On the Interaction Between Subscribers and a Telephone System

Arne Myskja
The Norwegian Institute of Technology, Trondheim, Norway

Finn Arve Aagesen
The Norwegian Telecommunications Administration, Kjeller, Norway

ABSTRACT

A teletraffic system and the set of its subscribers may be considered as two distinct subsystem of a complete system including both. The two subsystems, which are of quite different natures, are interconnected across an interface, and this interconnection conveys mutual feedback signals to the two subsystems.

Previous publications have treated these phenomena from a practical as well as from a theoretical point of view. The present paper concentrates on the fundamental statistical parameters of the subscriber subsystem and on alternative models based on these parameters. Failure rates and subscriber persistence are essential quantities, and a study of these is carried out, based on mathematical descriptions, including different effects of influence. The subscriber persistence and repetition intervals are measured for different A-subscriber categories and for different failure causes, by means of computerized data recording equipment. Different mathematical models are tested statistically versus experimental results obtained by observations on real traffic.

1. INTRODUCTION

The interaction between telephone subscribers and the telephone system is a matter of rather complex nature. This is very reasonable, since the feedback from a large and complicated technical system influences the set of calling subscribers, which is a non-technical and non-deterministic subsystem. The behaviour of this subsystem is even influenced by personal habits and psychological factors that are difficult to predict. Nevertheless, experience indicates that it is possible by statistical methods to develop a rather satisfactory description of the complicated matter.

A. Elldin [1] carried out a comprehensive discussion of the traffic process related to essential system parts, and in this discussion included the feedback mechanism as well as subscriber dependent properties, such as call urgency and subjective interpretation of the state of the system. However, the influence of these properties was not quantified, and the greater part of Elldin's paper is devoted to a theory of repeated call attempts based on the equation of states method, assuming Markov processes.

More recent investigations have added to the real knowledge of subscriber behaviour, and even made possible simplified theoretical treatment. Some papers in this field are [2]-[8]. Particularly mentioned is Pellieux's very simple model of the repetition rate as a function of the overall failure probability.

The approach of the present paper contains three main elements:
- a more detailed study of the main parameters in Elldin's fundamental theory
- observation results from automatic measurements on a group of subscribers
- statistical tests of some models compared with experimental results.

GENERAL SYMBOL LIST

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f, p )</td>
<td>failure rate, probability of failure</td>
</tr>
<tr>
<td>( F )</td>
<td>accumulated failure rate</td>
</tr>
<tr>
<td>( w )</td>
<td>persistence, probability of repetition</td>
</tr>
<tr>
<td>( \bar{w} )</td>
<td>accumulated persistence</td>
</tr>
<tr>
<td>( x )</td>
<td>attempt number, index</td>
</tr>
<tr>
<td>( r )</td>
<td>efficiency rate, ( r = 1 - f )</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>call multiplication factor</td>
</tr>
<tr>
<td>( \beta )</td>
<td>functional constant</td>
</tr>
<tr>
<td>( p )</td>
<td>probability</td>
</tr>
<tr>
<td>( A )</td>
<td>call rate, general and reattempt rate</td>
</tr>
<tr>
<td>( c )</td>
<td>call rate towards individual B-subscribers</td>
</tr>
<tr>
<td>( \bar{c} )</td>
<td>mean failure state duration</td>
</tr>
<tr>
<td>( Y )</td>
<td>parameter of failure state distribution</td>
</tr>
<tr>
<td>( z, \gamma )</td>
<td>parameter of free state distribution</td>
</tr>
<tr>
<td>( \vartheta, \delta, \varphi )</td>
<td>distribution functions</td>
</tr>
<tr>
<td>( \mu, \rho )</td>
<td>statistical moments</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>standard deviation</td>
</tr>
<tr>
<td>( E(y) )</td>
<td>expectation of ( y )</td>
</tr>
<tr>
<td>( k, m )</td>
<td>parameters in Beta-distribution</td>
</tr>
<tr>
<td>( r, q )</td>
<td>general variables</td>
</tr>
</tbody>
</table>

2. FUNDAMENTAL PARAMETERS AND MODELS

In Elldin's paper a simplified theoretical model of repeated calls is presented, based on two fundamental parameters:

1. \( f \) = failure probability for an arbitrary call.

(2.1)\( F_x = \frac{P_x}{1 - f} \), \( F_0 = 1 \)

Accumulated failure probability:

(2.2)\( W_x = \frac{P_x}{1 - w} \), \( W_0 = 1 \)

Accumulated retrial probability:

(2.3)\( P_x = F_x - 1 - (1 - e^{-\lambda x}) = (F_x - 1 - P_x)W_x \)

Probability of success after exactly \( x \) trials:

(2.4)\( Q_x = F_xW_x - 1 - (1 - W_x) \)

Probability of termination after exactly \( x \) trials:

(2.5)\( R_x = P_x + Q_x = F_x - 1 - W_x \)

Probability of abandonment after exactly \( x \) trials:
Probability of a call series ending successfully:

\[ P = \sum_{x=1}^{\infty} P_x = \sum_{x=1}^{\infty} (F_{x=1} - F_x)W_{x=1} \]  
(2.6)

Probability of a call series being abandoned:

\[ Q = \sum_{x=1}^{\infty} Q_x = \sum_{x=1}^{\infty} F_x(W_{x=1} - W_x) = 1 - P \]  
(2.7)

Expected number of trials, or average length of a call string:

\[ E(x) = \sum_{x=1}^{\infty} xR_x \]  
(2.8)

Relating these quantities to key concepts in LeGall-Pellieux's theory yields

\[ E(x) = \frac{8r}{P} = \frac{8r}{Q} = \frac{8r}{1 - Q} \]  
(2.9)

\[ EW = \frac{r}{P} \]  
(2.10)

A further discussion will be carried out in section 3.

While \( f \) and \( w \) are the fundamental parameters, the deduced parameters \( E(x) \) and \( P/E(x) \) are also useful traffic quantities with meaningful interpretations, and so are the combined quantities \( P=8r \) and \( Q=1-8r \).

In the following we shall carry out a study of the variation of \( f_x \) and \( w_x \).

### 2.1 DISCUSSION OF PARAMETERS

Even in a stationary traffic situation there are two reasons for variation of the failure rate \( f_x \) with the attempt number \( x \):

1. \( f \) is the conditional probability that an attempt fails, given that the previous \( x-1 \) attempts failed.
2. There will be a selection among destinations with increasing \( x \), which may lead to a variation in \( f \). The reason for this is that the destinations with a low failure rate will be reached successfully during the early attempts, whereas the remaining ones will have a higher failure rate.

A similar selection effect also applies to the persistence \( w_x \).

The cause of failure may belong to one of four main categories:

- Failure on the part of the calling subscriber
- System congestion, delay or fault
- Busy called subscriber
- Absent called subscriber

These categories are of different character, and they are listed in ascending order of state duration. (This does not quite apply to system faults, however, any special effects are considered negligible).

For the sake of simplicity, only one category of failure cause will be considered at a time. Overall effects can generally be obtained by linear combinations, since observations have shown that most call strings have uniform failure causes.

2.1.1. FAILURE INTERDEPENDENCE BETWEEN CALL ATTEMPTS

For the study of conditional failure probabilities, three distribution functions are of interest, namely that of

- Failure state duration, index \( f \)
- Inter-failure interval duration, index \( i \)
- Reretempt interval, index \( r \)

Frequency functions are denoted \( \phi(t) \), cumulative distribution functions \( \Psi(t) \) and survival functions \( \psi(t) \):

\[ \phi(t) = \int_0^t \phi(z)dz = 1 - \Psi(t) \]  
(2.12)

Figure 1.

With reference to fig. 1, the probability of hitting the same failure state in a second attempt, given that the first attempt failed, can be calculated by

\[ p_1 = 1 - \sum_{T_1=0}^{\infty} \int \Psi_f(t_1)\psi_f(t_1)dt_1 \]  
(2.13)

The probability of hitting the first interfailure interval in the second attempt is:

\[ p_2 = \sum_{T_1=0}^{\infty} \int \int \psi_f(t_1)\psi_f(t_1)\psi_f(t_2)dt_1dt_2 \]  
(2.14)

In general, failure states \( f \) and interfailure states \( i \) are interleaved, and the probability that the second attempt hits a state number \( j \) may be expressed by

\[ p_j = \sum_{T_1=0}^{\infty} \left( \begin{array}{c} p_{j-1} - p_j \\ j \\ \end{array} \right) \]  
(2.15)

where

\[ p_j = \sum_{T_1=0}^{\infty} \int \psi_f(t_1)i_p_i+j(2j-1)dt_1 \]  
(2.16)

Here

\[ \phi_j = \phi_j \text{ for } j \text{ odd} \]  

\[ \psi_j = \psi_j \text{ for } j \text{ even} \]

If the reattempt interval distribution can be expressed as a sum of exponential terms,

\[ \psi_f(z) = a_1e^{-\lambda z} + a_2e^{-\lambda_2 z} + \ldots + a_ne^{-\lambda_n z} \]  
(2.17)

then the solution of (2.16) can be expressed by the Laplace-transforms \( \phi_f^* (\lambda_k) \) with respect to \( \lambda_k \)

Each term number \( k \) is of the form

\[ \lambda_k = \frac{1}{2} \left( \begin{array}{c} -1 \n \kappa \end{array} \right) \phi_f^*(\lambda_k) \cdot \phi_i^*(\lambda_k) \]  
(2.18)

Till now failure states and interfailure states are assumed to have general distributions. Assuming that

\[ \phi_f(z) = ye^{-yz} \]  
(2.19)

the solution (2.18) takes the form

\[ I_{j,k} = a_k^\gamma \left( \frac{1}{2} \left( \begin{array}{c} -1 \n \kappa \end{array} \right) \phi_f^*(\lambda_k) \cdot \phi_i^*(\lambda_k) \right) \]  
(2.19)
From (2.15), since
\[ \frac{1}{\sum_{j=1}^{m-1}} = \gamma \]
and
\[ \sum_{j=1}^{m} a_k (\lambda_k + \delta) \gamma \](2.20)

From formula (2.18) one can observe that the solution will still be rather simple if (2.19) is replaced by Erlang m-distributions, whereas hyperexponential distributions of the form (2.17) will lead to more complicated sums of product terms.

It is easily verified that
\[ \sum_{j=1}^{m} p_j = 1 \]

The probability of failure in a second attempt, given failure in the first attempt is:
\[ f_2 = \frac{\sum_{m=1}^{\infty} \sum_{k=1}^{n} a_k (\lambda_k + \delta) \gamma \left( \sum_{j=1}^{m-1} \right)}{\sum_{m=1}^{\infty} \sum_{k=1}^{n} a_k (\lambda_k + \delta) \gamma \left( \sum_{j=1}^{m-1} \right)} \]

From formula (2.18) one can observe that the solution will still be rather simple if (2.19) is replaced by Erlang m-distributions, whereas hyperexponential distributions of the form (2.17) will lead to more complicated sums of product terms.

In practice the reattempt interval distribution can be approximated by a two-term function
\[ f_2 = b \frac{1}{1+\gamma} + (1-b) e^{\gamma} \]

so that (2.22) takes the form
\[ f_2 = b \frac{1}{1+\gamma} + (1-b) e^{\gamma} \]

(2.23)

Any first attempt is considered to be random, with a failure probability
\[ f_1 = \frac{1}{\gamma} + \frac{\delta}{1+\gamma} = \frac{1}{\gamma + \delta} \]

With constant parameters one always obtains
\[ f_2 > f_1 \]

Under the assumption of \( f_2(z) = \gamma c e^{-\gamma z} \) the probability of failure in a third attempt will have a form which is identical to (2.23) even though the parameters may be different.

The argument may be repeated, so that for an attempt number \( k \) the probability of failure is
\[ f_k = b_k \frac{1}{1+\gamma} + (1-b_k) e^{\gamma} \]

where \( b_k \) means a parameter relating to an \( x \)-th attempt.

2.1.2 THE FAILURE RATE SELECTION EFFECT

The effect treated in the previous section will now be neglected for assuming by instance that \( \gamma > \lambda_1, \lambda_2 \)

Furthermore is assumed a distribution \( g(z) \) of called subscribers according to traffic load \( z \) in the interval \( 0 \leq z \leq 1 \). Similar distribution may be envisaged for other failure causes than busy subscriber.

Assuming now that the persistence \( \omega_x \) is independent of \( z \) and that the fresh call rate per called subscriber \( c(z) \) is in general a function of \( z \), the following expressions are obtained:

\[ c_x = \frac{1}{\gamma} \frac{1}{z} \]

\[ b_x = \frac{1}{\gamma} \frac{1}{z} \]

\[ \rho_x = \frac{1}{\gamma} \frac{1}{z} \]

(2.28)

(2.29)

(2.30)

where \( c_x \) is call frequency, \( b_x \) is failure frequency and \( \rho_x \) is failure probability of \( x \)-th attempt.

Assuming now that \( c(z) \) can be modeled by
\[ c(z) = \frac{c}{z} \]

or
\[ c(z) = c' z \]

(2.31)

(2.32)

the expressions (2.23) and (2.24) are obtained, respectively, where \( \mu_k \) is the \( x \)-th moment of \( g(z) \):

\[ \mu_k = \frac{\mu_x}{z} \]

(2.33)

\[ \mu_x = \frac{\mu_1}{z} + \frac{\mu_0}{z} \]

(2.34)

For fresh calls then
\[ \mu_0 = \frac{\mu_1}{z} + \frac{\mu_0}{z} \]

(2.35)

\[ \mu_1 = \frac{\mu_2}{z} + \frac{\mu_1}{z} + \frac{\mu_2}{z^2} \]

(2.36)

\[ \mu_x = \frac{\mu_x}{z} \]

(2.37)

Given \( k \) and \( m \), and equating the fresh call failure rate from (2.24) with (2.26) and (2.27) respectively, leads to

\[ \frac{1}{\mu_x} = \frac{1}{\mu_0} + \frac{1}{\mu_1} \]

(2.38)

\[ \frac{1}{\mu_x} = \frac{1}{\mu_0} + \frac{1}{\mu_1} \]

(2.39)

For a discussion of formula (2.27) the parameters \( b, \lambda_1, \lambda_2 \) and \( \gamma \) are not substantially influenced by the selection effect, particularly when busy subscribers are concerned. For other failure causes this is not quite so evident. The parameter \( \delta \), however, may increase substantially with increasing \( z \).

2.1.3. PERSISTENCE \( \omega_x \): THE SELECTION EFFECT

For the persistence \( \omega_x \) there is no interdependence between the first call and the following ones similar to that of the failure probability. However, the same kind of selection effect exists.

Suppose that a certain degree of urgency is connected with each traffic demand i.e. call string. The urgency expresses itself by the probability of a reattempt after a failure. If all call strings are allocated along the interval \( 0 \leq z \leq 1 \) according to their reattempt probability \( z \), a corresponding probability distribution \( h(z) \) is obtained. For a given call string, \( z \) is assumed not to change with \( x \). It will furthermore be assumed that the failure proba-
bility is independent of the urgency. The persistence \( w_x \) can now be calculated:

\[
 w_x = \int_0^\infty z h(z) \, dz = \frac{\eta_x}{\eta_x - 1}
\]

(2.40)

where \( \eta_x \) is the \( x \)-th moment of \( h(z) \).

Introducing again the Beta distribution (2.35) one obtains

\[
 w_x = \frac{x+k-1}{x+k+m-1}
\]

(2.41)

2.1.4. CALCULATED RESULTS

In figures 2 and 3 are shown results for \( f_x \) according to equations (2.24), (2.27), (2.38) and (2.39).

The parameters in the restatement distribution are calculated according to an adaptation to a two-term hyperexponential distribution for residence-subscribers when "busy B-subscriber"; \( t_m = 1/\gamma \) is chosen to be \( \gamma = 0.084 \).

Fig. 2 illustrates limiting cases and a typical case for the exponential distribution for residence-subscribers when the selection effect is included, with limiting and typical curves are valid for all values of \( \Delta R \leq 1 \). The same applies to the persistence \( w_x \), as also illustrated.

In fig. 3 the selection effect as well the interdependence effect is included, with limiting and typical curves for the cases \( c(z) = c \) and \( c(z) = c' + z \), equations (2.31) and (2.32).

![Figure 2: Failure rate. Only selection effect included. Persistence: \( b=0.95, \lambda_1=6.87 \cdot 10^{-7} \text{s}^{-1}, \lambda_2=2.39 \cdot 10^{-4} \text{s}^{-1}, t_m = 180 \text{s.} \)](image)

![Figure 3: Failure rate. Selection effect and interdependence effect included. Persistence: \( b=0.95, \lambda_1=6.87 \cdot 10^{-7} \text{s}^{-1}, \lambda_2=2.39 \cdot 10^{-4} \text{s}^{-1}, t_m = 180 \text{s.} \)](image)

2.2. OBSERVATION RESULTS

The observations comprise measurements on real telephone traffic done by means of the recording equipment "CARAT" [10] (Computer Aided Recording and Analysis of Traffic). The measurements cover a group of 500 subscribers, which has been subdivided into 3 categories, A, B and C:

- A: PABX lines
- B: Business subscribers
- C: Residence subscribers
- D: Total group

Because of signalling limitations only local traffic is considered. The measurements comprise 93 working days between hours 8.00 and 16.00 in the period 29/5-74 to 21/8-75. Only calls with completed dialling are considered. The result of a call attempt may then be as follows:

i=1: The A-subscriber replaces his handset immediately after dialling.

i=2: The A-subscriber gives up waiting for a free path.

i=3: Time-out occurs while waiting for a free path.

i=4: B-subscriber busy.

i=5: No B-answer.

i=6: Successfull call.

Only the first 4 calls in a call string are considered. The symbol \( f_x, w, \) and \( \delta \) are used as previously, and a double index \( x_i \) means an \( x \)-th attempt with a result \( i \).

Series with mixed results are neglected. In addition are introduced

\[ n_{xi} \quad \text{number of attempts no.} \cdot x \text{ with result } i \]

\[ n_{yi} \quad \text{number of one day samples for some variable } \xi \]

\[ \mu_x \quad \text{a-percentage point in the standard normal distribution} \]

\[ E \quad \text{expectation} \]

\[ \text{Var} \quad \text{variance} \]

\[ \nu \quad \text{coefficient of variance} \]

\[ s \quad \text{skew} \]

Let \( z \) be a stochastic variable. Then

\[ v(z) = \frac{E(z-E(z))^2}{(E(z))^2} \]

and

\[ s(z) = \frac{E(z-E(z))^3}{(E(z))^2} \]

(2.43)

\[ \text{The symbol } "v" \text{ at the top of the operators symbolize estimators. Table 1 shows the number of calls that is observed from the different A-subscriber categories.} \]

<table>
<thead>
<tr>
<th>( x )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 1 )</td>
<td>340</td>
<td>505</td>
<td>535</td>
<td>6772</td>
<td>5707</td>
<td>64244</td>
</tr>
<tr>
<td>( 2 )</td>
<td>2</td>
<td>28</td>
<td>42</td>
<td>2365</td>
<td>990</td>
<td>D</td>
</tr>
<tr>
<td>( 3 )</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>324</td>
<td>160</td>
<td>-</td>
</tr>
<tr>
<td>( 4 )</td>
<td>159</td>
<td>151</td>
<td>253</td>
<td>3724</td>
<td>2892</td>
<td>A</td>
</tr>
<tr>
<td>( 5 )</td>
<td>0</td>
<td>12</td>
<td>34</td>
<td>1296</td>
<td>491</td>
<td>-</td>
</tr>
<tr>
<td>( 6 )</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>552</td>
<td>160</td>
<td>-</td>
</tr>
<tr>
<td>( 7 )</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>55</td>
<td>160</td>
<td>-</td>
</tr>
<tr>
<td>( \text{Total} )</td>
<td>558</td>
<td>717</td>
<td>879</td>
<td>1113</td>
<td>1318</td>
<td>10632</td>
</tr>
</tbody>
</table>

Table 1. Number of calls \( n_{xi} \).

\( x \) = repetition number

\( i \) = result of call attempt

2.2.1. PERSISTENCE \( w_{xi} \)

Table 2 shows the estimated mean and standard deviation of \( w_{xi} \) for the different A-subscriber categories. In spite of the fairly large sample of measurements from 500 subscribers in 93 days, the uncertainty in the material is considerable.

To get an estimate of the significance in the observed difference in \( w_{xi} \) with different causes of failure and different A-subscriber categories, suppose that \( w_{xi} \) is normally distributed \( N(\mu_{xi},\sigma_{xi}^2) \). To test the hypothesis \( H_0 \) against \( H_1 \),

\( H_0 : w_{xi} = w_{ij} \)

\( H_1 : w_{xi} \neq w_{ij} \)

ITC8
suppose that
\[ N = \frac{\hat{E}(w_{1j}) - \hat{E}(w_{1j})}{\sqrt{\text{var}(w_{1j})} + \sqrt{\text{var}(w_{1j})}}^{1/2} \] (2.44)

if \( H_0 \) is normally distributed \( N(0,1) \). \( H_0 \) is rejected with significance level \( \alpha \) when \( |N| > z_{\alpha/2} \), where \( z_{\alpha/2} \) is the \( \alpha/2 \)-percentage point in the standard normal distribution. With significance level \( \alpha = 0.05 \), \( z_{\alpha/2} = 1.96 \). There is no significant difference between failure causes 2 and 3.

The differences between the other causes are significant.

The subscriber behaviour is of significant interest, how­ever because it must be expected that these quantities will become considerable in situations with high traffic loads.

Failure cause 1 has no practical importance. Failure causes 2 and 3 are also of little practical importance in this observation. It is clear that failure cause 1 has no practical importance. Failure causes 2 and 3 are also of little practical importance in this observation. The subscriber behaviour is of significant interest, however because it must be expected that these quantities will become considerable in situations with high traffic loads.

Table 2. Persistence

<table>
<thead>
<tr>
<th>i</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.415</td>
<td>.784</td>
<td>.778</td>
<td>.702</td>
<td>.488</td>
<td>PABX (A)</td>
</tr>
<tr>
<td>2</td>
<td>.316</td>
<td>.207</td>
<td>.194</td>
<td>.129</td>
<td>.147</td>
<td>Residence</td>
</tr>
<tr>
<td>3</td>
<td>.91</td>
<td>86</td>
<td>78</td>
<td>93</td>
<td>93</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>.327</td>
<td>.669</td>
<td>.751</td>
<td>.661</td>
<td>.442</td>
<td>Business (B)</td>
</tr>
<tr>
<td>2</td>
<td>.373</td>
<td>.365</td>
<td>.286</td>
<td>.213</td>
<td>.179</td>
<td>93</td>
</tr>
<tr>
<td>3</td>
<td>.72</td>
<td>54</td>
<td>70</td>
<td>93</td>
<td>93</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>.305</td>
<td>.918</td>
<td>.824</td>
<td>.811</td>
<td>.506</td>
<td>Residence (C)</td>
</tr>
<tr>
<td>2</td>
<td>.443</td>
<td>.304</td>
<td>.284</td>
<td>.181</td>
<td>.138</td>
<td>46</td>
</tr>
<tr>
<td>3</td>
<td>.667</td>
<td>.769</td>
<td>.867</td>
<td>.559</td>
<td>.762</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>.397</td>
<td>.228</td>
<td>.386</td>
<td>.222</td>
<td>.369</td>
<td>82</td>
</tr>
</tbody>
</table>

Table 3. Repetition intervals (in sec.)

<table>
<thead>
<tr>
<th>i</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.004</td>
<td>.016</td>
<td>.007</td>
<td>.089</td>
<td>.081</td>
<td>PABX (A)</td>
</tr>
<tr>
<td>2</td>
<td>.014</td>
<td>.071</td>
<td>.104</td>
<td>.506</td>
<td>.387</td>
<td>Business (B)</td>
</tr>
<tr>
<td>3</td>
<td>.004</td>
<td>.007</td>
<td>.020</td>
<td>.125</td>
<td>.096</td>
<td>Residence (C)</td>
</tr>
</tbody>
</table>

2.2.3. FAILURE RATE \( f_x \)

As already seen in section 2.1.1, \( f_x \) is given when \( \Phi_1(x(t)) \), \( \Phi_{1,x}(x,t) \) and \( \Phi_{1,x}(x,t) \) are known. In the following some observation results are presented to illustrate the shown effects. Table 4 contains \( f_x \) and \( f_x \), where \( f_1 \) is the total average for attempts 1-3 with result 1.

It is clear that failure cause 1 has no practical importance. The subscriber behaviour in this situation is therefore of little interest. Failure causes 2 and 3 are also of little practical importance in this observation. The subscriber behaviour is of significant interest, however because it must be expected that these quantities will become considerable in situations with high traffic loads.

Table 4. The probability of failure.
Table 4 indicates always a strong increase in \( f_{xi} \) when \( x \) goes from 1 to 2. This is mainly caused by the interdependence effect. The data available show:

- a small increase in \( f_{xi} \) when \( x \) goes from 2 to 3 for the case of "busy B-subscriber".
- a fairly large increase in \( f_{xi} \) when \( x \) goes from 2 to 3 for the case of "no B-answer".

As observed in section 2.1.1., if \( f_{xi}(t) \) is n.e.d., \( f_{xi} \) will have the same expression for all \( x \geq 2 \). \( f_{xi}(t) \) for the busy condition is nearly n.e.d. Thus a small increase in \( f_{xi} \) due to the selection effect must be expected for \( x \geq 2 \). The observation results are then in accordance with the theory.

The cause of "no B-answer" is due to two situations:

i) The B-subscriber is not present
ii) Long answering-time because of different situations by the B-subscriber. These may be a highly loaded switchboard, short temporary absences etc.

The effect of this combination is a \( f_{xi}(t) \) which is hyper-exponential, or a "flat" distribution. The assumption that leads to \( f_{xi} \rightarrow f_{x} \) does not apply in this case. A detailed calculation is not carried out. However, an intuitive explanation is that situation i) will be more predominant in later attempts with this distribution than with n.e.d., thus leading to an increasing \( f_{xi} \) with \( x \).

This adds to the selection effect as calculated. Even the selection effect tends to influence the later results more in the "no answer" case than in the "busy" case. As can be seen from table 3 the coefficient of variation is small for this case, which actually means that the \((1-b)\)-term in equation (2.27) is greater than in the "busy" case. In practice the selection effect is concentrated to the \((1-b)\)-term.

For PABX-subscribers as well as for business-subscribers "busy B-subscriber" is the most important failure cause, while for residence subscribers "no B-answer" is predominant. In section 2.2.1. and 2.2.2. is seen that residence subscribers have the highest persistence and the shortest time between repetitions for each failure cause. However, an observation that does not take into account the failure cause, might lead to other and wrong conclusions, because the persistence is less and the repetition interval longer for "no B-answer" than for "B-subscriber busy".

### 3 SIMPLIFIED MODELS

#### 3.1 GENERAL

For an exact study of the stochastic properties of the complete arrival process, knowledge of the fundamental parameters and models presented in section 2 is necessary.

However, for less exact purposes approximate models may be sufficient. Suppose there is only one failure cause and only one subscriber category. Equation (2.8) may be written

\[
\beta = 1 - \frac{1}{\sum_{i=1}^{x} f_i w_i^x} \quad (3.1)
\]

and from (2.6) and (2.8)

\[
\omega = \sum_{i=1}^{x} f_i w_i^x \quad (3.2)
\]

Experiments have indicated a clear functional relationship between \( \beta \) and \( \omega \). The simplified models are of the form

\[
\beta \propto \omega \quad (3.3)
\]

The value of the simplified models depends on the possibility of evaluating \( r \) without having to use the more basic quantities \( w_i \) and \( f_i \), as the calculation of \( r \) is of the same complexity as the calculation of \( \beta \).

#### 3.2 TWO SIMPLE PROPOSED MODELS

If in (3.1) and (3.2) \( w_1 = \omega = \text{constant} \), then

\[
\beta = \frac{1}{1-(1-r) \omega} \quad (3.4)
\]

even for arbitrary \( f_i \). Thus (3.4), which will be called Eldén's simplified model, has a wider application than the simple \( f_x = \omega \) and \( \omega x = f = 1-r \).

Another proposal is the LeGall-Pellieux model

\[
\beta = r^a \quad a = \text{constant} \quad (3.5)
\]

The constant \( a \) has no physical interpretation in the deduction of the function, which is based on network arguments rather than on the fundamental parameters. It is possible, though, to interpret (3.5) even in a physical way by Taylor series expansion:

\[
\beta = r - a + \frac{1}{2} (1-r)^2 + \frac{1}{6} r^3 + \ldots \quad (3.6)
\]

Comparison illustrates that (3.6) is identical to (3.1) if one assume

\[
\omega = \frac{x+a-1}{x} \quad (3.7)
\]

\[
f_x = 1 - r \quad (3.8)
\]

Note the functional similarity between (2.41) and (3.7). According to (3.7) \( \alpha = w_1 \), which has a clear physical interpretation.

#### 3.3 EXPERIMENTAL ANALYSIS OF SIMPLIFIED MODELS

##### 3.3.1. STATISTICAL METHOD

This analysis is also based on data from the traffic recording system CARAT [10]. The measurements cover a group of 500 subscribers and comprise 39 working days in the time between 9.00 and 15.00 in the period 26/11-73 to 15/3-74. This analysis is independent of the one presented in section 2.2.

The source may be a PABX or a single line, while the destination is a trunk or a local line. The source-destination combination are:

- \( K \): PABX - trunk line
- \( P \): PABX - local line
- \( G \): Single line - trunk line
- \( N \): Single line - local line

The group contains 86 PABX-lines, 107 of the single lines are business lines, 1 is a coin telephone, 13 special purpose lines like operator trunk lines, 40 are not in use and the remaining 253 are residence lines.

This chapter is an extract of [11], where \( \beta \) is analysed for observation times of 10 and 30 minutes and 1, 3 and 6 hours, while the present results are based on 1 hour observation time.

The following symbols are used:

- \( r_i \) - discrete representation of \( r \). Resolution is 0.01.
- \( \beta_i, j \) - sample number of \( \beta \) when \( r_i, j = 1, 2, \ldots, N \).
- \( N_i \) - number of observations when \( r_i \).
- \( N_a \) - number of observations in a given area of \( r \)-axis.
- \( I \) - number of different points on the \( r \)-axis.
- \( f_{a,b_1,b_2} \) - \( q \)-percentage point in the Fisher-distribution with \( b_1 \) and \( b_2 \) degrees of freedom.

For comparison a third relation between \( \beta \) and \( r \) is supposed.

\[
\beta = \eta + (1-\eta) r \quad (3.9)
\]

Optimal values of \( \eta, \omega \) and \( \beta \) is found by the principle of least squares. Let \( \Theta, \Phi \) and \( \psi \) symbolize the quantities found by this principle.

Assuming that \( \beta_i \) is normally distributed \( N(\mu_i, \sigma^2_i) \) and that the measured values of \( r \) are exact, the hypotheses

\[
H_0^* : \beta = \frac{1}{1-(1-r)\omega} \quad (3.4)
\]
are tested against the alternative that there is another relation between $\beta$ and $r$. By using the maximum likelihood principle, $H_0^{***}$ is rejected when

$$
F^{***} = \frac{\sum_{i=1}^{N_1} \frac{(\hat{\theta}_{ij} - \hat{\theta}_{0}^2)}{(N-1)} \geq f_{\alpha, i-1, N-1}
$$

because $F^{***}$ is $F$-distributed with $i-1$ and $N-1$ degrees of freedom. Let $F'$ and $F''$ denote the analogous expressions for $H_0'$ and $H_0^{**}$, with $1/[(1-r_j)\hat{\theta}_{ij}]$ and $r_j$ respectively substituted in the numerator. If it is supposed that $F'$ and $F''$ are also $F$-distributed when $H_0'$ and $H_0^{**}$ respectively are correct.

### 3.1.2 RESULTS

The mean of the fresh call intensity for the different source-destination combinations is $E$ in the observation period is 50, 180, 31 and 163 calls per hour, respectively.

Table 5 shows the estimated values for the expectation of $\hat{\beta}$ in the different intervals on the r-axis. Table 6 shows the result of the model-testing. The significance level is 0.05. The linear model is rejected for all source-destination-combinations. Elldin's simplified model is rejected for combination G and H, while LeGall/Pellieux's model is rejected for combination G.

It is to remark that the different combinations give different values for the parameters $\omega$, $\alpha$ and $\eta$. The difference between the combinations is found significant in [11], however combination G might be equal to $F$ or $H$. It is also remarkable that the value that Pellieux has obtained in [8] for $\alpha$ is 0.72 while our value is 0.74.

### Table 5. Estimated values of $\hat{\beta}$

<table>
<thead>
<tr>
<th>$r$</th>
<th>$E$</th>
<th>$F$</th>
<th>$G$</th>
<th>$H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.35</td>
<td>1.95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.45</td>
<td>1.82</td>
<td>1.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.55</td>
<td>1.53</td>
<td>1.44</td>
<td>1.39</td>
<td></td>
</tr>
<tr>
<td>0.65</td>
<td>1.38</td>
<td>1.33</td>
<td>1.31</td>
<td>1.29</td>
</tr>
<tr>
<td>0.75</td>
<td>1.26</td>
<td>1.22</td>
<td>1.21</td>
<td>1.19</td>
</tr>
<tr>
<td>0.85</td>
<td>1.16</td>
<td>1.15</td>
<td>1.11</td>
<td>1.12</td>
</tr>
</tbody>
</table>

Table 6. Results of model testing.

<table>
<thead>
<tr>
<th>$h_1, h_2$</th>
<th>43,190</th>
<th>26,207</th>
<th>52,181</th>
<th>26,207</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_{0.05,h_1,h_2}$</td>
<td>1.4</td>
<td>1.5</td>
<td>1.4</td>
<td>1.5</td>
</tr>
<tr>
<td>$\beta = 1/(1-(1-r)\omega)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reject?</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>$F'$</td>
<td>1.2</td>
<td>.7</td>
<td>2.1</td>
<td>1.7</td>
</tr>
<tr>
<td>$\hat{\omega}$</td>
<td>.80</td>
<td>.73</td>
<td>.72</td>
<td>.65</td>
</tr>
<tr>
<td>$\beta = r^{-\alpha}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reject?</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>$F''$</td>
<td>1.0</td>
<td>.6</td>
<td>2.3</td>
<td>1.4</td>
</tr>
<tr>
<td>$\hat{\alpha}$</td>
<td>.74</td>
<td>.69</td>
<td>.62</td>
<td>.60</td>
</tr>
<tr>
<td>$\beta = \eta +(1-n)r$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reject?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>$F'''$</td>
<td>4.1</td>
<td>2.1</td>
<td>4.7</td>
<td>3.2</td>
</tr>
<tr>
<td>$\hat{\eta}$</td>
<td>2.20</td>
<td>1.91</td>
<td>2.09</td>
<td>1.82</td>
</tr>
</tbody>
</table>

The point with the simplified models is that they should be easy to use in practical situations with acceptable accuracy. A comparison between table 5 and the computed values from Elldin's simplified model and LeGall/Pellieux's model with $\hat{\alpha}$ and $\hat{\omega}$ from table 6 gives a maximum difference between the models and the observations of approx, 4 %.

The practical difference between the models seems to be small, and they both seem to be acceptable in practice.

It should be stressed, however, that the presented values of $\alpha$ and $\omega$ are only valid in situations with the same call mix and the same subscriber group composition. As seen in chapter 2.2 the persistence is different for different failure causes and for the different A-subscriber categories. If the single lines were only residence subscribers, and if they had the same call mix as the PABX-lines, the values of $\omega$ and $\alpha$ would be greater for single lines than for PABX-lines.

The previous results are based on the assumption that all hours between 9.00 and 15.00 gives the same $\beta$ when $r$ is fixed. However, this assumption is not quite correct, fig. 4 shows $\hat{\alpha}$ estimated for different hours of the day. $\hat{\alpha}$ shows an increasing tendency because of an accumulation of call demands. For all source-destination combinations one has a maximum between 14.00 and 15.00. The difference between the maximum and the mean values of $\alpha$ is approx 0.05.

![Figure 4](image-url)
4. CONCLUSION

The present study of the interrelationship between a telephone exchange as a technical system and its subscribers as a social environment largely follows the same approach of statistical description as presented in some earlier papers in this field.

Beginning with a probabilistic model with failure and reattempt probabilities as basic quantities, a further study of the main effects influencing these quantities is carried out. This study indicates that the strong increase in failure probability with the attempt number can be explained partly by the conditional probability of failure in the following attempts, given first attempt failure, and partly by a repeated selection among destinations. This selection effect also applies to the reattempt probability. Experimental results largely affirm the model calculations.

Furthermore, a statistical study of simplified models of repeated calls is carried out. The close interrelationship between the two main models is pointed out, and the models are tested statistically versus observation results. The tests indicate that the models may cover considerable variations fairly well, though the type of source-destination combinations as well as the time of day influence the parameters of the models.

5. ACKNOWLEDGEMENT

The authors acknowledge the great value of the cooperation with people within local as well as central branches of the Norwegian Telecommunications Administration. This has made available vital observation data for the present study.

Several people have contributed with ideas or practical assistance. Particularly mentioned is A.0stlie for suggestions of considerable value and O.Helle for programming assistance.

6. REFERENCES


