\mu[1 - \gamma(s)]} \tag{42}$$

where $\gamma(s)$ is the root with smallest absolute value in z of the equation

$$z = \varphi(s + m\mu(1 - z)). \tag{43}$$

$\gamma(s)$ is given by (7) with $r = 1$. By Lagrange expansion we find that

$$\int_0^\infty e^{-sx}\, dG(x) = \frac{m\mu}{s + m\mu} + s\sum_{n=1}^{\infty} \frac{(-1)^n(m\mu)^n}{n!} \cdot \frac{d^{n-1}}{ds^{n-1}}\left(\frac{[\varphi(s + m\mu)]^n}{(s + m\mu)^2}\right) \tag{44}$$

whence (41) follows by inversion.

By using the expansion (36) we get from (42) that

$$\int_0^\infty x\, dW(x) = \frac{A}{m\mu(1 - \omega)^2} \tag{45}$$

and

$$\int_0^\infty x^2\, dW(x) = \frac{2A}{(m\mu)^2(1 - \omega)^3[1 + m\mu\varphi'(m\mu(1 - \omega))]}. \tag{46}$$

REFERENCES

1. Erlang, A. K., Post Office Elect. Engrs. Jour., **10,** 1917–18, p. 189.
2. Kendall, D. G., Annals of Math. Statistics, **24,** 1953, p. 338.
3. Pollaczek, F., C. R. Acad. Sci., Paris, **236,** 1953, p. 578.
4. Takács, L., Acta Math. Acad. Sci. Hungar., **8,** 1957, p. 325.
5. Mellor, S. D., Post Office Elect. Engrs. Jour., **35,** 1942, p. 53.
6. Vaulot, E., C. R. Acad. Sci., Paris, **22,** 1946, p. 268.
7. Palm, C., Tek. Meddelanden Från. Kungl. Telegrafstyrelsen, Stockholm, **110,** 1946, p. 70.
8. Palm, C., Tele. No. 1, 1957, p. 68.
9. Pollaczek, F., C. R. Acad. Sci., Paris, **222,** 1946, p. 353.
10. Riordan, J., B.S.T.J., **32,** 1953, p. 100.
11. Wilkinson, R. I., B.S.T.J., **32,** 1953, p. 360.
12. LeRoy, J., Ann. des Télécom., **12,** No. 1–2, 1957, p. 2.
13. Vaulot, E., C. R. Acad. Sci., Paris, **238,** 1954, p. 1188.
14. Riordan, J., B.S.T.J., **40,** 1961, p. 785.
15. Wishart, D. M. G., Operations Research, **8,** 1960, p. 591.
16. Takács, L., Operations Research, **8,** 1960, p. 231.