# MODELING OF EFFECTIVE PROCESS OF NETWORK MAINTAINING BASED ON STATISTICAL DATA

Oleksandr Kuzomin, Nataliya Goloviy, Yaser Dayoub

*Abstract: This paper is dedicated to modelling of network maintaining based on live example – maintaining ATM banking network, where any problems are mean money loss. A full analysis is made in order to estimate valuable and not-valuable parameters based on complex analysis of available data. Correlation analysis helps to estimate provided data and to produce a complex solution of increasing network maintaining effectiveness.*

*Keywords: effectiveness estimation, statistical processes, correlation analysis*

*ACM Classification Keywords: H.3.4 Systems and Software*

## Introduction

Choosing equipment of this or that supplier, as a rule, pay attention to following parameters: the price, productivity, reliability and convenience. For a network of ATM where breakage or device stop means the lost money, service is the extremely important.

Equipment downtimes because of malfunction and repair, breaking business process, are capable to worsen economic indicators of banks sharply. Therefore, corresponding requirements also are shown to service services - to provide competitive advantage at modern level it is necessary to contain in an appropriate technical condition the equipment at the minimum losses of manufacture.

For the organization service and warranty service the system representing a complex of organizational-technical actions which realization would allow to make the basic part of works in advance known terms providing the prevention of progressing wear process of details technicians in operation and reduction of losses because of its malfunction and repair should be created.

At the heart of such system creation attempt of process modelling which usually not to do without application empiric and statistical the data about refusals and changes of parameters of a condition of elements and specifications of time for performance of preventive operations taking into account specificity of manufacture and increasing complexity of the applied equipment lies.

## Problem Statement

It is necessary to generate general idea about the complex character of the studied phenomenon expressed, in particular, in interrelations and interconditionality of separate features. To conduct research of the internal reasons forming specificity of the studied phenomenon, and to reveal the generalised factors which stand up for corresponding concrete indicators. Proceeding from the formulated problem definition, for an investigated subject domain it is necessary to study dependence between the features describing a condition of an ATM network.

An objective of this research is modelling of service process of an ATM network on the basis of statistical data for maintenance of its reliability and stability.

## Data matrix forming

During conducted research as set of investigated objects the ATM network, being on service of the company which have some feature set (indicators, characteristics) is considered. The features characterising investigated objects, carry both quantitative, and qualitative character. On the basis of the researches conducted and also expert appraisal, the features characterising each unit of a banking network were conducted.

| Para-meter | Definition | Type | Para-meter | Definition | Type |
|---|---|---|---|---|---|
| *m1* | Service Life | quantitative | *m13* | Load | quantitative |
| *m2* | Operating time | quantitative | *m14* | Quantity of ATM's in region | quantitative |
| *m3* | Service refusals | quantitative | *m15* | Quantity of engineers in region | quantitative |
| *m4* | Overdue failures | quantitative | *m16* | Quantities of temperature parameters infringement | quantitative |
| *m5* | Time in service | quantitative | *m17* | Quantities of power losses | quantitative |
| *m6* | Average time of malfunctioning | quantitative | *m18* | Quantities of grounding parameters losses | quantitative |
| *m7* | Quantity of spare parts replacements | quantitative | *m19* | Quantities of installation problems | quantitative |
| *m8* | Qualifications of the engineer | quantitative | *m20* | Quantities of usage ploblems | quantitative |
| *m9* | Quantity of installations | quantitative | *m21* | Client | qualitative |
| *m10* | Novelty | quantitative | *m22* | Type (Rent/Owned) | qualitative |
| *m11* | Reaction Time | quantitative | *m23* | Warranty | qualitative |
| *m12* | Repair time | quantitative | m24 | Region | qualitative |

Table 1. Characteristics of ATM

Here we have *n* objects (ATM), each of them is characterized by set of *m* features. Let's define as $x_{ij}$ value of *j-th* feature for *i*-th object. In that case data could be represented as a table called as "data matrix".

| Data matix | | | | | | |
|---|---|---|---|---|---|---|
| *Object number* | *Feature number* | | | | | |
| | *1* | *2* | *3* | … | … | *m* |
| *1* | $x_{11}$ | $x_{12}$ | $x_{13}$ | … | … | $x_{1m}$ |
| *2* | $x_{21}$ | $x_{22}$ | $x_{23}$ | … | … | $x_{2m}$ |
| *3* | $x_{31}$ | $x_{32}$ | $x_{33}$ | … | … | $x_{3m}$ |
| … | … | … | … | … | … | … |
| *n* | $x_{n1}$ | $x_{n2}$ | $x_{n3}$ | … | … | $x_{nm}$ |

Table 2. Data matrix sample

It consist of n lines (according to number of objects) and m columns (according to number of features), analysed sets of data consists of 908 objects.

## Correlation analysis

Studying of dependences between the features describing the difficult phenomenon or process, invariably brings an attention to the question: what reasons which have caused given structure of communications? Difficult dependences of such system of features are rather difficult for "untangling", as interrelations of features (correlations measured, for example, in the factors) can be interpreted not only as dependences of one features on others, but also as dependences on the certain latent parameters defining change of the whole groups of correlated features.

We will generate some statistical hypotheses about dependence of the chosen features. We believe that maintenance of reliability of a network of cash dispenses from outside the service organization consists in time reduction of ATM malfunction and reduction of average time of malfunction. We receive that reliability of a network is defined by indicators *m5-m6*. In turn these parameters depend on quantity of demands, delay of works and replaced spare parts (see Table 1). We investigate their dependence on other indicators for what we will carry out the correlation analysis.

We will calculate for existing quantitative features factors of correlation and we will construct a correlation matrix.

The structure of the experimental data which features in a context of a solved problem are described by mathematical model, is reflected by means of two basic categories of mutual relations between elements of a matrix of the data – similarity and distinction categories.

The matrix sets the relation "feature-featyre" and represents a two-dimensional symmetric square matrix of the size $m \times m$

$$R = \begin{pmatrix} r_{11},\ldots,r_{1m} \\ r_{21},\ldots,r_{2m} \\ \ldots\ldots\ldots\ldots \\ r_{m1},\ldots,r_{mm} \end{pmatrix},$$

where $R_{ij}$ — measure of connection between features $m_i$ и $m_j$ .

The considerable quantity of measures of communication between features is known. They differ both volume of calculations, and those aspects of communication which they reflect. Considering specificity of an investigated problem, we will consider two representation liaison groups between features. In the first group the covariation principle , and in the second — a principle of an associativity of features [2] is used.

Proceeding from the first principle, the conclusion about communication presence between variables becomes in that case when the increase in value of one variable is accompanied by steady increase or reduction of values another. In mathematical expression the problem is reduced to covariation calculation, that is accompanying change of numerical values of signs. Here to concern factor of correlation of Pirson (r) which represents product of the moments and is a measure of linear communication of two variables. It is calculated under the formula

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{(n-1)S_x S_y},$$

where $x_i$, $y_i$ - values of two variables, $\bar{x}$, $\bar{y}$ - their average values, $S_x$, $S_y$ - their average square deviation, *n* – amount of pairs of values.

Many measures of communication differ from the resulted factor of correlation of Pirsona the external form, but are, as a matter of fact, the algebraic transformation of this factor considering specificity (type) of compared signs. For example, the factor of rank correlations of Spirmena (rs), often applied to the analysis of numeric variables, represents algebraic simplification of factor of Pirsona r.

The same it is possible to tell about dotbiserial correlation factor (rpb) which serves as a communication measure between dichotomizing and quantitative variables. Some other factors, in particular tetrachoric (rtet) and biserial correlation factor (rbis), it is possible to interpret correlation factor as approximation for certain types of signs [5]. The more correlation factor on the module, the more strongly linear dependence. Value of factor of correlation lies between-1 and +1. If the tendency of increase of one size is observed at growth another speak about positive correlated sizes if the tendency of increase in one size is observed at reduction another speak about negative correlated sizes.

We will exclude from consideration parameters which unequivocally do not depend from each other. The greatest interest represents calculation and the correlation analysis between dependent (m3-m7) and independent (m1 - m2, m8-m20) parameters. It is possible to consider insignificant factors as zero and to take into consideration only the significant. We will reject insignificant parameters :quantity of installations and infringement of parameters of operation. If the correlation factor is calculated on the basis of the selective data it is not excluded that its nonzero value is not reflexion of the valid communication between signs, and is simply received as a result of specificity of the given sample (whereas the correlation factor is equal in a general totality to zero, i.e. linear communication between signs is not present).

Thus, it is necessary to understand, how r value is far from zero. If value of factor of correlation is calculated under the selective data, for an estimation of its value in a general totality as usually, the confidential interval is under construction.

For construction of a confidential interval the error of factor of correlation r under the formula is calculated:

$$\sigma_r = \frac{\sqrt{1-r^2}}{\sqrt{n-2}}$$

Then correlation coefficient error is multiplied by parameter *t*, which depends on probability *P* ( for probability 95% *t* = 2), in order to found maximal error.

At last, the confidential interval r - σr is under construction and check is spent, whether zero value will get to this interval. If the zero does not get to a confidential interval, means, with high probability in a general totality there can not be a zero value of factor of correlation, i.e. communication between signs exists and in a general totality. In that case the correlation factor is statistically significant.

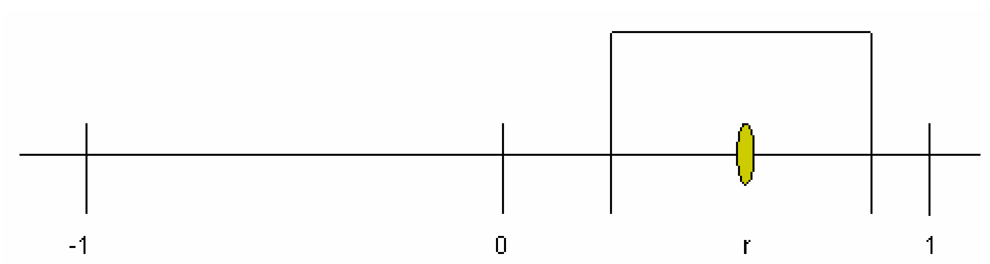


Figure 2. Confidential interval excluding zero

If the zero gets to a confidential interval, means, with high probability in a general totality there can be a zero correlation, i.e. absence of communication. In that case the correlation factor is statistically insignificant.

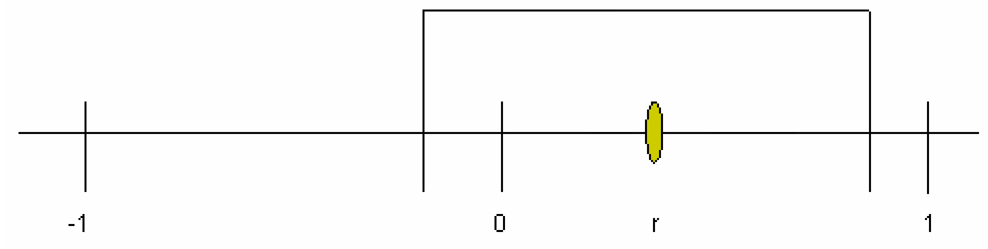


Figure 3. Confidential interval including zero

## Conclusions

On the basis of the received results it is possible to draw following conclusions:

- The quantity of refusals depends on parameters:
  - Service life (-)
  - Operating time (+)
  - Qualifications of the engineer (-)
  - Novelty (-)
  - Reaction Time (+)
  - Performance Time (+)
  - Quantities of temperature parameters infringement (+)
  - Quantities of pressure parameters infringement (+)
  - Quantities of the unit parameters infringement (+)
- Quantity of the delayed demands depends on :
  - Service life (-)
  - Operating time (+)
  - Novelty (-)
  - Load (-)
  - Quantities of engineers in region (+)
  - Quantities of cash dispenses in region (+)
  - Quantities temperature parameters infringement (+)
  - Quantities of pressure parameters infringement (+)
- Time of ATM repair most depends on following parameters:
  - Quantities of cash dispenses in region, where ATM is located (+)
  - Quantities of engineers in region (+)
  - Load(-)
- Average time of malfunction depends on:
  - Service life (+)
  - Operating time (-)
  - Qualifications of the engineer (-)
  - Novelty (+)
  - Loading (-)
  - Quantities of engineers in region (+)

  - Quantities of cash dispenses in region (+)
- Quantity of spare parts replacements depends on:
  - Service life (-)
  - Operating time (+)
  - Qualifications of the engineer (-)
  - Novelty (-)
  - Loading (-)
  - Quantities of engineers in region (-)
  - Quantities of cash dispenses in region (-)
  - Quantities temperature parameters infringement (+)
  - Quantities of pressure parameters infringement (+)
  - Quantities of grounding parameters infringement(+)

The feature near to each parameter testifies about sizes correlation – positive and negative.

Provided information allows greatly reducing losses due to network processing and estimating the common plan of ATM service works

## Authors' Information

*Oleksandr Kuzomin – Head of Innovation marketing department, professor in informatics dept.; Kharkiv National University of Radio electronics, e-mail: kuzy@kture.kharkov.ua*

*Nataliya Goloviy – Postgraduate, Informatics dept.; Kharkiv National University of Radio electronics, e-mail: rica1982@mail.ru*

*Yaser Dayoub – Postgraduate, Informatics dept.; Kharkiv National University of Radio electronics, e-mail: yaser_daeob@hotmail.com*