ELECTRONIC JOURNAL
OF INTERNATIONAL GROUP
ON RELIABILITY

JOURNAL IS REGISTERED
IN THE LIBRARY
OF THE U.S. CONGRESS

●

ISSN 1932-2321

●

VOL.1 NO.4
DECEMBER, 2008

Gnedenko Forum Publications

Reliability & Risk Analysis Theory & Applications

№4
2008

Special Issue on International Scientific School "MODELLING and ANALYSIS of SAFETY and RISK in COMPLEX SYSTEMS" (MA SR)

San Diego

ISSN 1932-2321

http://www.gnedenko-forum.org/Journal/index.htm

Journal of International Group on Reliability

RELIABILITY & RISK ANALYSIS: THEORY & APPLICATIONS

Vol.1 No.4,
December, 2008

**Special Issue on International Scientific School
"MODELLING and ANALYSIS of SAFETY
and RISK in COMPLEX SYSTEMS"
(MA SR)**

San Diego
2008

Journal Council

Send your paper

e-Journal ***Reliability: Theory & Applications*** publishes papers, reviews, memoirs, and bibliographical materials on Reliability, Quality Control, Safety, Survivability and Maintenance.

Theoretical papers have to contain new problems, finger practical applications and should not be overloaded with clumsy formal solutions.

Priority is given to descriptions of case studies.

General requirements for presented papers

1. Papers have to be presented in English in MSWord format. (Times New Roman, 12 pt , 1.5 intervals).
2. The total volume of the paper (with illustrations) can be up to 15 pages.
3. A presented paper has to be spell-checked.
4. For those whose language is not English, we kindly recommend to use professional linguistic proofs before sending a paper to the journal.

* * *

The Editor has the right to change the paper title and make editorial corrections.

The authors keep all rights and after the publication can use their materials (re-publish it or present at conferences).

Publication in this e-Journal is equal to publication in other International scientific journals.

Papers directed by Members of the Editorial Boards are accepted without referring.

The Editor has the right to change the paper title and make editorial corrections.

The authors keep all rights and after the publication can use their materials (re-publish it or present at conferences).

Send your papers to

the Editor-in-Chief ,
Igor Ushakov
igorushakov@gmail.com

or

the Deputy Editor,
Alexander Bochkov
a.bochkov@gmail.com

Table of Contents

The article deals with the possibility of system availability prediction using the simulation modelling. The system availability determined with system faultlessness and system maintainability is expressed by various parameters of mean time between the failures and the mean time of single elements repair. The system simulations are carried out with more parameters MTBF and MTTR, the results of the simulation course gives a real idea about the system behaviour in time and about changes of the values of asymptotic system availability factor.

This paper deals with optimal control interval determination using minimization the financial costs. It clears conceptual, mathematical and simulation model of the problem solution. It enumerates and evaluates results of the simulation.

This paper deals with possibility of simulation of reliability block diagrams, failure trees analysis as a time dependent analysis using Matlab/Simulnk

A novel method is proposed for hard optimization type of problem wherein an exact optimal solution is increasingly difficult in terms of run time and memory requirements. Especially for the cases when search graph has higher number of nodes and more number of paths, which increase as factorial of node number. This is based on Simulated Electrical Network Approach (SENA) proposed here, in which the graph is modeled as an electrical network and current distribution is found which is used as a directive for search decisions. The proposed algorithm results in an approximate method that achieves average accuracy of 99.89% to reach close to the most optimal path that is found by ranking all possible paths. Conversely, it can eliminate on average 99.89% paths in polynomial time from consideration if one requires finding the most optimal one.

Probabilistic methods of risk optimization are applied to specify the most effective arrangements of road tunnels. The total consequences of alternative arrangements are assessed using Bayesian networks supplemented by decision and utility nodes. It appears that the optimization may provide valuable information for a rational decision concerning number of escape routes. Discount rate seems to affect the total consequences and the optimum arrangements of the tunnels more significantly than number of escape routes.

The problem of development of Boolean models of a reliability for systems, including elements with many states is considered on the basis of multivalued logic, algebra of trains, algebra of groups of incompatible events and classical logistic-probabilistic method (LPM). The inexpediency of development of Boolean models of a reliability on the basis of multivalued logic is displayed. The numerical examples demonstrating serviceability of LPM and their new possibilities are demonstrated. The perspective of development of methods of an evaluation of effectiveness of operation at different levels of operation rate by formulation of a set of different tasks, solved by the same LPM is underlined

The stages of development of Management and Risk are described. The scenario management of risks of accidents and catastrophes in complex systems on the stages of designing, debugging and exploitation test and exploitation itself are considered. In the scenario management of accidents and catastrophes risks the personnel and the General designer are taken into account. The uniform approach to the modelling of risks in technical, economic and organisational systems is presented on the basis of substantial description of a SCENARIO of an accident or a catastrophe, and then the construction of models of the risk for the purpose of analysis and management. As the intellectual core for the risk quantitative evaluation and analysis and the scenario management of accidents and catastrophes risk, LP-methods and risk LP-models with groups of incompatible events are used.

Safety Management is intended to create order out of disorder, to reduce the “information entropy”, for the purpose of improved safety. Our purpose here and now is to try to introduce some predictability and insight into the risk or occurrence of apparently random events, where a general risk prediction we adopt a fundamental must be testable against the world’s existing data. The risk management issues are clear, given the classic features of major human involvement and contribution to accidents, errors and outcomes occurring with modern technological systems. Prior incidents and prior knowledge and experience must be fully incorporated or learned from. If we do not know where we are on the learning curve, we also do not know the probability of such an event, and we have no objective measure of the “safety culture”. Emphasis on defining and finding so-called “lack of safety culture” has resulted in an extensive and detailed study of the safety management and process safety of many global corporations. We utilize the concepts adopted in thermodynamics and Information Theory to establish the information entropy as a finite, physically based and useful measure of risk in technological systems. The results that we demonstrate show that the risk is dynamic, and can be utilized for management and predictive risk analysis purposes.

This article analyses the traffic accident rate on roads and highways and possibilities of risk evaluation related to traffic accident occurrence based on factors that were the causes of accidents. A new term – risk of traffic accident occurrence is a product of probability of accident occurrence and its impacts. The results are presented by way of example that uses selected statistical data of the Czech Republic traffic accident rate between 1993 - 2001. The article provides a brief methodological procedure of evaluation of the traffic accident rate using the risk of traffic accident occurrence.

This paper presents the scope and some main results of a European project on the ASSessment of Uncertainties in Risk ANalysis of Chemical Establishments (ASSURANCE). The project aims at identifying the uncertainties associated with risk analysis of major industrial hazards and assessing the way these uncertainties can affect the final outcome of risk studies and of the relevant decisions based on that outcome. In order to achieve this goal, a number of benchmark exercises/case studies have been performed by the partners and the results were analysed in a modular and structured way. A reference plant served as the basis for a realistic description of these case studies. For this particular project an ammonia storage plant was selected, consisting of cryogenic and pressurised storage tanks, together with import loading/unloading facilities and the relevant piping. This installation was analysed independently by each partner, using common input data and boundary conditions, but different methods, tools and assumptions. The results were then compared and discrepancies identified, discussed and explained.

Inherent safety of the new generation airships, based on some fundamental laws of Space, is discussed in some detail. An algorithm is proposed to analyze risks, resulting from hazards not compensated by "inherent safety". Then a thoroughly verified statistical model of learning is used to evaluate results of airship flight testing-the probability of mission success and its confidence limit. The results can be used as a part of evidence for airship airworthiness certification.

To keep the fatigue ageing failure probability of an aircraft fleet on or below the certain level an inspection program is appointed to discover fatigue cracks before they decrease the residual strength of the airframe lower the level allowed by regulations. In this article the Minimax approach with the use one- and two-parametric Monte Carlo modelling for calculating failure probability in the interval between inspections is offered.

A general approach for analysing spatial survival in the plane is suggested. Two types of harmful random events are considered: points with fixed coordinates and moving points. A small normally or tangentially oriented interval is moving along a fixed route in the plane, crossing points of initial Poisson random processes. Each crossing leads to termination of the process with a given probability. The probability of passing the route without termination is derived. A safety at sea application is discussed.

Problem of representation of human preferences among uncertain outcomes by functionals (risk measures) is being considered in the paper. Some known risk measures are presented: expected utility, distorted probability and value-at-risk. Properties of the measures are stated and interrelations between them are established. A number of methods for obtaining new risk measures from known ones are also proposed: calculating mixtures and extremal values over given families of risk measures.

In this paper the results of the researches in identification of the logical and probabilistic (LP) risk models with groups of incompatible events are presented. The dependence of the criterion function on several parameters has been investigated. The parameters include: the total number of optimisations, the amplitude of parameters increments, the initial value of the criterion function (CF), the choice of identical or different amplitudes of increments for different parameters, objects risks distribution. An effective technology of defining the global extreme in the identification of LP-risk model for the calculation time, appreciable to practice has been suggested.

The application of RAMS techniques in all the phases of the lifecycle of each type of installation will surely guarantee its adequate exploitation in terms of production continuity and quality of the obtained products in the respect of prefixed constraints on the security of the working staff, safety and environment impact. In this frame, a particular importance must be attributed to the use of those techniques as support to quality assurance applied in the planning and building phases of the installation and of the products obtained by it. The present paper will include a short description of a method for the application of those techniques in this phase of the lifecycle and of the results that may be obtained by its application in shoes manufacturing, in particular those types where the technical requirements are higher, as it is the cases of certified products like "safety" footwear.

SCIENTIFIC SCHOOL «MODELING AND ANALYSIS OF SAFETY AND RISK IN COMPLEX SYSTEMS» - ACTUAL APPROACH TO ACTUAL PROBLEMS

Karasev V.V.
●
Candidate of Technical Science, senior research worker,
Russia, Saint-Petersburg, IPME RAS

Solojentsev E.D.
●
Doctor of Technical Science, Professor,
Honoured Worker of Science of Russian Federation
Russia, Saint-Petersburg, IPME RAS

All wise men should be in a dialog
Titus Maccius Plautus (254–184 B.C.)
Roman comedy writer

Dear colleagues !

We are living in interesting time, in period of global changes all sides of human life: economics, politics, environmental conditions, ideologies, ideals, culture and cultural wealth.

Scientific and technological advance grows rapidly, obtaining of new knowledge increases twice every ten years. 80% per cent of all scientists, who lived ever on Earth, are our contemporaries. Every minute 2000 pages of new knowledge are added to scientific heritage. To study these materials we need to read it uninterruptedly during 5 years. About 500 thousands new books are published every year.

The development of information technology, communications and Internet has a great influence on knowledge industry's growth because information plays important role. Well-timed information allows understand modern trends in development of economics, science and society, follow newest discoveries and elaborations, make correct decision in changeable conditions.

Therefore, bridging of scientists, researchers and developers is extremely important task and scientific conferences and symposiums are best means for communication between members of scientific society.

Of course, the main source of information where modern scientists get actual information about newest developments and research results is periodicals and Internet, but, it's needed to note, papers, which are published in periodicals (especially in reviewed ones), as a rule, contain finished research with final results. On the other hand, conference gives unique chance to study a work at initial stage of its research and understand tendencies of modern science at their beginning. Scientist can be informed at first hand about last developments, not described in periodicals, and who and in what research is engaged at present, pick up new interesting ideas.

Besides, lively productive talk with author of new approach or new method is more productive for disputants than one-sided examination of his (her) work in journal or another periodicals, where the volume of the paper and rules of publishing usually are limited by requirements of publishers.

Scientific School «Modeling and Analysis of Safety and Risk in Complex Systems» (MASR) is held annually since 2001 in Saint-Petersburg with financial support of Russian Fund for Basic Research (RFBR). Traditionally, there are two sections: "Risk in Engineering and Ecology" and "Risk in Economics and Finances".

Every year scientists and specialists from more than 15 countries arrive in Saint-Petersburg to make a lecture, meet colleagues, put questions and get answers, to discuss perspectives of the collaboration.

Scientific School (MASR) is focused on actual problems of quantitative estimation and analysis of risk and safety in various areas of business and engineering. Scientists and specialists in nuclear power engineering, navy, aviation and space vehicles, economists from investment companies, financial institutions, banking and business are taking part in the School.

During holding of eight Scientific Schools MASR in period 2001-2008 years more than 500 scientist from 34 countries have took part. More than 650 scientific papers were published in eight MASR Proceedings.

Idea of holding of Scientific School «Modeling and Analysis of Safety and Risk in Complex Systems» have appeared for a good reason. Russian scientists, leading specialists in reliability theory and safety theory for complex systems, K.V. Frolov and I.A. Ryabinin have initiated MASR.

Idea of MASR was been supported by director of Institute of Problems of Mechanical Engineering of Russian Academy of Sciences (RAS), Doctor in Technical Science Vladimir Pavlovich Bulatov (1937–2002), Corresponding Member of RAS Nikolay Andreevich Machutov and Academician of RAS Konstantin Vasilevich Frolov.

Too many specialists were engaged in area of estimation and analysis of risk in technical and economical systems but they worked in isolation. Many unique developments and decisions remained in closed laboratories, departments and companies and were not available to anybody. Scientific School was organized to get possibility for scientists to share experience with world scientific society and for other scientists to study these developments.

First School MASR–2001 was held in June 18–22, 2001 in Institute of Problems of Mechanical Engineering of RAS and Saint-Petersburg Institute of Informatics of RAS with financial support of RFBR and had excellent success. Actual problems of estimation and analysis of risk were discussed and ways to solve these problems were offered.

However, MASR have become popular not only among specialists, engaged in specialized problems. In MASR topics new problems were opened, which were actual at the end XX – beginning XXI centuries not only in engineering but in economics and business also.

Topics of the School involve problems of the capital market risk, including market risk, security portfolio risk, operational risk, credit risk. Methods of modeling and analysis of the risk in engineering, development of quantitative risk analysis methods, optimization of risk models, distribution of resources to risk management, are discussed. The problem of risk analysis and efficiency of social and economical processes under statistical data is formulated. Two types of events are distinguished; the appearance of system's state and failure of system's state.

Today, main goal of International Scientific School is a chance for scientists and specialists to share results of their theoretical and practical research with colleagues in area of modeling and analysis of risk, and, to establish interdisciplinary connections for development of universal risk theory.

Other goal of International Scientific School is introduction of mathematicians, economists and managers to logical and probabilistic theory of the risk with groups of incompatible events in problems of the classification (credits), investment (security portfolio), efficiency (social processes), management of the company, corruption and bribes; consideration of the connection of databases and knowledge bases in above-mentioned problems, transition from VaR models to logical and probabilistic models of the risk of classification.

Great benefit of Scientific School is the possibility to learn newest developments in methods of estimation, analysis and management of risks. Works, presented at Scientific School, give some participants new ideas for Ph.D. thesis and thesis for a Doctor's degree. With using of results of presented works, several Ph.D. thesis and thesis for a Doctor's degree were written and successfully defended. MASR participants have written also several scientific monographs.

It's necessary to note, the scientific importance of papers, presented by participants, is high enough and annually increasing. In July, 2003 the special issue of the known scientific reviewed journal "Automation and Remote Control" № 7, vol. 64, was published. This issue contains selected MASR papers that were published in 2001-2003 years.

At present moment we are glad to offer you selected MASR papers that were published in MASR Proceedings during period 2001-2008 years.

In conclusion, we are very thankful to MASR participants those valuable material was used as a basis for this issue, and to all our participants for their interest and activity. Also, we are very appreciate to organizers of e-forum Gnedenko and Alexander Bochkov personally, who provide good information support and promotion of Scientific School MASR.

PREDICTION OF NO – FAILURE SYSTEM OPERATION

Alexej Chovanec
●
Faculty of Special Technology / Alexander Dubcek University in Trencin
Studentska 1, 911 50 Trencin, Slovak republic
e-mail: chovanec@tnuni.sk

Abstract: The paper is focused on analytical approach to prediction of ability and resources of simulation. It deals with simulation experiments with static approach except using time and selection decisive event.

Keywords: reliability, life cycle, reliability block diagram, faultlessness, probability, static simulation, stochastic, element, failure - free operation

1. INTRODUCTION

Systematic attention in all stages of the product's life cycle is supposed in order to assure the technique reliability security.

The reliability bases are forming in the first periods of the technique creation. The rational determination of qualitative and quantitative demands on the reliability so called *specific demand on the reliability* is the decisive task.

It includes the set of the activities realised in the stage of the conception and demands determination and the stage of the design and development, which aim is the specification of the demands of product reliability as the unit and also the reliability of single parts.

Predictive analyses of the reliability of faultlessness indicators observing and prediction, preparedness, maintainability and safety of the system are applied.

The reliability analyses are realised also in the stage of the usage, operation and the maintenance for the evaluation and determination or specifying the indicators of the reliability and for the assessment whether the specified demands were fulfilled.

The analysis of the system reliability is the process, which is usually realised on the system model. The set of information about the properties of the system model is the final product of this process. The model can be modified during the analysis.

Analysis has to have clearly determined rules and processes so as the process of the analysis to be repeatable and in order to always lead to the same results.

For the concrete case we have to choose a suitable analytical method, which enables to (F):

- ➢ model and evaluate the reliability problems in the required range,
- ➢ make direct, systematic qualitative and quantitative analysis,
- ➢ predict numeral values of the reliability indicators.

In the present practise the most often used methods of reliability analysis are:

- o Reliability Block Diagrams,
- o Fault Tree Analysis,
- o Markov analysis,
- o Failure mode and effects analysis, FMEA,
- o Simulation methods.

The simulation modelling is suitably applied within the frame of the separation (allocation) of the demands on single parts of the product (or parts of the process of maintenance providing). The aim is to determine the demands on the critical reliability or of each product part in such a way so that the product as the complex fulfils the determined demands.

The allocation of the demands on the reliability is closely associated with the process of the design and evaluation of the product and its steps can be repeated in connection with the changes of the design or on the basis of optimising studies, feedback from the operation.

The indicators of the reliability on the lower levels of the product structure can differ from those, which were defined in the product specification. For example, the repaired product can be built up from the parts, which had been never repaired.

2. RELIABILITY BLOCK DIAGRAMME

Prior estimations of the reliability indicators are obtained above all by calculation in the stage of projection. They can be received also by following of the critical elements reliability in the period of evaluation and testing and specified as the result of the feed-back information form the statistic plotting of operation failures and the degradation of elements plotted during the repair actions.

At the specification of demands on the reliability generally the numerical values of partial reliability properties for the technique as the unit are defined.

The reliability of the technique as a unit is the reflection of the reliability level of single groups, subgroups, functional pairs and components, from which the technique consist of. That's why also in the case of reliability prediction the decomposition of reliability demands is necessary.

From the whole level of the reliability the demands of the lower levels of constructional elements are defined. The effort is to get to the level of the elements in the serial or parallel structure.

Due to it, the transformation of the complicate structure to the simple one is sometimes made in order to be able to express the mathematical level of the reliability.

On the contrary, at the well-known values of the reliability of the decisive elements, it is possible to determine the final level of the system reliability.

The probability reliability analysis - also called the reliability block diagram (RBD - Reliability Block Diagram), whose bases are the Bool algebra of events, logical, oriented and undirected graphs, calculus of probabilities, is the basic tool of reliability observation.

The probability reliability analysis is the method of taking into account the probabilities of events occurring in complicate systems, which represent various arranged structures formed with elements.

The structure of the system can be expressed by the equation:

$$M_k = \{ E_1, E_2, ... E_k \} \qquad (1)$$

Single symbols indicate:

M_k – mechanical system with „k“ elements,

E_i – i element of the system.

The structure of the system elements can be serial, parallel or combined.

2.1 SYSTEMS WITH SERIAL CONNECTION OF ELEMENTS

The elements of the set are arranged one after another and they are each other independent. The failure of unique element causes the loss of operational capability of the whole system. The system is functional if all elements are in the state of operational capability.

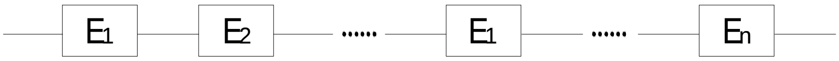


Fig. 2.1 Scheme of the system with the serial connection of the elements

If we mark the faultlessness of i-element E_i as R_i, then the faultlessness of the system is the product of the faultlessnesses of all elements:

$$\boldsymbol{R_S} = \boldsymbol{R}_1 \times \boldsymbol{R}_2 \times \boldsymbol{R}_3 \cdots \boldsymbol{R}_n = \prod_{i=1}^{n} \boldsymbol{R_i} \qquad (2)$$

The faultlessness of the serial system is lower than that of the most faultless element of the system. The reliability of the failure-free operation of the serial system with the number of elements decreases and the probability of the failure creation increases.

2.2 SYSTEMS WITH PARALLEL CONNECTION OF THE ELEMENTS

The increase of the reliability of the elements and the systems is ensured with the use of parallel (advance, reserve, redundant) elements. The failure occurs only in the case all elements of the system, basic and also advance, are damaged. The system is serviceable if at least one of the equal, independent elements is functional.

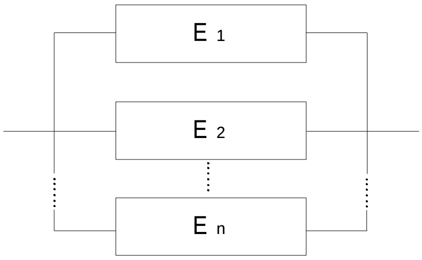


Fig. 2.2 Scheme of the system with parallel connection of the elements

The faultlessness of the parallel system can be calculated from the expression:

$$\boldsymbol{Rp} = 1 - \boldsymbol{F_S} = 1 - (1 - \boldsymbol{R}_1)(1 - \boldsymbol{R}_2)...(1 - \boldsymbol{R_n}) = 1 - \prod_{\boldsymbol{i}=1}^{\boldsymbol{n}} (1 - \boldsymbol{R_i}) \qquad (3)$$

If all elements have the same faultlessness R, then:

$$\boldsymbol{R_S} = 1 - (1 - \boldsymbol{R})^{\boldsymbol{n}} \qquad (4)$$

Also some specific cases of backup exist, for example if K elements are enough from N elements with same faultlessness for the security of the operational capability of the system.

Such a system is marked as k from n and the faultlessness of such a system can be calculated according to:

$$\boldsymbol{R_S}(\boldsymbol{k}, \boldsymbol{n}, \boldsymbol{R}) = \sum_{\boldsymbol{r}=\boldsymbol{k}}^{\boldsymbol{n}} \binom{\boldsymbol{n}}{\boldsymbol{r}} \boldsymbol{R}^{r} (1 - \boldsymbol{R})^{\boldsymbol{n}-\boldsymbol{r}} \qquad (5)$$

If the elements have various faultlessness, we will use the equation:

$$\boldsymbol{R_S} = \boldsymbol{R}_1 \boldsymbol{R}_2 + \boldsymbol{R}_2 \boldsymbol{R}_3 + \boldsymbol{R}_1 \boldsymbol{R}_3 - 2\boldsymbol{R}_1 \boldsymbol{R}_2 \boldsymbol{R}_3 \qquad (6)$$

2.3 SYSTEMS WITH SERIAL-PARALLEL CONNECTION OF ELEMENTS

The majority of real systems consists of serial and parallel connected subsystems, which are called combined. The final faultlessness of the whole system is calculated from the previous mentioned equations by the suitable dividing of the system to serial or parallel subsystems.

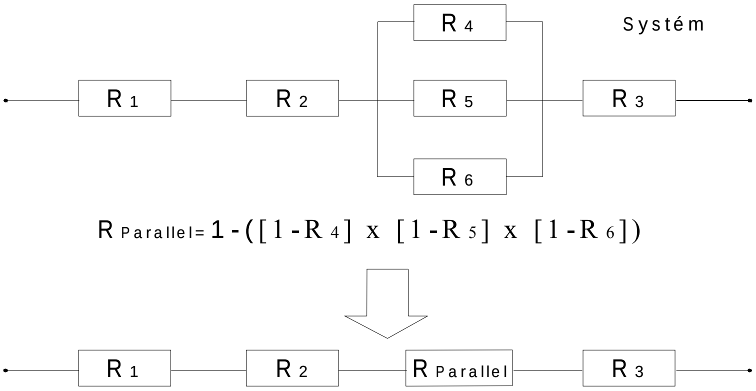


Fig. 2.3 Scheme of combined configuration of elements calculation

3. SIMULATION MODELLING OF FAULTLESSNESS

We can be express the static approach of the creation of simulation model of the probability reliability analysis by the following notional model:

- Serial – parallel sequenced system is characterized by the number of serial subsystems and number of the subsystems sequenced parallel.
- Each subsystem is defined by the number of elements and each element by the value of the probability of the failure-free operation.
- In the state vector of the system $X_{Mk(t)} = (X_{1(t)}, X_{2(t)}, \ldots, X_{k(t)})$ and its elements $X_{i(t)}$ the time of system activity t is constant, deterministically defined value. Simulation experiment is defined by the number of realisations representing constant time intervals between the failures.
- The elements of the system are characterized by two basic states:
  – failure-free – serviceable state P_A,
  – failure state P_B.

  The basic states create the whole group of events and it is given $P_A + P_B = 1$

  The states of the elements can be expressed by logical nil or by one depending whether the case occurred or did not occur, for example:

  $P_A = 1$ - the element is in failure state,
  $P_B = 0$ - the element is in failure-free state.

  The generally causal change of nil and one in the system is the state quantity $X_{Ei(t)}$ expressed by the expression:

$$\{X_1(t) : t \geq 0 \} \rightarrow E_1,$$
$$\{X_2(t) : t \geq 0 \} \rightarrow E_2,$$
$$\vdots$$
$$\{X_k(t) : t \geq 0 \} \rightarrow E_k.$$

- The probabilities of elements failure-free operation are constant in time. They can be taken for and modelled as causal values of equal separation of the probability of failure- free operation of elements E_i expressed by the mean value.
- The generation of the event "failure" can be generated from the equal division in the range 0 - 1.

  $P_A + P_B = 1$ - the whole group of events
  $P_A = 0{,}97$ - probability of failure-free operation
  $P_B = 0{,}03$ - probability of failure generation

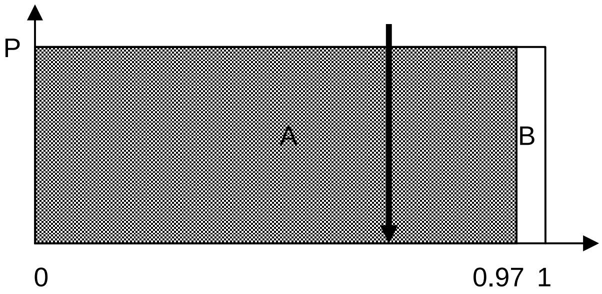
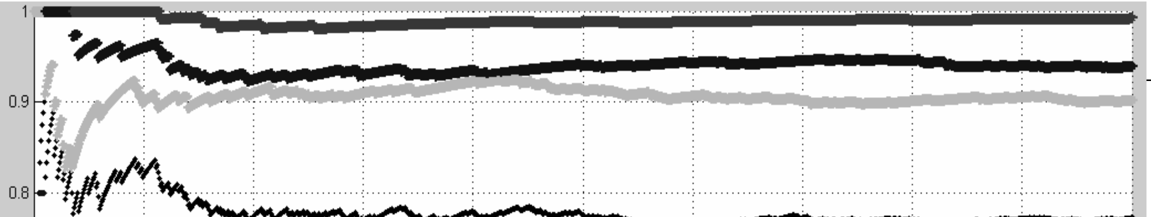


Fig. 3.1 The figuration of events occurring generation

- If the generated value exceeds the determined value of the failure-free operation probability, then the failure of the element occurs and vice - versa.
- At the parallel subsystem, one function element is enough so as the subsystem to be function, at the serial subsystem, all elements have to be function.
- The system is functional if all subsystems are functional.
- The number of events, which represents the failure - free operation of the element or of the system is marked n.
- If we simulate the generation of the element or system failure N-times then the final probability of the failure-free operation is defined by the expression R=N/n.
- The program collects the output characteristics of the elements and system.



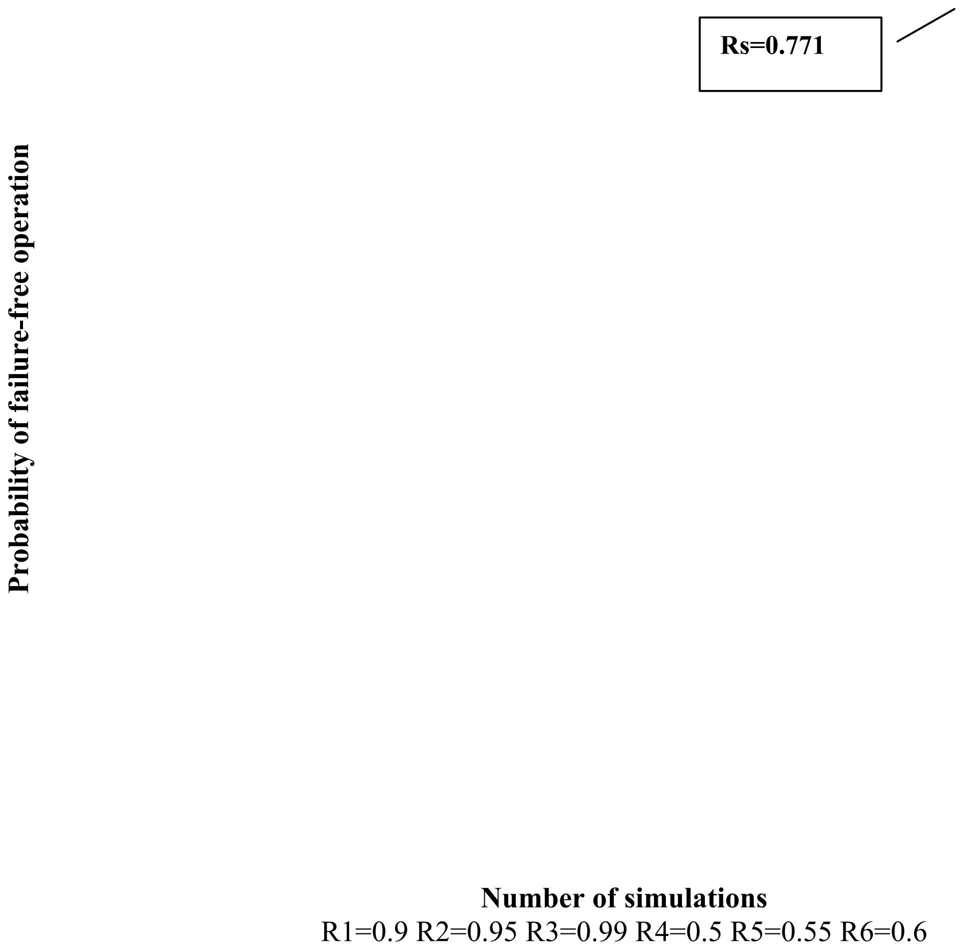


Fig. 3.2 Input data, function development and results of simulation experiments from static simulation model with structure according to fig. 2.3

The static approach does not enable to determine the final level of the reliability at various dividing rules of elements probability and it does not give any real idea about the failure creation in time.

Analytical ways of calculations of probability reliability analysis have restricted usage for the same statistic divisions of the probability of elements failure-free operation, which can be mathematically expressed. It is most often the case of exponential or normal probability separation.

We are not able to solve analytically the more complicated structures, other various rules of probability dividing, due to the lack and the complexity of the mathematical apparatus.

The stochastic approach of the probability of the reliability analysis modelling can remove the above-mentioned drawbacks.

If we can express the probability of the elements failure-free operation by the parameters of the division of causal variable intervals between the failures, the total probability of the failure-free operation can be expressed by the summary statistic parameters determined from the values of the total amount of generated data.

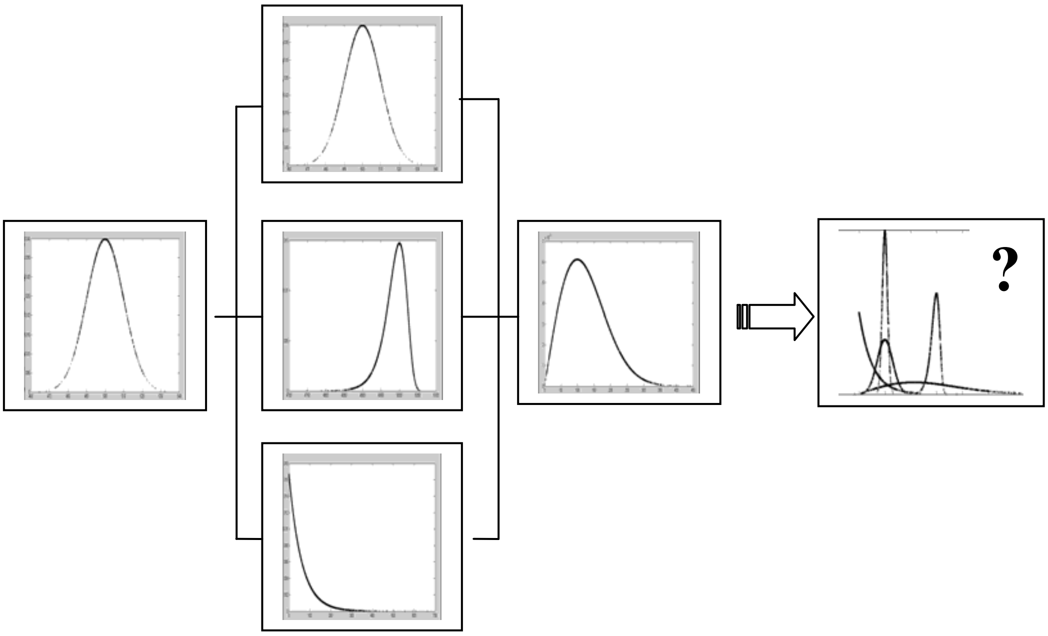


Fig. 3.3 The figure of the stochastic philosophy of modelling

3.1 STOCHASTIC SIMULATION MODEL OF FAULTLESSNESS WITH THE CHOICE OF DECISIVE EVENT MÔŽEME VYJADRIŤ NASLEDOVNÝM POJMOVÝM MODELOM

- The system is divided into subsystems with serial and parallel structures. The values of the periods between single elements failures are generated.
- The highest values of failure generation times are chosen from the times of failure creation of the subsystem from the parallel elements and they are used for the integration into the serial structure of the system.
- The lowest value of the failure generation time is chosen from the times of failure creation of the serial connected subsystems.
- The process of generating of events choice is repeated until the end of the realization number.
- Statistic data are collected about operation time, total number of elements failure and other necessary data

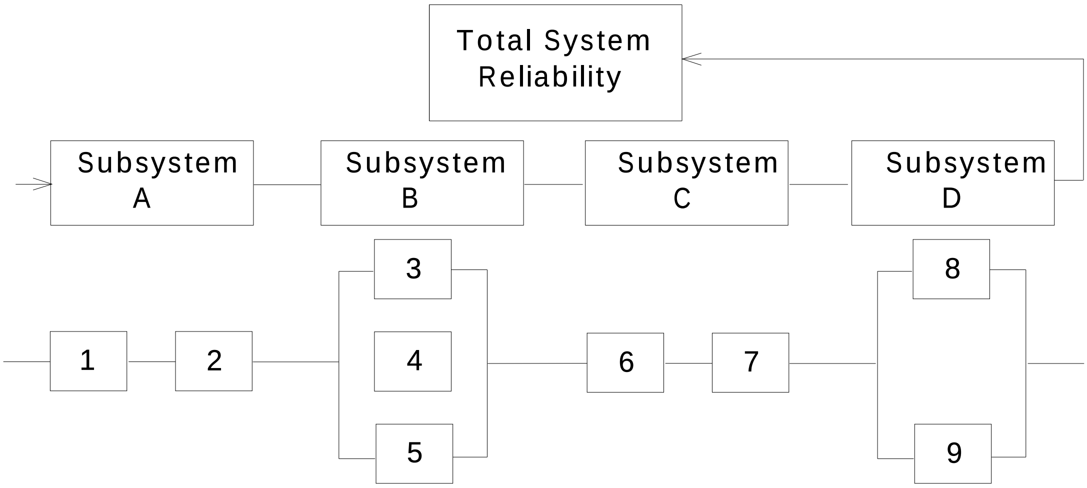


Fig. 3.4 System decay with combined structure

System is made of 6 elements with series-parallel structure according to *fig.2.3*.

Elements are decribed with time to failure distributions according to computer language MATLAB:

R1=exprnd(260);	R2=normrnd(80,18);
R3=wblrnd(90,2);	R4=exprnd(50);
R5=normrnd(50,10);	R6=wblrnd(20,1.5);

Simulation experiment was made for 10000 simulations. During simulation cycle, maximum event was chosen from parallel structure. This event was used in series structure and minimal event was chosen for system failure.

Computed times to failure of elements 2,4,6 are depicted with probability density function and cumulative density function on *fig.3.5* and *fig.3.6*.

Probability density function and cumulative distribution function of the decisive event – mean time to failure is depicted on *fig. 3.5* and *fig. 3.6*.

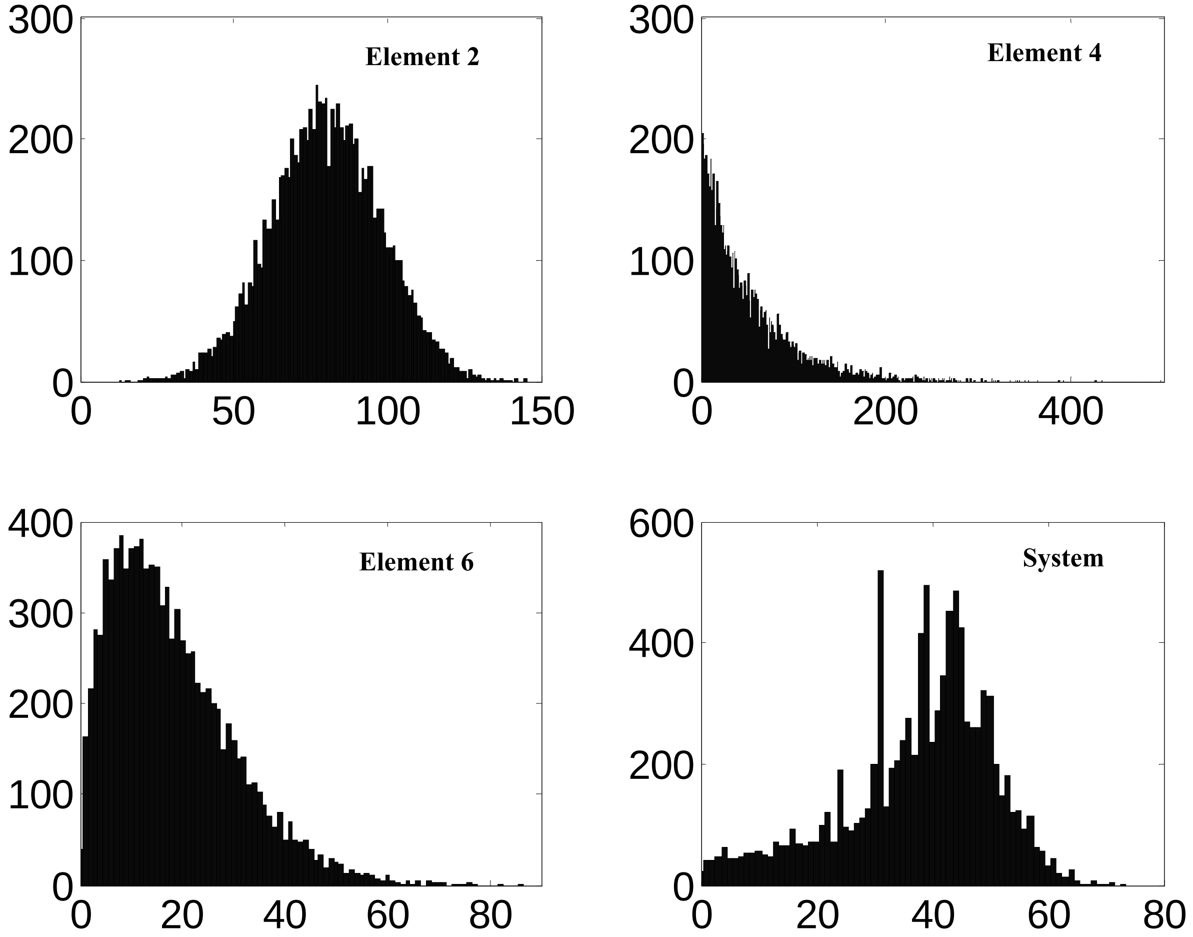


Fig. 3.5 Histogram of generated values of times between failures of elements 2,4,6 and value of system's time between failures

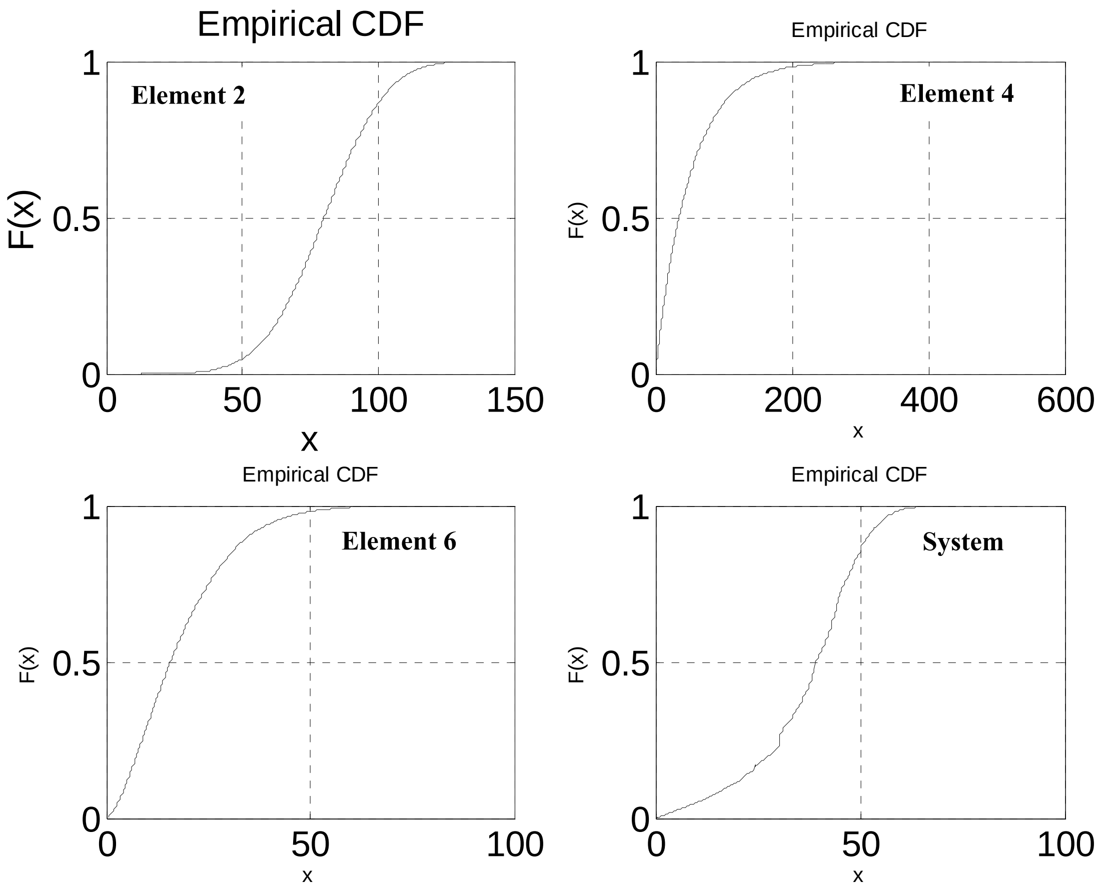


Fig. 3.6 Distribution functions of time between failures of elements 2,4,6 and mean time between failures of the system

4. CONCLUSION

Static a stochastic quantitative analysis ensures the calculation (estimation) of quantitative numerical values of chosen reliability indicators. The numerical value of the indicator is obtained by the experimentation with the model with the help of the computer technique, at the consideration of elementary effects, which structurally join the model into behaviour and analytical stages of the system.

The model and all inputs have stochastic character, also the result of the analysis is stochastic, loaded by a single rate of uncertainty, which is possible to decrease but not totally remove.

Calculation using analytical methods of probability intersection of several phenomena with different kinds of probability distributions is not possible.

References

1. Holub, R., Vintr, Z.: Aplikované techniky spolehlivosti. Část 1: Specifikace požadavků na spolehlivost. [Skripta] Brno: Vojenská akademie 2002.
2. US Army TM 5-698-3 Reliability primer for command, control, communications, computer, intelligence, surveillance, and reconnaisance (C4ISR) facilities, 2003
3. Leitner, B. : Optimum design of Track Maintenance machine Frames by Matlab. In: Zborník medzinárodnej konferencie "TRANSPORT 2003, VTU Todora Kableshkova, Sofia 2003.

This article was elaborated at the tasks realisation of granted project VEGA 1/2094/05

PREDICTION OF THE SYSTEM AVAILABILITY USING SIMULATION MODELING

Alexej Chovanec
●
Faculty of Special Technology / Alexander Dubcek University in Trencin
Studentska 1, 911 50 Trencin, Slovak republic
e. mail: chovanec@tnuni.sk

Abstrakt: The article deals with the possibility of system availability prediction using the simulation modelling. The system availability determined with system faultlessness and system maintainability is expressed by various parameters of mean time between the failures and the mean time of single elements repair. The system simulations are carried out with more parameters MTBF and MTTR, the results of the simulation course gives a real idea about the system behaviour in time and about changes of the values of asymptotic system availability factor.

Keywords: sampling size, fault time, interval between failure, normal distribution, financial costs, simulation experiment, optimisation process, probability density, optimal maintenance interval

1. INTRODUCTION

At the design of mechanical system consisting of several subsystems or elements we must generally predict the final reliability level characterised by the system availability. The estimation, which arises from the well-known, eventually estimated availability level of single subsystems is one of the possible ways.

The aim is to determine the system availability from the understanding of the factors of faultlessness partial properties, maintainability and arranging of single components maintenance.
The so-called states analysis, in which the system can occur, is the base for the design of the system availability model. The system can be in many and various states, whereas each of them is determined by a specific combination of single elements states.

Similarly, each system elements can occur in various states, which are randomly changing. The process, when the states of the studied objects are randomly changing in the time, is called Markov random process.

The most often the states in mechanical systems are expressed by a two - state model. The system, depending on the state of single elements, can occur either in function or in non-functional state. If the transition between these states is randomly changing and they can occur in arbitrary time, then this random process is usually called **common process of recovery** (1).

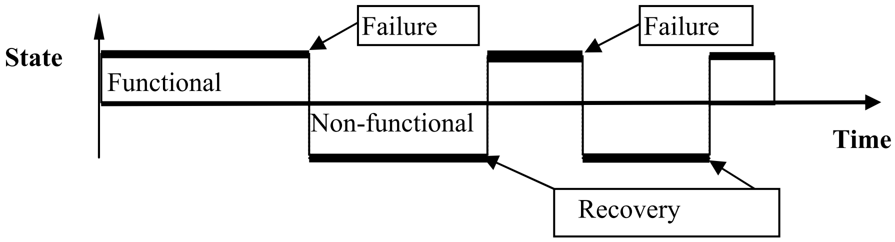


Fig.1.1 Common process of recovery

2. MECHANICAL SYSTEMS AVAILABILITY

The reliability of the objects being repairing is characterised above all by the availability indicators, which completely describe their faultlessness and maintainability.

The availability indicator is a function or numerical value used for the description of the probability distribution of a concrete studied (random) quantity, which characterises the object availability. The state of the object, which is randomly changing in time, is generally such a quantity.

The probability, in which state the object (element, system) occurs in the certain time, is described for the operation state by the **function of immediate availability *A(t)*** or for the non-functional (disable to operate) state by the function of immediate unavailability *U(t)*.

The functions *A(t)* and *U(t)* are inter-complementary, the sum of their values is in the certain time equal with 1 (the probability that the object will occur in the one state or in another one is equal with the certainty).

The function of the immediate availability *A(t)* expresses the probability that the object is in the state when it is able to perform the requested function in the given conditions and in given time providing that the requested outer conditions are ensured.

This indicator is not very often used in the practice because not the immediate level of object availability but the level of its availability detached to a certain time interval is usually the subject of the interest.

For the availability description the following indicator are used:

a) **Mean availability coefficient**, which expresses the mean value of the immediate availability in the certain time interval (t_1, t_2):

$$\overline{A}(t_1,t_2)=\frac{1}{t_2-t_1}\cdot\int_{t_1}^{t_2}A(t)\cdot \mathrm{d}t\,. \qquad (1)$$

b) **Asymptotic (stabilized) availability coefficient** represents the limit of the immediate function of the availability for t → ∞.

$$A=\lim_{t\to\infty}A(t)\,. \qquad (2)$$

Asymptotic availability coefficient *A* can be expressed by equation:

$$A=\frac{MTBF}{MTTR+MTBF}\,, \qquad (3)$$

where: *MTBF* – mean time between failures, *MTTR* – mean time to repair.

It expresses the probability that the object, which is in the stabilized operation regime ”operation – recovery”, will be in the arbitrary time in state of operation capability (apart from the planned time during which the usage of the object is not planned, for example the planned prevention repair).

c) **operation availability coefficient** expresses the ratio of the total operation time in the usable state and the total time including the downtimes (4).

$$A_o=\frac{MUT}{MUT+MDT}\,, \qquad (4)$$

where: *MUT*- mean uptime, *MDT*- mean downtime.

d) **achieved availability coefficient** is expressed with the help of the mean time between maintenance MTBM and the mean time of maintenance downtime $\overline{M}$:

$$A_A=\frac{MTBM}{MTBM+\overline{\overline{M}}}\,. \qquad (5)$$

In the technical practice the asymptotic availability coefficient is used for the stabilized recovery process. It is very often used on condition that:

- logistic, administration and technical delays are neglected,
- the distributions of the random variable for the faultlessness with parameter λ and the maintainability with parameter μ are exponential.

If the distribution of periods between the faults and periods to the recovery has exponential character, we can express the asymptotic availability coefficient of the object as (3):

$$A=\frac{\mu}{\lambda+\mu}\,, \qquad (6)$$

where: λ - intensity of faults, μ - intensity of repair.

The asymptotic availability coefficient, which characterises a certain stabilized availability level, which the object is successively approaching with the increased operation time, is the most suitable from the mentioned indicators for the complete description of the object availability. All other statistical models created on the base of stochastic principles always lead to non-constant availability function A_i, i.e. to the availability function dependent on the operation time *t*.

3. SIMULATION APPROACH TO AVAILABILITY MODELLING

The concept of deterministic availability models arises from the idea that time functions to the fault and time necessary for fault removal at the element failure Ei are same distributions of parameters probabilities like those, which appear in them.

They most often lead to exponential, eventually Weibull probabilities distribution. The time curves of fault rate and reparations, eventually other stochastic influences during the reliability of complicated systems ensuring in real operation (3), are not taken into account in these models.

In the real case the operation reliability, respectively its partial properties are connected with processes, which are necessary for the failure removal (control process, supplying system, repairing process, etc.). That's why also the model may have several states and distributions of random variables.

These facts can be expressed by simulation modelling.

We utilise the fact that the probabilities of the components time of fault occurrence and the time of fault removal are quantities with significantly stochastic character, which can appear in wide range of values.

The proposed solution can be expressed like this:

a) System $M_{k\Psi}$ is decomposed into subsystems or elements

$$M_{k1\Psi} = \{ m_1, \Psi_1, m_2, \Psi_2, ... m_k, \Psi_k, ... m_s, \Psi_s \}_\Psi$$

The partial systems are analysed separately and the results are utilised for the final valuation of the system.

b) Statistic rules of model subsystems (elements) can be described by:
   - probability distribution of fault occurrence intervals,
   - probability distribution of active repair time,
   - eventually probability distribution of other downtimes.

c) We determine, which states are important for the system analysis and which we want to express by the simulation. Some states can be united.

d) We determine the outputs, which we can statistically elaborate and visualize by means of the graph.

e) We construct the computation simulation model and realise the experiments, which are then evaluated.

Tab. 3.1 Maintenance periods, which can represent model states with its own probability distribution of random variable

Maintenance time						
Prevention maintenance time		Maintenance time after failure				
Logistic downtime delay	Active Prevention maintenance time	Active maintenance time after failure				Logistic downtime delay
		Technical downtime delay	Fault localisation time	Active repair time	Checking time	
	Active maintenance time					

4. MODELLING OF SYSTEM AVAILABILITY BY DISCRETE SIMULATION WITH VARIABLE TIME STEP

The formation of discrete simulation model with variable time step and the realisation of simulation experiment predict the execution of the following activities.

1. Input of the starting conditions and specification of variables values in the initial simulation time TIME = 0, input of the simulation period TEND. Elements state $L_{(i)} = 0$, system state S = 0.
2. Generating of intervals of failure occurrence of single elements of the system from the probabilities distributions of periods between the failures $x_{(i)}$ (i = 1,2,......N).
3. Sequencing of faults occurrence and choosing of the first event by searching the minimum from values $x_{(i)}$ for i = 1, 2,..., N.
4. Element state $L_{(i)}$ change to $L_{(j)} = 1$ the element is defective and the repair is realised. System state S change to S = 1 system is non-functional.
5. Shift the time axis by the interval of the first fault CAS=CAS+ $x_{(i)}$.
6. Generate element maintenance realisation period from the probability distribution of maintenance time $y_{(i)}$ (i = 1,2,......, K).
7. Shift the time axis by the maintenance time CAS=CAS+ $y_{(i)}$. Element state $L_{(i)}$ change to $L_{(j)} = 0$ the element is serviceable. System state S change to S = 0 system is non-functional.
8. Generate new interval $x_{(i)}$ of element *I* failure, which was returned to the serviceable state.
9. Calculations of the elements and system availability.
10. Condition testing of the simulation process finishing, if the value of the simulated time reaches the predefined value TEND, otherwise repeat points 3-10
11. Collect and elaborate by statistical methods the data of input and output quantities.
12. The results outputs on the display and printer. End of the simulation experiment.

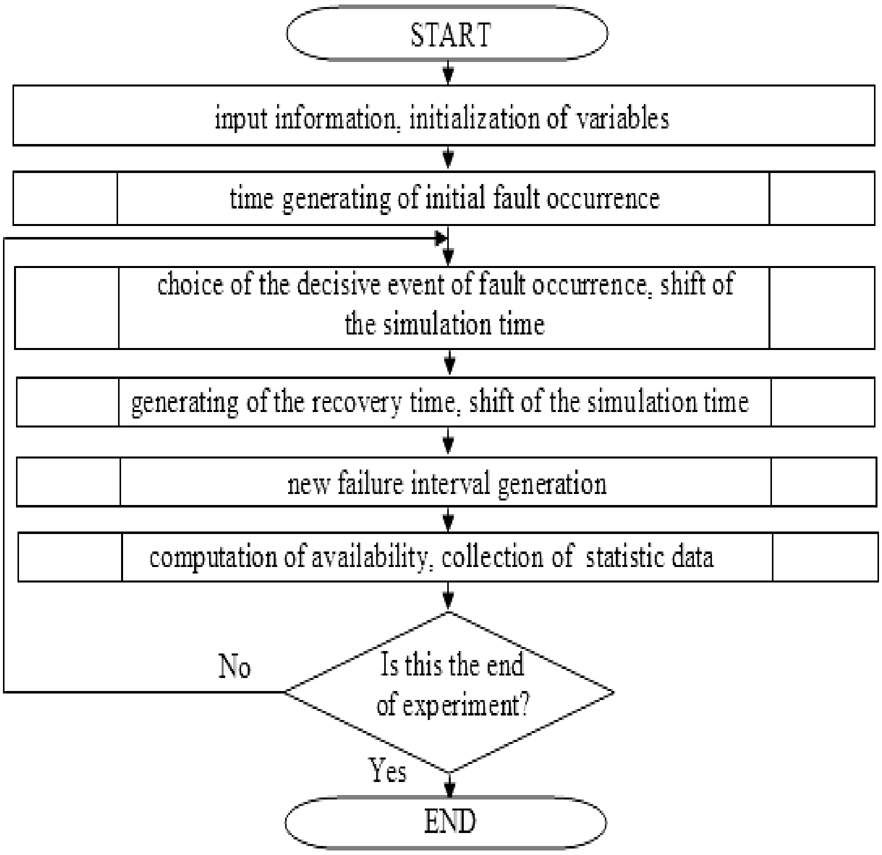

START
input information, initialization of variables
time generating of initial fault occurrence
choice of the decisive event of fault occurrence, shift of the simulation time
generating of the recovery time, shift of the simulation time
new failure interval generation
computation of availability, collection of statistic data
Is this the end of experiment?
No
Yes
END

Fig.4.1 Flow diagram

Simulation model is constructed for easy observation of the dynamic maintenance states of the elements and system from the simulation results

The intervals of maintenance are limited by red rectangles *fig.4.2,f ig.4.3*.

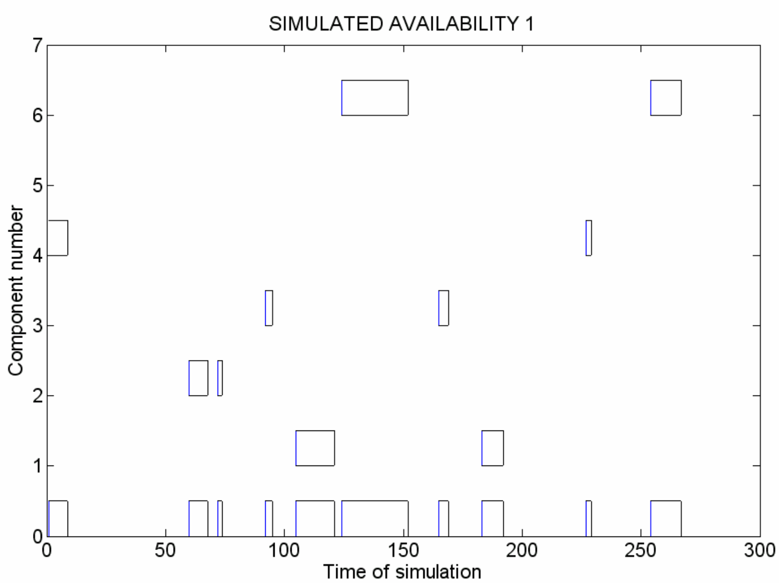
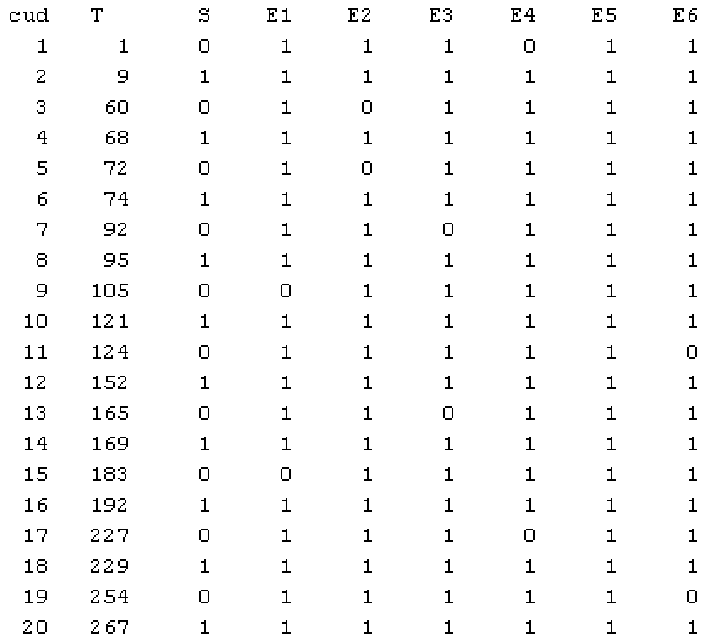


Fig. 4.2 The process of experiment with the simulation duration of 267 hours with indication of elements and system maintenance time



```
cud   T       S     E1    E2    E3    E4    E5    E6
  1     1     0     1     1     1     0     1     1
  2     9     1     1     1     1     1     1     1
  3    60     0     1     0     1     1     1     1
  4    68     1     1     1     1     1     1     1
  5    72     0     1     0     1     1     1     1
  6    74     1     1     1     1     1     1     1
  7    92     0     1     1     0     1     1     1
  8    95     1     1     1     1     1     1     1
  9   105     0     0     1     1     1     1     1
 10   121     1     1     1     1     1     1     1
 11   124     0     1     1     1     1     1     0
 12   152     1     1     1     1     1     1     1
 13   165     0     1     1     0     1     1     1
 14   169     1     1     1     1     1     1     1
 15   183     0     0     1     1     1     1     1
 16   192     1     1     1     1     1     1     1
 17   227     0     1     1     1     0     1     1
 18   229     1     1     1     1     1     1     1
 19   254     0     1     1     1     1     1     0
 20   267     1     1     1     1     1     1     1
```

Fig. 4.3 Process of the experiment with output statement of events, time and states of elements and system

In the case of simulations with longer simulation time the intervals with low predicative value are indicated *fig. 4.4* and that’s why we further evaluate the graphs shown below.

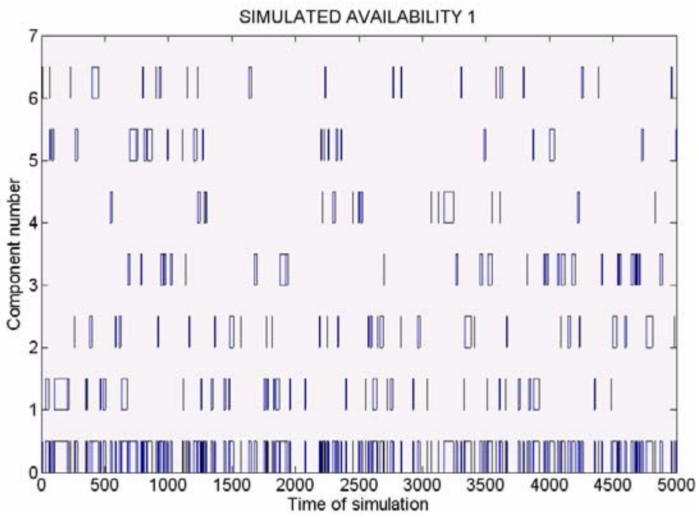


Fig. 4.4 The process of experiment with simulation duration of 5000 hours

It is possible to follow graphically the value of the asymptotic availability coefficient in dependence on time *fig.4.5*.

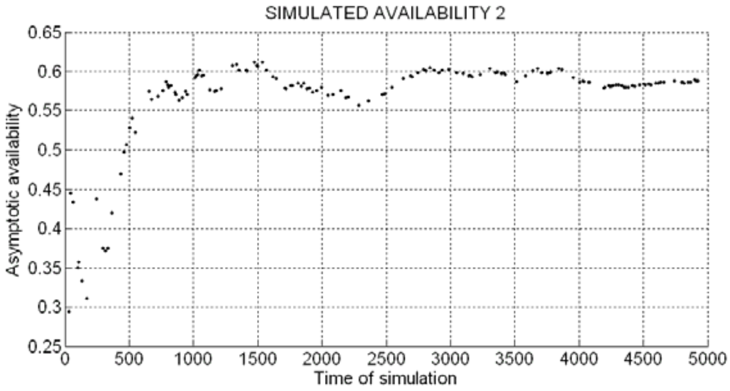


Fig.4.5 Rise and stabilisation of asymptotic availability coefficient

The rise of the asymptotic availability coefficient value is very interesting. In comparison to the published statements (3) the rise time to the stabilized state is quite long. After using same input values, experiments give results with same dissipation, so is necessary to realize more number of the simulations.

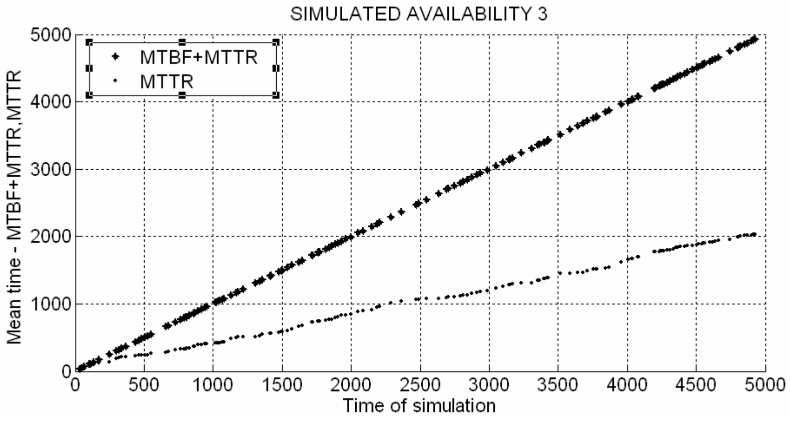


Fig.4.6 Participation of MTBF and MTTR during system observation time

Graph in *fig.4.6* shows the time period of the system use MTBF+MTTR, quotient of the no-failure state and quotient of the maintenance MTTR realisation.

The limitation of the areas under the curves evaluated from the bottom to the top shows the quotient of the maintenance time and the quotient of the system's serviceable state from the total time of study. The quotient of the maintenance lower, the reliability is higher.

5. RESULTS AND CONCLUSIONS FROM SIMULATION EXPERIMENTS

The simulation experiments are made for the serial mechanical system with six elements. The mean time between the faults and the mean time to repair is of exponential probability distribution.

Experiments are realized with parameter values depicted in tab. 1.

- Mean time between the failures is expressed by exponential distribution with parameter MTBF,
- Mean time to repair is expressed by the exponential distribution with parameter MTTR.

In experiments 1 - 4 the mean time to repair parameter was decreasing.

Tab. 5.1 Input parameters of simulation experiments

Parameter	Element of the system					
In hours	E1	E2	E3	E4	E5	E6
MTBF	100	140	110	160	120	150
MTTR 1	15	12	16	11	14	10
MTTR 2	7.5	6	8	5.5	7	5
MTTR 3	3.75	3	4	2.75	3.5	2.5
MTTR 4	1.5	1.2	1.6	1.1	1.4	1

Results from the first experiment are shown in *fig. 4.5, fig. 4.6*. Quotient of the maintenance time is high and asymptotic availability coefficient doesn't reach a value 0.6.

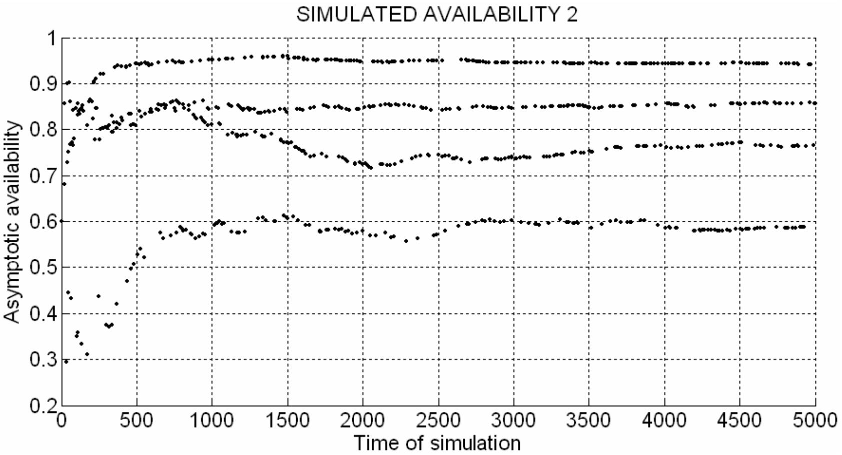


Fig. 5.1 Asymptotic availability coefficient developments

Other experiments shows rising of asymptotic availability coefficient up to value 0.95, while lowering MTTR of elements *fig.5.1.*

From the shown experiment results the confirmation of the mathematical expression base of availability coefficients is clear that the reached availability level is determined by two components – by faultlessness and maintenance.

The relationship between faultlessness, maintenance and availability is shown in table 5.2.

Tab. 5.2 Relationship between faultlessness, maintenance and availability

Faultlessness	**Maintenance**	**Availability**
Expressed by parameter MTBF	Expressed by parameter MTTR	Expressed by parameter A
MTBF - Constant	MTTR - Decreases	A - Increases
MTBF - Constant	MTTR - Increases	A - Decreases
MTBF – Increases	MTTR - Constant	A - Increases
MTBF - Decreases	MTTR - Constant	A - Decreases

6. CONCLUSION

It expresses the possibilities of availability increase of constructed and operating devices. It is possible to increase the availability practically only by shortening of intervals of the device maintenance components.

Simulation modelling of availability prediction is advantageous for it's possibility to watch dynamic process using graphical outputs. This outputs gives more illustrative image about random processes. Output values can be easily compared. Parameters of the system can be tunable according this comparison, so the system response will be appropriate.

References

4. Holub, R., Vintr, Z.: Aplikované techniky spolehlivosti. Část 1: Specifikace požadavků na spolehlivost. [Skripta] Brno: Vojenská akademie 2002.
5. Holub, R., Vintr, Z.: Základy spolehlivosti. [Skripta] Brno: Vojenská akademie 2002.
6. Vintr, Z.: Možnosti predikce pohotovosti technických systémů. In: TD 2004 – DIAGON 2004 – Sborník přednášek 27. mezinárodní konference. Zlín: Academia centrum UTB, 2004, str. 85 – 90. ISBN 80-7318-195-6.
7. US Army TM 5-698-3 Reliability primer for command, control, communications, computer, intelligence, surveillance, and reconnaisance (C4ISR) facilities, 2003
8. Crowe, D., Feinberg, A.: Design for reliability. CRC Press. New York. 2001
9. Chovanec, A.: Modelovanie a simulácia diskrétnych stochastických procesov, TNU AD v Trenčíne v spolupráci s vydavateľstvom GERŠI. Trenčín 2004, ISBN 80-8075-009-2.
10. Leitner, B. : Optimum design of Track Maintenance machine Frames by Matlab. In: Zborník medzinárodnej konferencie "TRANSPORT 2003, VTU Todora Kableshkova, Sofia 2003, str. 157 - 160.
11. Balog, J.: Diagnostikovanie a prognózovanie stavov pri údržbách strojov. SM3 – 1, Prípadová studia, Slovenská poľnohospodárska univerzita v Nitre. Nitra 2002.

This article was elaborated at the tasks realisation of granted project VEGA 1/2094/05

COST OPTIMIZATION FOR REALISATION OF MAINTENANCE COST

Alexej Chovanec
●
Faculty of Special Technology / Alexander Dubcek University in Trencin
Studentska 1, 911 50 Trencin, Slovak republic
e-mail: chovanec@tnuni.sk

Anton Ambrozy
●
Faculty of Special Technology / Alexander Dubcek University in Trencin
Studentska 1, 911 50 Trencin, Slovak republic
e-mail: ambrozy@tnuni.sk

Abstract: This paper deals with optimal control interval determination using minimization the financial costs. It clears conceptual, mathematical and simulation model of the problem solution. It enumerates and evaluates results of the simulation.

Keywords: sampling size, fault time, interval between failure, normal distribution, financial costs, simulation experiment, optimisation process, probability density, optimal maintenance interval

1. INTRODUCTION

In the stage of the technique usage some precautions of dependability support are realised at the technique user, expressed in the Complex care program, which determines:

- strategy of approaches to the realisation of the complex care control,
- methods of the activities in the realisation precautions,
- organisation work forms of executive workers,
- types and technological limitations of complex care tasks.

The non-operating state and maintenance costs minimizing, but also optimising of:

- ➢ dependability,
- ➢ safety,
- ➢ production quality,
- ➢ profit.

The optimal determination of time periods for maintenance making is important task of the maintenance systems control. The maintenance interval represents the maintenance period between single maintenance levels (the period for realisation of the maintenance of respective level).

At the optimisation of maintenance intervals we need the verified and usable data, which can be acquired from one or more sources, for example:

- Previous knowledge from similar systems, if the maintenance actions can be suitable also for the new designed product. The maintenance programs of similar products can provide after the critical analysis suitable base.
- The data from all kinds of tests at the producer can provide information about the efficiency, usefulness and the effectiveness of designed maintenance programs of the new product and its components.
- If no previous verified knowledge about the fault rate of other systems exist or if the previous and new systems are not similar enough, the maintenance interval can be determined by estimation (expert method). The experiences of producer, user specialists are utilised and it is progressing in accordance with the knowledge in the dependability, operation, operation conditions etc.

- If verified data about the failure rates of single system components exist, it is possible to use calculations in accordance with the restoration theory and rules of complicate systems dependability for determination of the maintenance intervals.

2. COST OPTIMISING PRINCIPLE

The optimal determination of time periods for maintenance making is important task of the maintenance systems control. The maintenance interval represents the maintenance period between single maintenance levels (the period for realisation of the maintenance of respective level).

At the optimisation of maintenance intervals we need the verified and usable data, which can be acquired from one or more sources, for example:

- Previous knowledge from similar systems, if the maintenance actions can be suitable also for the new designed product. The maintenance programs of similar products can provide after the critical analysis suitable base.
- The data from all kinds of tests at the producer can provide information about the efficiency, usefulness and the effectiveness of designed maintenance programs of the new product and its components.
- If no previous verified knowledge about the fault rate of other systems exist or if the previous and new systems are not similar enough, the maintenance interval can be determined by estimation (expert method). The experiences of producer, user specialists are utilised and it is progressing in accordance with the knowledge in the dependability, operation, operation conditions etc.
- If verified data about the failure rates of single system components exist, it is possible to use calculations in accordance with the restoration theory and rules of complicate systems dependability for determination of the maintenance intervals.

Determination of maintenance action interval thus so that the total unit costs of the object's life cycle should be minimized is the aim of the optimisation.

The total unit costs N_c at vanishing the costs components, which are not suggestible by the maintenance actions, consist of three components:

1. Supplying costs of the object (N_o), which are constant (independent on the maintenance interval time),
2. Costs of preventive maintenance (N_p),
3. Costs of the corrective maintenance (N_n) of the object (their amount depend on the maintenance interval)

The expression of total unit costs:

$$Nc_{(t)} = 1/t\ (N_o + N_{p(t)} + N_{n(t)}\ ) \qquad (1)$$

where: t – is the maintenance interval, which we are looking for.

The expression is rearranged into the form, where each term on the right side expresses the relative costs:

$$Nc_{(t)} = N_o\ /t+ N_{p(t)}\ /t+ N_{n(t\ )}/t) \qquad (2)$$

The relative costs for the preventive maintenance are constant in time and that's why we directly replace them by symbol N_p :

$$Nc_{(t)} = N_o\,/t+ N_p+ N_{n(t)}/t) \qquad (3)$$

The optimising task lies in finding such a maintenance interval t, which ensures that the total unit costs expressed by the previous equation will be minimal.

The suitable (convex) form of the curve $Nc_{(t)}$ and sufficiently sharp and expressive minimum in point $D(t_{opt};\ Nc_{(t)\ min})$, which enables relatively exact identification of the optimum point position, are the conditions of the application successfulness of the optimising method.

If the function minimum $Nc_{(t)}$ is flat and inexpressive, the determination of the optimal maintenance interval is more difficult, but it enables easier combination of maintenance actions without negative impact in economical field.

All variants of economical optimising of the technique life cycle or maintenance intervals are based on a similar principle.

The suitable (convex) form of the curve $Nc_{(t)}$ and sufficiently sharp and expressive minimum in point $D(t_{opt}; Nc_{(t)\ min})$, which enables relatively exact identification of the optimum point position, are the conditions of the application successfulness of the optimising method.

3. OPTIMISING MODEL OF PERIODIC CONTROLS INTERVAL

The device Z is formed by elements $P_1, P_2, \ldots\ldots P_K$, which are in operable or failure state.

The diagnostic operations with costs per hour C_1 are made for the detection of failure state occurrence in periodical times $t_{K1}, t_{K2}, \ldots t_{KN}$. During the diagnostic inspection all or several elements can be diagnosed.

In the case of failure occurrence until the closest control of the operability the device causes with its operation the loss per hour C_2.

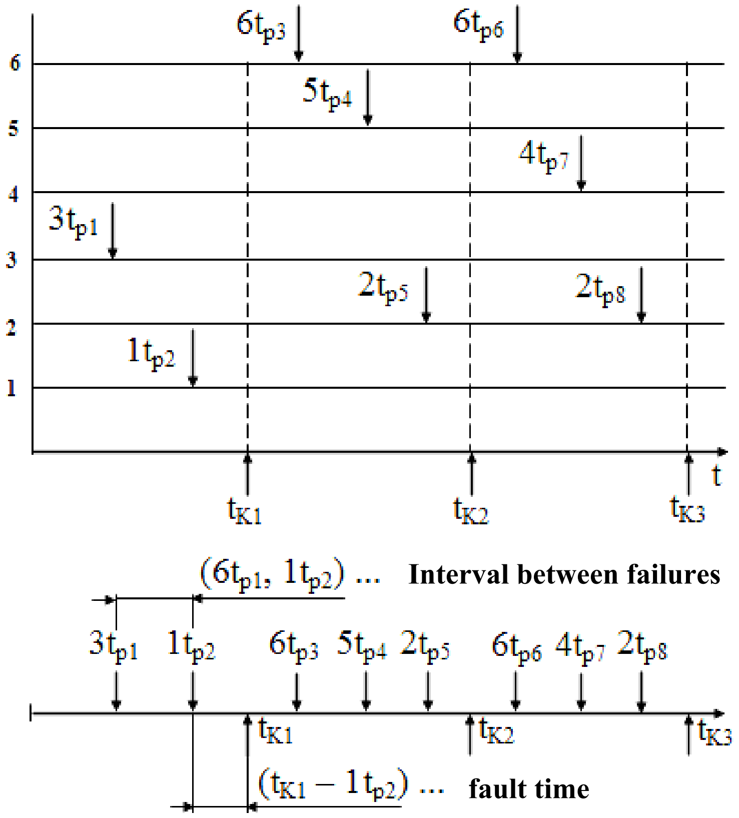

$6t_{p3}$
$6t_{p6}$
$5t_{p4}$
$4t_{p7}$
$3t_{p1}$
$2t_{p5}$
$2t_{p8}$
$1t_{p2}$
6
5
4
3
2
1
t
t_{K1}
t_{K2}
t_{K3}
$(6t_{p1}, 1t_{p2})$ ... Interval between failures
$3t_{p1}$ $1t_{p2}$ $6t_{p3}$ $5t_{p4}$ $2t_{p5}$ $6t_{p6}$ $4t_{p7}$ $2t_{p8}$
t_{K1}
t_{K2}
t_{K3}
$(t_{K1} - 1t_{p2})$ ... fault time

Fig. 3.1 Scheme of failures moments and diagnostic controls
t_{Ki} – control time, I t_{pi} – time of failure occurrence on the I – element

Operation times to the failure occurrence of single elements have the probability distribution of continuous type with distribution function:

$$F_{(t)} = P\,(\,T \leq t\,) \qquad (4)$$

and with probability density $f_{(t)} = \dfrac{d}{dt}\,F\,(t).$

Due to the failures, which appear in times $t_{P1}, t_{P2}, \ldots\ldots t_{PM}$, due to the work in unsatisfactory conditions or due to non operating state, losses rise up to the closest diagnostic control.

$$N_P = \Sigma\, C_2\, .\, (\,t_{Ki} - t_{Pi}\,) \qquad (5)$$

If no failure occurs, losses rise from the unavailing usage of the diagnostic device C_1.

The sum of costs of planned diagnostic controls and losses from the device operation in the fault state

$$N_C = \sum C_1 + C_2 \, . \, (t_{Ki} - t_{Pj}) \qquad (6)$$

must be minimal and so also the mean costs of one cycle of periodic control.

4. SIMULATION MODEL OF MAINTENANCE INTERVALS OPTIMIZATION

In the case of knowledge of the mean time values of failures occurrence, deterministic calculation for various control intervals is made according to previous relations and the minimal value of test function is being searched.

If we have an elaborated statistical set of data and information about failures in the form of type and distribution parameters of times between failures in the operation units, we can utilise the simulation approach.

Flow diagram *fig. 4.4.* shows the process described below.

1. Insertion of input data (elements number, unit controlling costs C_1, unit costs - losses C_2, time of simulation end in operation units, distribution parameters of periods between failures of elements, set-up of initial values of variables).
2. We generate the size of intervals between failures. We determine the smallest and the biggest interval between failures and the step of tested cycle.
3. We realise the cycle of optimisation experiment and change the step (increase or decrease of the control interval).
4. We determine the number of controls. We calculate the prevention controlling activity costs Np. We calculate the operation times in fault state and the loss size Nn. We calculate the total costs Nc.
5. We process the output data.
6. We verify the conditions of experiment continuation.
7. We diagrammatise the size of costs depending on the size of control intervals.

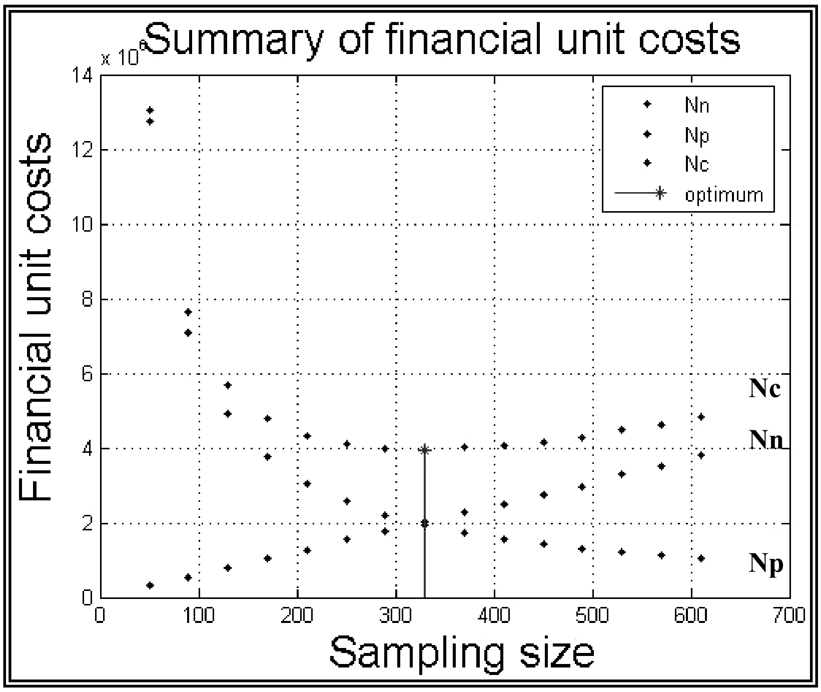


Fig. 4.1 Diagram *of optimal size of control interval for the simulation experiment 1 (input parameters tab.5.1)*

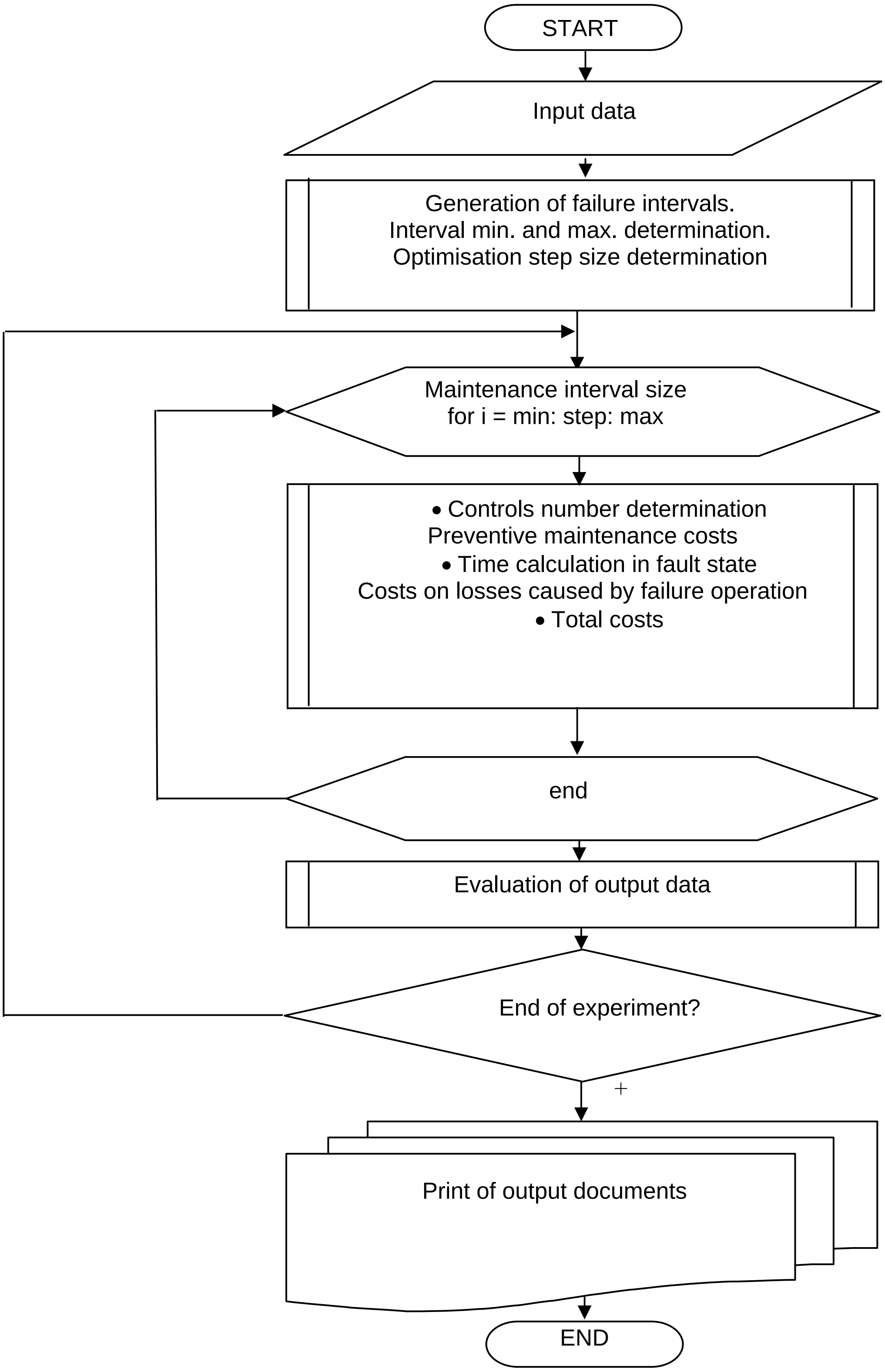

START
Input data
Generation of failure intervals.
Interval min. and max. determination.
Optimisation step size determination
Maintenance interval size
for i = min: step: max
• Controls number determination
Preventive maintenance costs
• Time calculation in fault state
Costs on losses caused by failure operation
• Total costs
end
Evaluation of output data
End of experiment?
+
Print of output documents
END

Fig. 4.2 Flow diagram

5. REALISATION AND EVALUATION OF RESULTS OF SIMULATION EXPERIMENTS

The simulation experiments enable to follow the changes of the output parameters values depending on the change of input parameters.

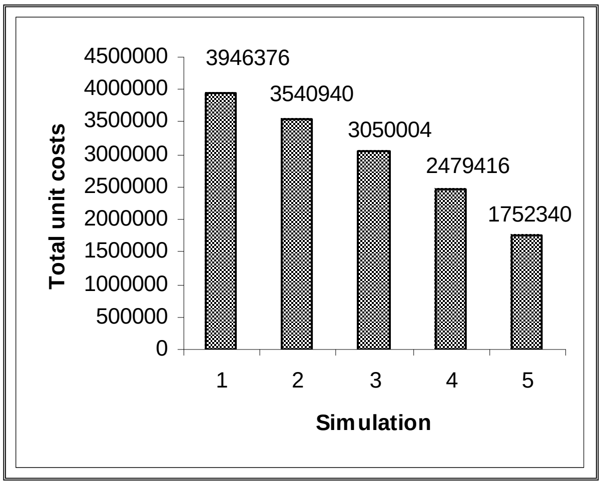
Which is the change of total costs and size of control interval in dependence on the change of unit costs values on control operations at constant costs of losses from the failure operation?

Tab. 4.1 represents the input and output parameters of simulations. Each experiment provides the determination visualisation of the control interval optimal size in the studied range of possible control interval. The optimisation process and its range for the simulation experiment 1 are shown in *fig. 4.3.*

Tab. 5.1 Input and output parameters of simulations

INPUT:					
Generation of failure times: Normal distribution:	MU=1200 SIGMA=200				
Simulation:	**1.**	**2.**	**3.**	**4.**	**5.**
Control operation unit costs	5000	4000	3000	2000	1000
Losses caused by failure operation	120				

OUTPUT:					
Simulation:	**1.**	**2.**	**3.**	**4.**	**5.**
Interval between controls	330	290	250	210	130
Total unit costs	3946376	3540940	3050004	2479416	1752340



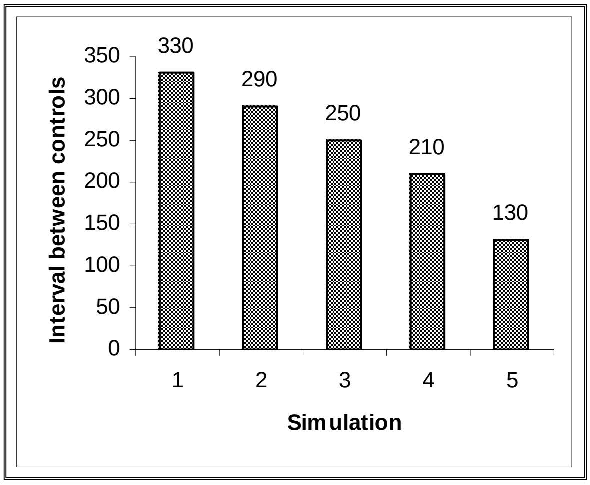


Fig. 5.1 Reduction of total costs and size of optimal control interval at decreasing of control operations costs according to tab. 5.1

Which is the change of total costs and control interval size depending on the change of unit costs values occurring due to the failure operation at constant costs of control operations?

Input and output parameters, optimisation process, increase of total costs and reduction of optimal control interval size at costs increase due to the failure operation is shown in tables and graphs below.

Tab. 5.2 Input and output parameters of simulations

INPUT:					
Generation of failure times: Normal distribution:	MU=1200				
	SIGMA=200				
Simulation:	**1.**	**2.**	**3.**	**4.**	**5.**
Losses caused by failure operation	120	220	320	420	520
Control operation unit costs	5000				

OUTPUT:					
Simulation:	**1.**	**2.**	**3.**	**4.**	**5.**
Interval between controls	330	250	210	170	160
Total unit costs	3936012	5327617	6408536	7411640	8174268

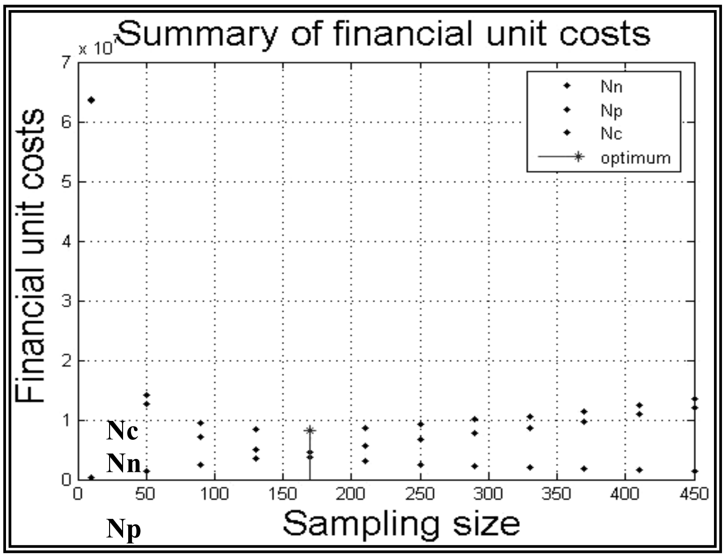


Fig. 5.2 Diagram of optimal size of control interval for simulation experiment 5 (input parameters tab. 5.2)

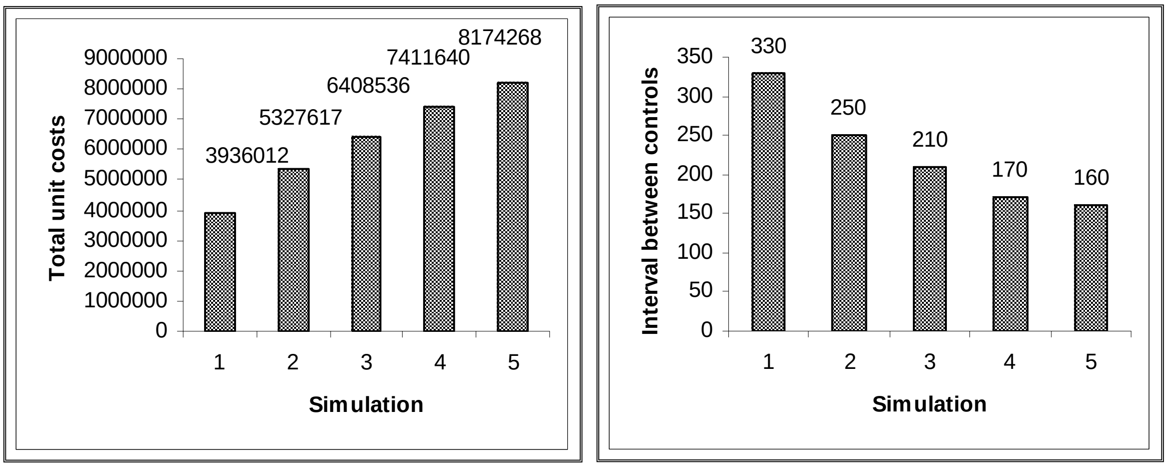


Fig. 5.3 Increase of total costs and the reduction of optimal control interval size at increase of costs due to failure operation according to tab. 5.2

The following result from the results of simulation experiments:

1. The total costs decrease with the decreasing values of unit control costs.
2. The size of optimal control interval decreases with the decreasing value of the unit control costs and the number of control activities intervals increases.
3. The total costs will increase and the size of the optimal control interval will decrease with the increasing value of the unit costs from losses caused by the failure activity.

6. CONCLUSION

The simulation model optimises the interval of maintenance action so that the total costs (on control and costs caused by failure) should be minimised.

The processing of simulation model arises from deterministic process of resolution of the mentioned problem. The simulation model was verified by processes of classical computer methods.

The algorithm representing the simulation model is drawn so that it could be easily modifiable. It enables to remove the lacks of classical computer methods and to get the required results in short time.

The obtained results can be applied to the determination of the number of control activities intervals and the size of control activities interval. They enable to gain an idea about the number of expended financial resources for the control operations, about the costs of the control operations and total financial costs.

The simulation model creates the bases for further development of problems resolution possibilities being connected with the determination of the optimal control interval size and minimising of total costs. It expands the range of application utilisation of simulation modelling for solution of problems related to the operation of specific technical systems. Generally it can be the mobile technique, elaboration device or constructional groups of technical devices, which during their operation get into the failure state.

The simulation model provides general graphical outputs of costs changes in dependence on the size of control interval and the number of control intervals. It can be used for the graphical and didactical support of the explanation of optimisation problems of maintenance intervals.

REFERENCES

1. HOLUB, R. – VINTR, Z.: Aplikovné techniky spolehlivosti. Část 1: Specifikace požadavků na spolehlivost. [Skripta]: Vojenská akademie 2002.
2. FURCH, J.: Možné přístupy pro stanovení optimálního intervalu doby životnosti vozidla. In: Sborník mezinárodní konference“ OPOTŘEBENÍ SPOLEHLIVOST DIAGNOSTIKA 2005, Univerzita obrany, Brno 2005, Česká republika, str. 49 – 54, ISBN 80-7231-026-7
3. MENČÍK, J,: Optimalizace návrhu technických objektů z hlediska nákladů. In: Sborník mezinárodní konference“ OPOTŘEBENÍ SPOLEHLIVOST DIAGNOSTIKA 2005, Univerzita obrany, Brno 2005, Česká republika, str. 125 – 130, ISBN 80-7231-026-7

This article was elaborated at the tasks realisation of granted project VEGA 1/2094/05

SIMULATION OF FTA IN SIMULINK

Alexej Chovanec
●
Faculty of Special Technology / Alexander Dubcek University in Trencin
Studentska 1, 911 50 Trencin, Slovak republic
E-mail: chovanec@tnuni.sk

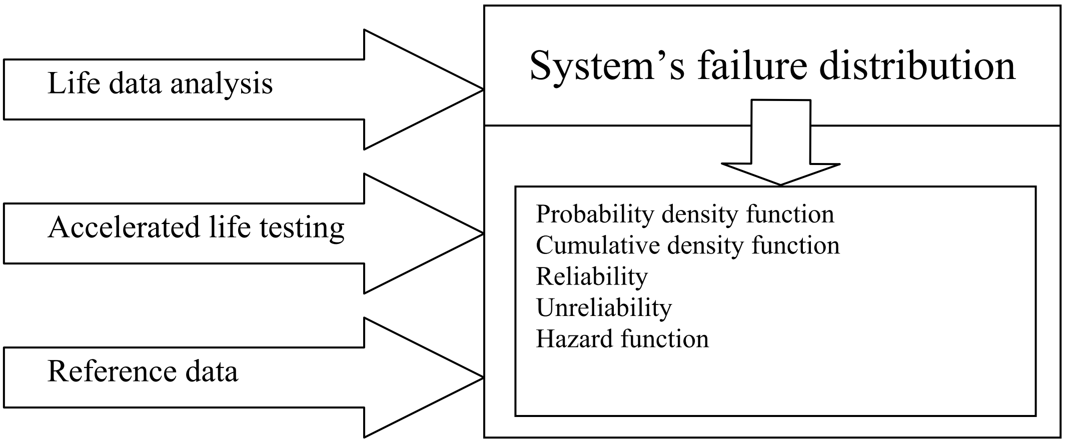
Jozef Bucha
●
Faculty of Special Technology / Alexander Dubcek University in Trencin
Studentska 1, 911 50 Trencin, Slovak republic
E-mail: bucha@tnuni.sk

Abstract: This paper deals with possibility of simulation of reliability block diagrams, failure trees analysis as a time dependent analysis using Matlab/Simulnk.

Keywords: failure distribution, Simulink, reliability, probability density function, cumulative density function, reliability block diagrams, failure tree analysis

1. INTRODUCTION

In life data analysis and accelerated life testing data analysis, the objective is to obtain a life distribution that describes the times-to-failure of a component, subassembly, assembly or system. The analysis to determine the life distribution is based on the time of successful operation or time-to-failure data of the item, either under use conditions or from accelerated life tests.



The main objective of system reliability is the construction of a model (life distribution) that represents the times-to-failure of the entire system based on the life distributions of the system's elements. These elements can be components assemblies, sub-systems etc. There are many specific reasons for looking at component data to estimate the overall system reliability. One of the most important is that in many situations it is easier and less expensive to test single elements rather than entire systems, also properties of failure distributions of single elements can easily tuned and then changes of overall system distribution can be compared.

In general, most problems in reliability engineering deal with quantitative measures, such as the time-to-failure of a product, or qualitative measures, such as whether a product is defective or non-defective. We can then use a *random variable X* to denote these possible measures. In the case of times-to-failure, our random variable *X* is the time-to-failure of the product and can take on an infinite number of possible values in a range from 0 to infinity. Product can be found failed at any time after time 0, thus *X* can take on any value in this range. In this case, our random variable *X* is said to be a *continuous random variable*.

The probability density function (*pdf*) and cumulative distribution function (*cdf*) are two of the most important statistical functions in reliability and are very closely related. When these functions are known,, other reliability such as reliability *R(t)*, unreliability *Q(t)*, hazard function *h(t)* can be computed and obtained.

On the Figures 1.1-2 are depicted examples of *pdf* and *cdf* of the normal distribution.

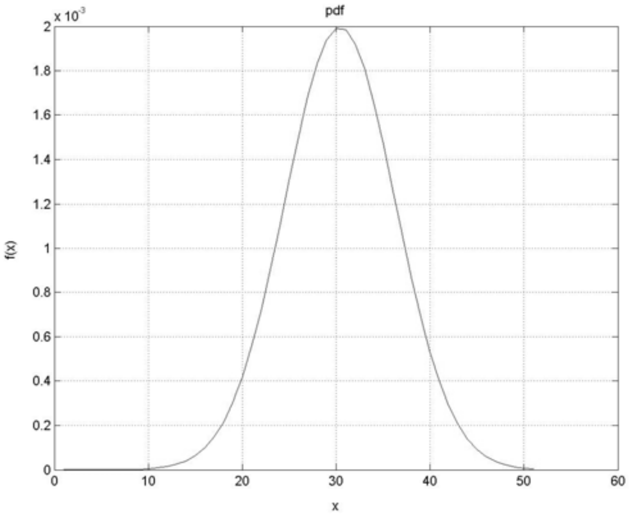


Fig. 1.1 Example of probability distribution function

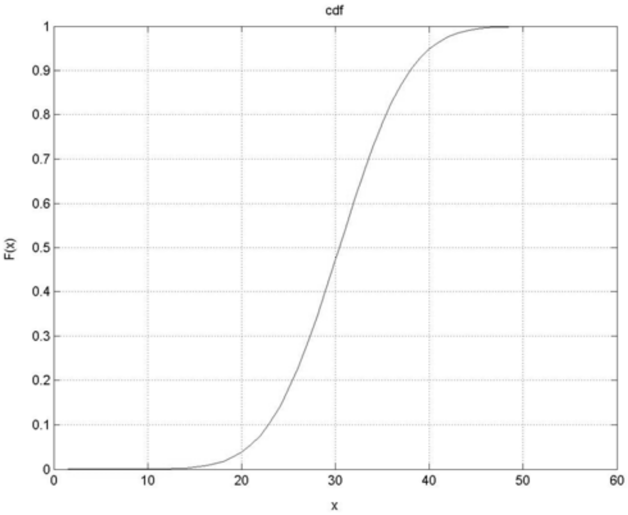


Fig. 1.1 Example of cumulative distribution function

The mathematical relationship between the *pdf* and *cdf* is given by eq. (1) and eq. (2)

$$F(x) = \int_{0}^{s} f(s)ds \tag{1}$$

Conversely

$$f(x) = \frac{d(F(x))}{dx} \quad (2)$$

The *cdf* is the area under the probability density function up to a value of *x*. The total area under the *pdf* (Figure 1.1) is always equal to 1, or mathematically:

$$\int_0^\infty f(x)dx = 1 \quad (3)$$

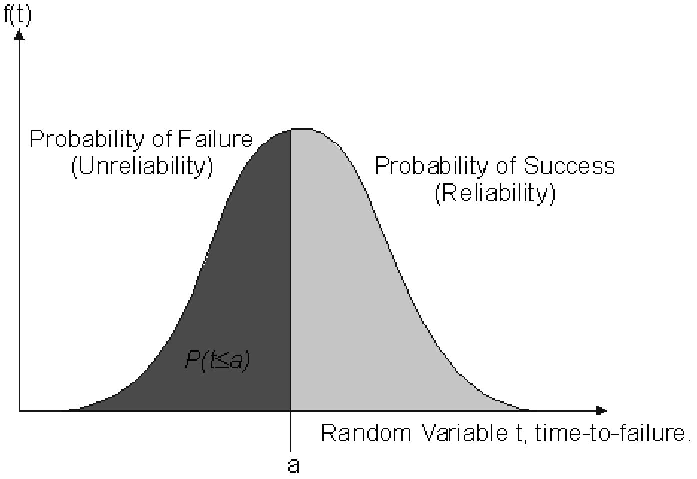


Fig. 1.3 Reliability and unreliability as areas under pdf

Other function as reliability *R(t)* and unreliability *Q(t)*can be computed according eq. (4-6)

$$Q(t) = F(t) = \int_0^t f(s)ds \quad (4)$$

$$Q(t) + R(t) = 1 \quad (5)$$

$$R(t) = \int_t^\infty f(s)ds \quad (6)$$

Other important function is failure rate also known as a hazard function. This function enables the determination of the number of failures per time. This function can be computed according eq. (7)

$$\lambda(t) = h(t) = \frac{f(t)}{R(t)} \quad (7)$$

2. APPROACHES OF SYSTEM RELIABILITY

In theory and in praxis exists two basic approaches (categories of approaches):

- Analytical calculations
    1. Static analytical calculations
    2. Time-dependent calculations
- Simulation calculations

Two types of analytical calculations can be performed using RBD or FTA: static reliability calculations and time-dependent reliability calculations. Systems can contain static blocks, time-dependent blocks or a mixture of the two.

Static analytical calculations are performed on RBD or failure trees that contain static blocks. A static block can be interpreted either as a block with a reliability value that is known only at a given time (but the block's entire distribution is unknown) or as a block with a probability of success that is constant with time. Static calculations can only be performed in the analytical mode and not in the simulation calculations.

Time-dependent analysis looks at reliability as a function of time. That is, a known failure distribution is assigned to each component. The time scale can be any quantifiable time measure, such as years, months, hours, minutes or seconds, and also units that are not directly related to time.

If one includes information on the repair and maintenance characteristics of the components and resources available in the system, other information can also be analyzed/obtained, such as i.e. system availabilty, maintability etc. This can be accomplished through discrete event simulation.

In simulation, random failure times from each component's failure distribution are generated. These failure times are then combined in accordance with the way the components are reliability-wise arranged within the system. The overall results are analyzed in order to determine the behavior of the entire system.

3. FAULT TREE ANALYSIS, RELIABILITY BLOCK DIAGRAMS

Block diagrams are widely used in engineering in many different forms. Fault trees and reliability block diagrams are both symbolic analytical logic techniques that can be applied to analyze system reliability and related characteristics. They can also be used to describe the interrelation between the components and to define the system.

When blocks are connected with direction lines, that represent the reliability relationship between these blocks, it's referred as reliability block diagram (RBD). Example of RBD is depicted on fig. 3.1.

A fault tree diagram follows a top-down structure and represents a graphical model of the pathways within a system that can lead to a foreseeable, undesirable loss event (or a failure). The pathways interconnect contributory events and conditions using standard logic symbols (AND, OR, etc.). Fault tree diagrams consist of gates and events connected with lines. Example of RBD is depicted on fig. 3.2.

The most fundamental difference between fault tree diagrams and reliability block diagrams is that you work in the "success space" in an RBD while you work in the "failure space" in a fault tree. In other words, the RBD looks at success combinations while the fault tree looks at failure combinations. In addition, fault trees have traditionally been used to analyze fixed probabilities (*i.e.* each event that comprises the tree has a fixed probability of occurring) while RBDs may include time-varying distributions for the success (reliability equation) and other properties, such as repair/restoration distributions. In general (and with some specific exceptions), a fault tree can be easily converted to an RBD. However, it is generally more difficult to convert an RBD into a fault tree, especially if one allows for highly complex configurations.On fig. 3.2 is converted RBD from fig. 3.1

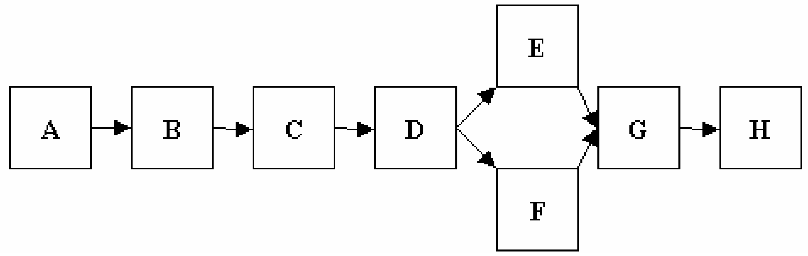


Fig. 3.1 Example of Reliability block diagram

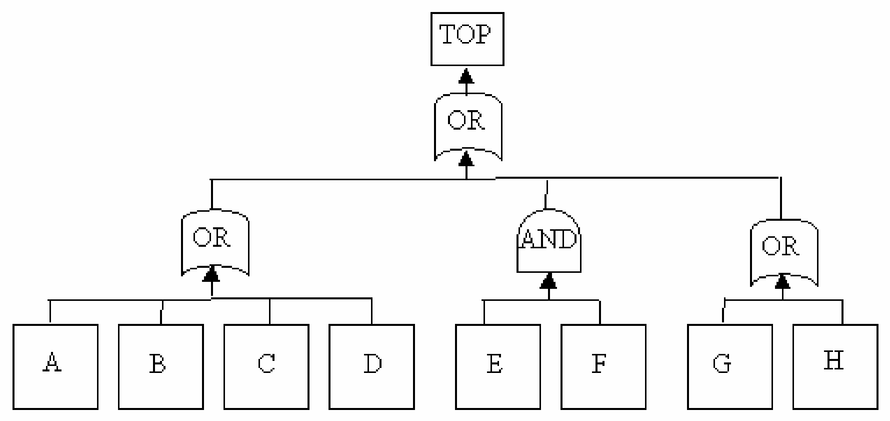


Fig. 3.2 Example of Failure tree

4. SIMULINK MODEL OF FAILURE TREE

In computer programming language Malab/Simulink, were constructed blocks for distributions (WEIBULL, NORMAL, EXPONENTIAL) and also blocks for AND gate and OR gate. Distributions contains all important outputs for computation of system reliability according failure tree or RBD diagram. Distributions outputs are: *f(t)* (*PDF*), *F(t) CDF, R(t), Q(t)* and *h(t).*

AND gate was made according fig.4.1, OR gate was made according fig. 4.2.

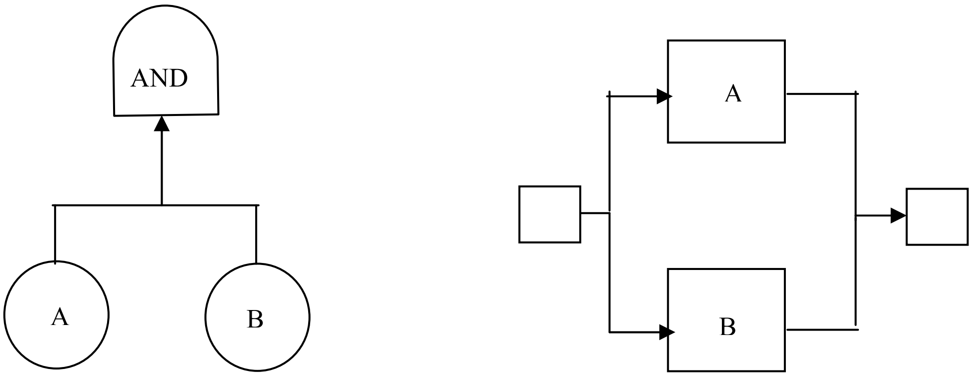


Fig. 4.1 FTA and RBD representation of parallel connection

The reliability equation for either configuration depicted on fig. 4.1

$$R_S = R_A + R_B - R_A R_B \quad (8)$$

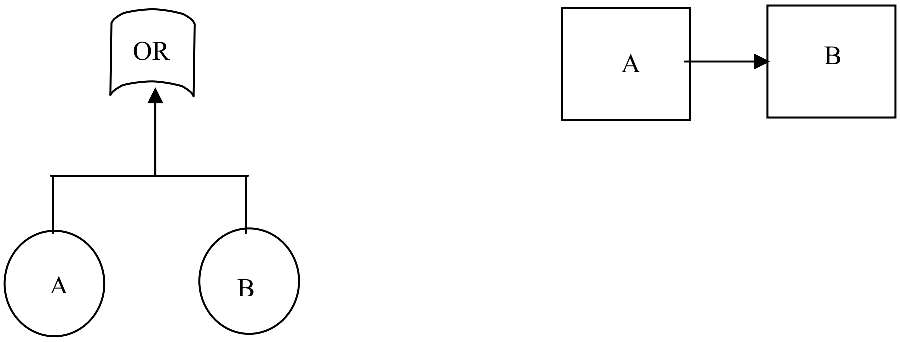


Fig. 4.2 FTA and RBD representation of serial connection

The reliability equation for either configuration depicted on fig. 4.2

$$R_S = R_A R_B \tag{9}$$

When more input events is needed in fig. 4.1 or fig. 4.2, the eq. (8) or (9) is automatically changed for correct input events.

As a example RBD, FTA depicted on fig. 3.1 and fig. 3.2 was simulated in Simulink. Table of distributions is shown in fig. 4.3

	Block							
	A	B	C	D	E	F	G	H
Failure distrib.	Weibull	Exponential	Normal	Weibull	Weibull	Weibull	Exponential	Normal
Param.	β=1,5	m=10000	σ=200	β=1,5	β=3	β=1,5	m=100000	σ=50
	η=1000		μ=1000	η=10000	η=1000	η=5000		μ=5000

Fig. 4.3 Table of failure distributions used in example depicted on fig.3.1 or fig. 3.2

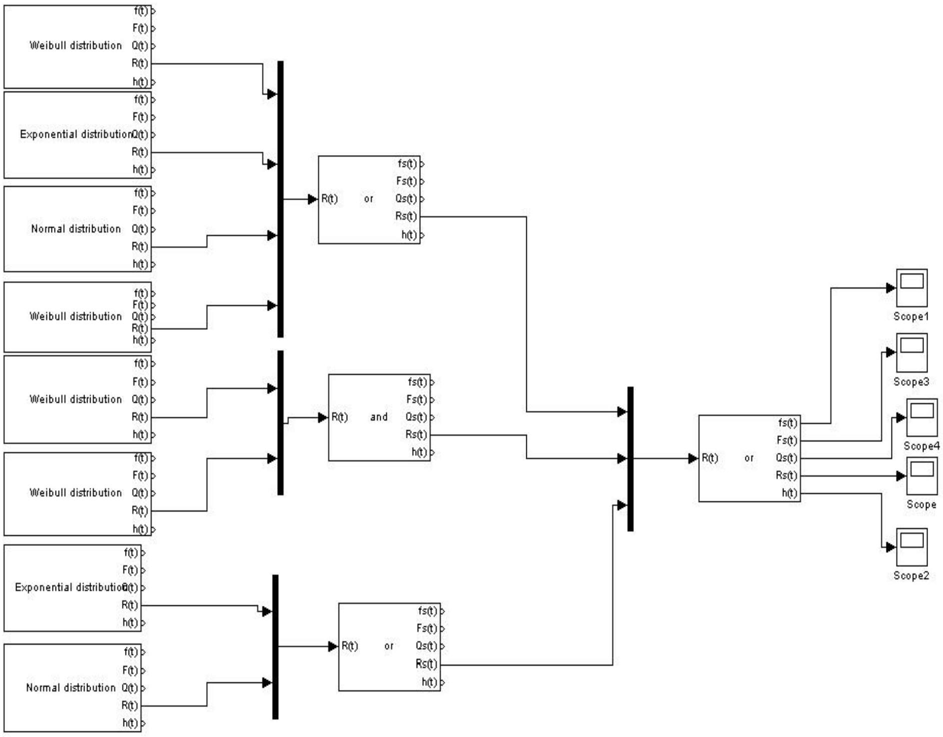


Fig. 4.4 Simulink model of FTA shown in fig. 3.2

5. SIMULATION RESULTS

On the fig. 5.1 up to fig. 5.4 are shown simulation results. Simulation was made in simulink with constant time step, using method ode5 (Dormand-Prince). The end of simulation was at time 1000 tu.

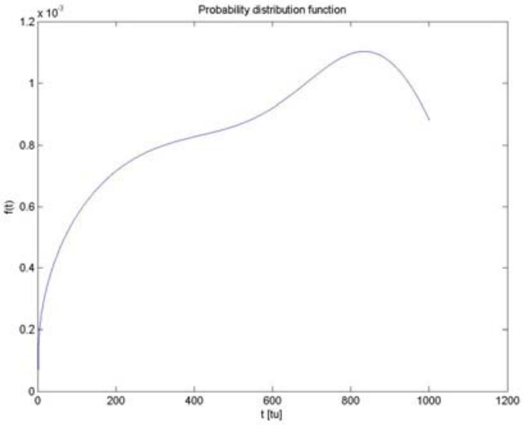


Fig. 5.1 PDF of the system

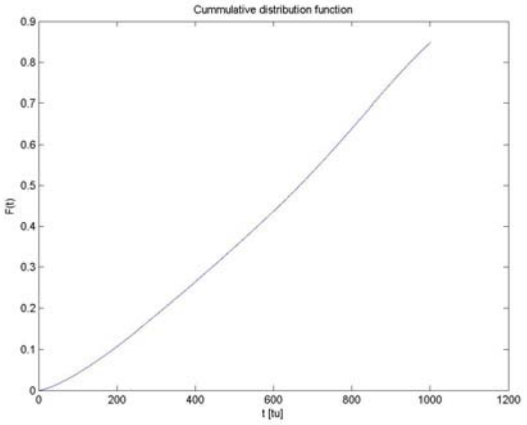


Fig. 5.2 CDF of the system

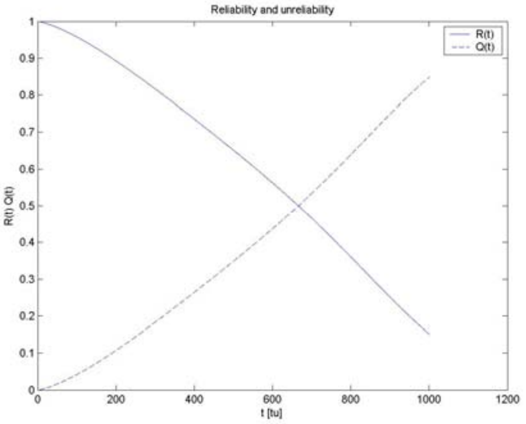


Fig. 5.3 R(t) and Q(t) of the system

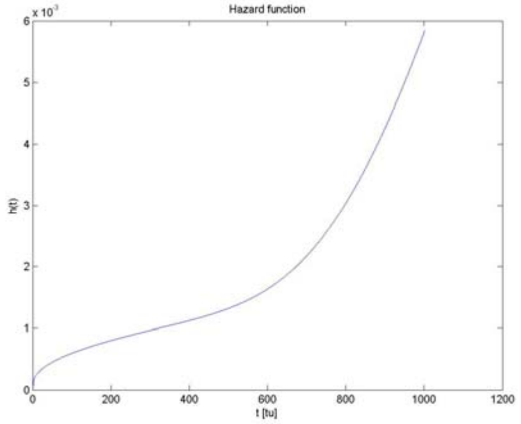


Fig. 5.4 Hazard function of the system

6. CONCLUSION

The purpose of this paper was to shown possibility of simulation reliability block diagrams or failure tree analyses in Matlab/Simulink. The main advantage of the programming in Simulink instead of Matlab is possibility to create appropriate blocks and then easily change FTA or RBD diagrams, Simulink also provide repeatable simulations with various input failure distributions and observing the changes of system failure distributions. The outputs from simulations can be easily processed in other computer programms.

REFERENCES

1. HOLUB, R. – VINTR, Z.: Aplikovné techniky spolehlivosti. Část 1: Specifikace požadavků na spolehlivost. [Skripta]: Vojenská akademie 2002.
2. SYSTEM ANALYSIS REFERENCE: Reliability, availability and optimalization, e-textbook, www.reliasoft.com, 2005
3. LIFE DATA ANALYSIS REFERENCE:e-textbook, www.reliasoft.com, 2005
4. CROWE, D., FEINBERG, A.: Design for reliability. CRC Press. New York. 2001
5. LEITNER, B. : Optimum design of Track Maintenance machine Frames by Matlab. In: Zborník medzinárodnej konferencie "TRANSPORT 2003, VTU Todora Kableshkova, Sofia 2003, str. 157 - 160.

Reviever: : doc. Ing. Bohuš Leitner, PhD.
This article was elaborated at the tasks realisation of granted project VEGA 1/2094/05

SIMULATED ELECTRICAL NETWORK APPROACH (SENA) TO HARD OPTIMIZATION PROBLEMS

Himanshu Dutt Sharma
●
Department of Electrical & Electronics Engineering,
Birla Institute of Technology & science, Pilani, (Raj.) INDIA.333031
E-mail:hdsharma@bits-pilani.ac.in

Bangale Shreyas Madhukarao
●
Department of Mechanical Engineering,
Birla Institute of Technology & Science, Pilani, (Raj.) INDIA.333031
E-mail:f2001413@bits-pilani.ac.in

ABSTRACT

A novel method is proposed for hard optimization type of problem wherein an exact optimal solution is increasingly difficult in terms of run time and memory requirements. Especially for the cases when search graph has higher number of nodes and more number of paths, which increase as factorial of node number. This is based on Simulated Electrical Network Approach (SENA) proposed here, in which the graph is modeled as an electrical network and current distribution is found which is used as a directive for search decisions. The proposed algorithm results in an approximate method that achieves average accuracy of 99.89% to reach close to the most optimal path that is found by ranking all possible paths. Conversely, it can eliminate on average 99.89% paths in polynomial time from consideration if one requires finding the most optimal one.

Key Words: Optimization, NP-Hard, TSP, Shortest path.

1 INTRODUCTION

Because of the intrinsic difficulty in finding polynomial time exact solution algorithms for NP-hard optimization problems, the research has moved to the approximate solutions to these problems and development of approximate algorithms is direction of research. Some of these aim at fast approximate optimization. Traveling salesman's problem is a classical problem of this class.

There is a lot of work going in the area of optimization and planning of shortest path. The shortest-paths problem involves a weighted, possibly directed graph described by the set of edges and vertices $\{E, V\}$. Given a source vertex, s, the goal is to find the shortest existing path between s and any of the other vertices in the graph. There are two types algorithms proposed depending upon the programming involved i.e. sequential and parallel. These types of problems involve weighted graphs and can be applied to Euclidean or non-Euclidean cases. Among these TSP (Traveling Salesman's Problem) stands as one of the most difficult and sought after problem and has remained a challenge for many algorithm planners and also serves as a problem for testing optimization algorithms efficiency. There are many approximations to solve this problem, as any polynomial time algorithm does not find exact solution yet [1]. The approximation algorithm using the triangle inequality is well known developed by Christofides [2]. Artificial intelligence based techniques are also developed to search for optimal paths e.g. genetic algorithms, simulated annealing, and neural nets are examples of these [3]. Some Other attempts to solve TSP include generalization in which, for each city, a neighborhood is specified in which the salesperson can meet the client is also approximable for a variety of neighborhood types such as unit segments, unit circles, and unit rectangles [4]. Another generalization in which the salesperson has to rearrange some objects while following the route is approximable within 2.5 [5]. A prize-collecting variation in which a penalty is associated with each vertex and the goal is to minimize the cost of the tour and the vertices not in the tour [6]. A variation in which vertices can be revisited and the goal is to minimize the sum of the latencies of all vertices, where the latency of a vertex c is the length of the tour from the starting point to c, is approximable within 29 and is APX-complete [7]. A combination of this problem and the matching problem, also called Printed Circuit Board Assembly, is approximable within 2.5 [8]. Finally, the variation in which a Hamiltonian path is looked for rather than a tour is also approximable within 1.5 in the general case while if both end vertices are specified it is approximable within 5/3 [9].

In the proposed Simulated Electrical Network Approach (SENA), to prove that it indeed gives paths in optimal range, some of the most complex sets of input graph are taken i.e. fully connected, non-Euclidean; require traversing all nodes once and only once. The test instances were randomly generated. This is similar to solving a TSP instance in terms of complexity. Cases are tested for varying number of nodes. In these types of graphs, the exact search techniques show exponential complexity for the algorithms, and therefore cannot be applied for the cases with large node numbers. Whereas this technique reduces time complexity to polynomial time and can be used as a good approximate technique for optimization. For large number of nodes its average performance returns a path, which is within top 1.1% of optimal paths. Hence the method introduced in the next section proves to be very effective in producing an approximate solution to the shortest path type problem in a polynomial time.

2 SIMULATED ELECTRICAL NETWORK APPROACH: ANALOGY WITH CURRENT FLOW

In this approach, at first the cost of each link of a given search graph is modeled as a branch resistance of an electrical network. A voltage source is added between pre-specified START and DESTINATION nodes and then current distribution is found in the transformed electrical network. In this electrical network model, the new approach relies on the observation that current flow in an electrical network follows a fundamental rule: maximum amount of current tries to take minimum resistance path, which is the key to eliminate most of the paths which do not come anywhere close to the least resistive (cost) ones. Following this common observation one can bring a sort of foresight in the network for path search type problems. To get the substance of the approach in its simplest form, consider the following simple network's example:

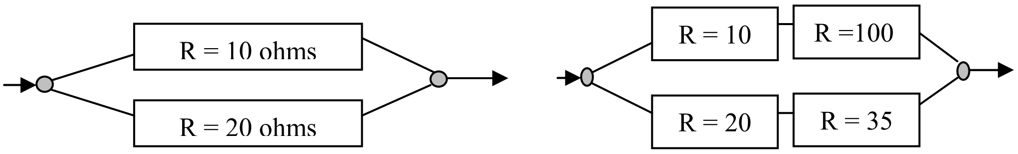


Figure_1 & 2

a) The current division in figure_1 is such that the larger current passes through the lesser resistance path. Therefore a decision can be made at node-1 of choosing a link, which has highest current flowing among all possibilities from that node to other nodes. This in effect equips a path planner with a foresight.

b) The Figure_2 is a modified case of (a). Although the resistances connected to node number '1' remains the same but at a later node, the introduced resistances change the situation in a way that the overall cost of the earlier lesser-cost path is now high but the connections as seen from the first node remain the same. To a simple path planner it becomes necessary to scan the whole set of links coming in path, otherwise it will misguide and a local minimum situation is most likely to be achieved. Whereas a current flow based approach would be a better decision maker in such cases. It will still give the correct path through R=20,35 part of the circuit. A full appraisal of proposed method in large fully connected networks is done in following sections.

3 ALGORITHM FOR SENA

This approach is applicable to both Euclidean and Non-Euclidean type of graphs, as the costs are not chosen on Euclidean basis. The algorithm is tested for fully connected, non-Euclidean; require traversing all nodes once and only once before reaching to a pre-specified destination. Here, It is to be noticed that these type of problems are as hard as TSP. The approach involves three major steps:
- Modeling the given graph in Electrical circuit;

- Solving the modeled electrical network for currents in each branch;
- Simple decision making by looking at current-magnitudes from each node;

And a straightforward very close solution to the exact shortest path is achieved.
The detail procedure adopted is as follows:

- Generate random instances of graphs;
- Convert cost to resistance i.e. $R_{ij}= C_{ij}$; where i & j are nodes
- Identify the source node as start point, make it positive terminal of battery;
- Identify the destination node make it a negative terminal of battery;
- Specify a voltage and solve it for current in each branch;
- From each node go to the next node to which maximum current flows;
- Destination node is traversed only in the last step.

A path thus found is indeed very close to exact optimal. See table figure_3.

4 RESULTS

In the graph shown in figure_3, we generated as many as 25 different fully connected graph and exact solutions for each node number, varying the number of nodes as N = 6,7,8,9,10,11,12,13. The search performed for shortest path had to get a path from possible number of paths ranging as few as 24 for N=6 to as many as 39,916,800 for N=13.

Table1: % Optimization Achieved for Number of Nodes in Graphs

Trials	Node 6	Node 7	Node 8	Node 9	Node 10	Node 11	Node 12	Node 13
1	73.91	94.96	60.64	99.27	84.37	98.4	99.96	99.27
2	73.91	96.64	91.38	98.47	95.19	99.91	99.86	99.99
3	69.56	65.55	90.33	94.30	99.91	99.99	99.99	99.96
4	78.26	88.24	94.30	96.49	83.80	99.39	99.98	99.99
5	91.30	96.64	96.11	99.76	83.89	97.77	99.99	99.99
6	82.61	95.80	97.36	90.67	96.28	96.48	99.83	99.76
7	**100**	44.54	97.08	98.33	98.83	97.97	99.41	99.99
8	82.61	**100**	98.89	99.80	99.68	99.29	99.93	99.98
9	60.87	86.55	98.47	77.40	96.38	99.99	99.98	99.97
10	95.65	94.96	74.55	92.94	99.78	99.99	96.89	99.99
11	**100**	95.80	98.05	**100**	99.72	99.97	99.83	99.99
12	86.96	**100**	94.02	96.47	93.33	99.99	99.81	99.66
13	82.61	**100**	79.00	96.57	99.91	99.99	99.99	99.97
14	30.43	97.48	93.60	75.00	97.80	99.82	99.94	99.51
15	**100**	**100**	80.67	99.44	93.34	99.66	99.48	99.92
16	**100**	93.28	90.33	99.98	99.73	99.85	98.52	99.8
17	95.65	83.19	95.83	98.83	99.15	98.84	99.41	99.99
18	78.26	**100**	90.40	83.51	99.99	98.7	99.89	99.67
19	69.56	**100**	98.89	66.24	99.74	99.99	90.45	99.99
20	**100**	98.32	95.27	93.65	99.84	99.68	99.86	**100**
21	82.61	99.16	99.30	94.90	99.64	97.22	99.84	**100**
22	69.57	77.31	95.41	95.61	98.95	99.37	99.83	99.99
23	95.65	83.19	97.50	**100**	99.26	99.29	99.39	99.97
24	30.43	90.76	61.75	91.70	91.97	99.98	99.89	99.99
25	91.30	66.39	91.80	99.94	99.85	99.81	99.95	99.99

Figure_3

Table2: Overall Analysis of Results Figure_4

SR. NO.	No. of Nodes	No. of Samples	Total No. Of paths	Best result ranking	Average Result (Within x% of Optimal)	Variance	Average Eliminated Paths (%) $\overline{X}$
1	6	25	24	1st	19.19	18.67	80.81
2	7	25	120	1st	10.05	13.44	89.95
3	8	25	720	1st	9.33	10.62	90.67
4	9	25	5040	1st	6.46	8.68	93.54
5	10	25	40,320	14th	3.59	5.12	96.41
6	11	25	362,880	26th	0.76	0.956	99.24
7	12	25	3,628,800	176th	0.73	0.91	99.27
8	13	25	39,916,800	1st	0.11	0.181	99.89

Figure_4

The table in figure_4 and Plot of average of results converging towards the exact optimal solution in figure_5 along with standard deviation show the effectiveness of the algorithm. Following points are inferred from these,

1 As number of paths increase, the results are the bold line in graph, which refers to the steep *Average convergence* of results towards optimal solution. As seen, it is coming in close range of within 0.11 % of the most optimal path. *The range x% refers:* “the percentage of number of paths lying in between the achieved solution and the exact optimal solution”. This comparison range is achieved, by finding all possibilities of paths and costs, then sorting these in increasing order for all instances.

2 The fourth column in table *best result ranking* shows that for higher number of nodes in our randomly generated instances, the best result was as close as 99.9928% to the most optimal path i.e. 26th out of 362,880 paths. 99.9979% to the most optimal path i.e. 176th out of 3,628,800 paths. For lesser number of nodes many times it found the best path i.e. 100% optimization e.g. 1st out of 720 paths. Also at random sometimes it hits best path for large number of nodes i.e. twice in case of 13 nodes.

3 The monotonous nature of the Average optimization curve and Standard deviation curve in figure_5 shows that there is natural tendency of this algorithm to converge towards optimal result as the number of nodes increase.

4 The nature of small and reducing standard deviation reveals that the samples had smaller and smaller deviation in their average performance upon increasing the node numbers.

5 The last column of table_2 shows elimination capacity of algorithm. It is also getting better for higher node numbers, e.g. for N=13 on an average it can eliminate ~99.89% paths which are away from optimal.

6 Following graph in fig_5 clearly depicts that optimization as well as deviation nearly reaches to the best as nodes are increased. Thus hypothesis proves to be an excellent approximate algorithm for optimization in highly complex, large number of nodes, fully connected graphs of Euclidean or Non- Euclidean type.

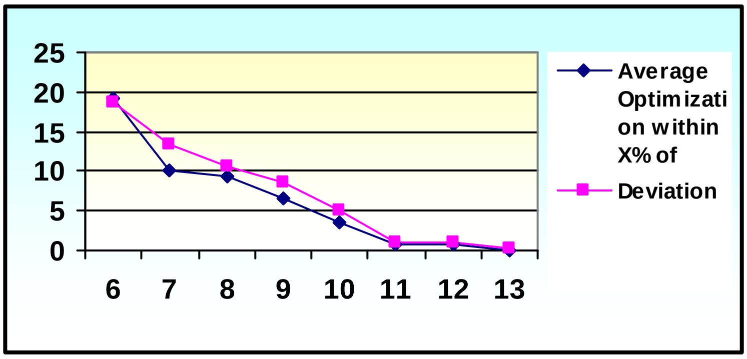


Figure_5

5 ERROR CALCULATION (ACCEPTANCE REGION OF μ)

To validate results further, the maximum error in $\overline{X}$ is calculated, given the probability as 99%. This will also give the estimate of actual population mean μ. Twenty-five samples are taken for each node number for optimization. For these 25 samples, the mean estimates of $\overline{X}$ are $\overline{X}$ = 80.81 for (nodes 6), $\overline{X}$ = 89.95 for (nodes 7), $\overline{X}$ = 90.33 for (nodes 8), $\overline{X}$ = 93.54 for (nodes 9), $\overline{X}$ = 96.41 for (nodes 10), $\overline{X}$ = 99.24 for (nodes 11), $\overline{X}$ = 99.27 for (nodes 12), $\overline{X}$ = 99.89 for (nodes 13). Therefore In our case: n = 25,P = 0.99,1-∝ = 0.99,∝ = 0.01 and $Z_{\propto/2}$ = $Z_{0.005}$ = 2 .575. Then Maximum error of estimate is given by the following formula. $E = Z_{\alpha/2}(\sigma/\sqrt{n})$

Figure_6

Nodes	$Z_{\propto/2}$	σ	Mean based on trials $\overline{X}$ %	E= $\lvert \overline{X} - \mu \rvert$ (probabilistic estimation in error with 99% confidence)	Lowest possible value of μ %	Highest possible value of μ %
6	2.575	18.67	80.81	9.61	71.2	90.42
7	2.575	13.44	89.95	6.92	83.03	96.87
8	2.575	10.62	90.33	5.47	84.86	95.8
9	2.575	8.68	93.54	4.47	89.07	98.01
10	2.575	5.127	96.41	2.64	93.77	99.05
11	2.575	0.956	99.24	0.49	98.75	99.73
12	2.575	0.91	99.27	0.47	98.80	99.74
13	2.575	0.181	99.89	0.09	99.80	99.98

From above graph and table figure_6&7, it is concluded that actual mean i.e. average optimization achieved in the proposed algorithm will be greater than 99% for nodes above N=10, with 99% confidence. It strongly supports the SENA's validity.

Figure_7

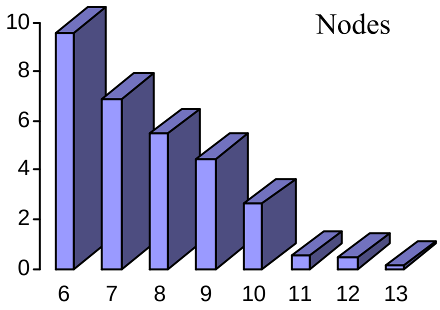
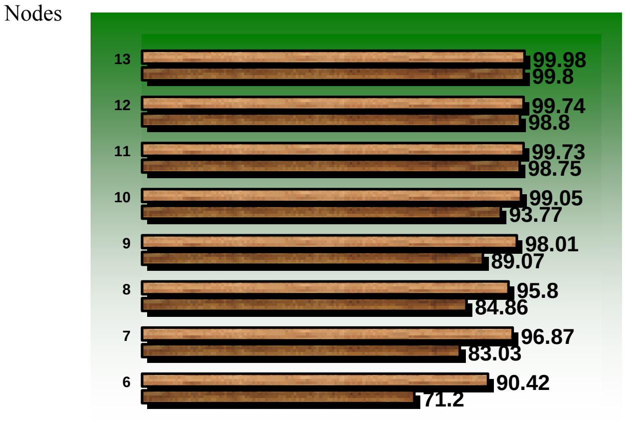


Figure_8



V

6 PROBABILITY OF ACHIEVING OPTIMIZATION

Given are the values of $\overline{X}$ for different nodes for a sample of 25 with variance (σ) values, following table finds probability of getting optimization between 90% to 100 % for different number of nodes using the following probability function: $F(Z) = F((\overline{X} - \mu)/\sigma * n^{1/2})$

Nodes	$\overline{X}$	μ Interval	Variance σ	Trials n	F (Z) for μ = 90	F (Z) for μ = 100	P= F (Z) $_{90}$ – F (Z) $_{100}$
6	80.81	90 - 100	18.67	25	0.0069	0.001	0.0059
7	89.95	90 –100	13.44	25	0.496	0.0001	0.496
8	90.33	90 –100	10.62	25	0.5596	0.0002	0.5596
9	94.56	90 –100	8.68	25	0.9788	0	0.9788
10	96.41	90 –100	5.127	25	0.99999	0	0.9999 ≈ 1
11	99.24	90 - 100	0.956	25	1	0	≈ 1
12	99.27	90 - 100	0.91	25	1	0	≈ 1
13	99.89	90 - 100	0.18	25	1	0	≈ 1

Figure_9

The above-tabulated results are plotted and it shows that as number of nodes increases the probability of getting optimization between the said regions is approaching 1.

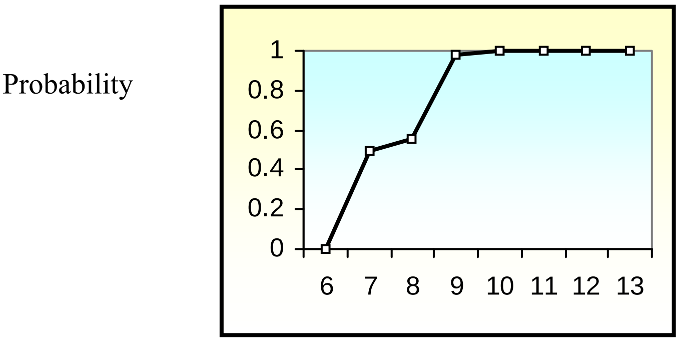


Figure_10

7 INVERSE LOGIC FOR SENA (AN ALTERNATIVE APPROACH)

An alternative to the one stated above is proposed here. **Path is searched on the basis of inverse technique where inverse values of costs are taken and path is searched for minimum current. Costs are then added normally for checking optimization.** After applying the normal probabilistic hypothesis, it came to know that inverse technique optimization increases significantly for the cases where optimization is less in normal SENA technique. Comparison is sufficient for 10 numbers of nodes and prediction can be done for further nodes on the basis of probability and statistics.

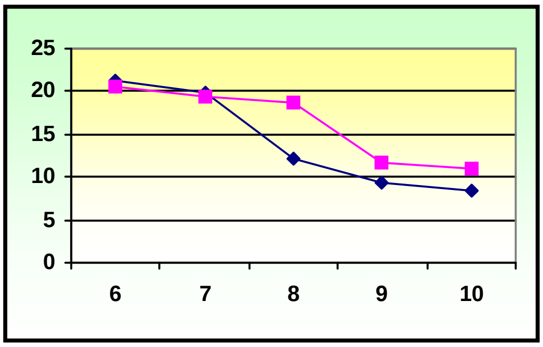
TRIALS	NODES 6	NODES 7	NODES 8	NODES 9	NODES 10
1	95.65	79.83	95.41	78.29	99.10
2	95.65	57.98	99.17	99.05	99.22
3	91.30	89.08	90.68	99.35	99.70
4	69.57	97.48	98.19	95.42	88.03
5	91.30	**100**	99.43	97.20	99.69
6	86.96	51.26	92.63	95.36	99.94
7	95.65	98.32	64.39	95.48	95.92
8	82.61	95.80	88.48	71.18	99.94
9	73.91	82.35	95.13	73.69	96.58
10	65.22	83.19	97.07	99.40	56.34
11	56.52	94.96	95.83	99.01	99.34
12	26.09	98.32	92.49	88.27	94.11
13	82.61	97.48	90.82	57.93	92.39
14	56.57	96.64	92.91	85.29	98.76
15	95.65	79.83	66.90	99.62	92.39
16	95.65	80.14	34.63	97.50	99.74
17	56.52	96.64	98.61	99.82	87.25
18	86.96	78.15	99.86	94.17	98.59
19	60.87	97.48	**100**	87.34	91.20
20	**100**	85.71	96.49	94.43	98.31
21	95.65	70.59	99.58	98.79	64.04
22	86.96	99.16	99.72	65.59	88.30
23	95.65	87.39	28.93	99.72	99.81
24	30.43	21.08	90.13	97.62	96.94
25	95.65	60.50	89.85	95.10	91.66

Figure_11: % Optimization Achieved for Number of Nodes in Graph

Figure_12: Overall Analysis of Results

Nodes	Optimization%	Deviation σ
6	78.77	20.48
7	80.14	19.38
8	87.89	18.77
9	90.58	11.69
10	91.66	10.89

Figure_13



Curve clearly depicts that optimization as well as deviation nearly reaches to 10% as we further increase the nodes. PINK line denotes efficiency while BLUE line indicates deviation. Thus hypothesis proves better.

8. APPLICABILITY OF INVERSE TECHNIQUE FOR OPTIMIZATION

At first, It looks as if inverse technique is not so useful because of its inconsistent results. But it is quite useful when acyclic technique is giving bad results in terms of optimization. 25 trials were taken and graphs for various nodes were plotted. It is concluded that when acyclic graphs were giving depressions for particular regions, at that time inverse was at its peak i.e. better optimizing. For example, we have given the visual reference to it in the following figure_14, 15, and 16. Following Charts show comparison of trials between normal and inverse technique (an example is taken for 9 nodes) as well as distribution about superiority of inverse and acyclic techniques over each other for similar 25 trials. Pointer shows for example; how inverse succeeds over normal SENA for a given case.

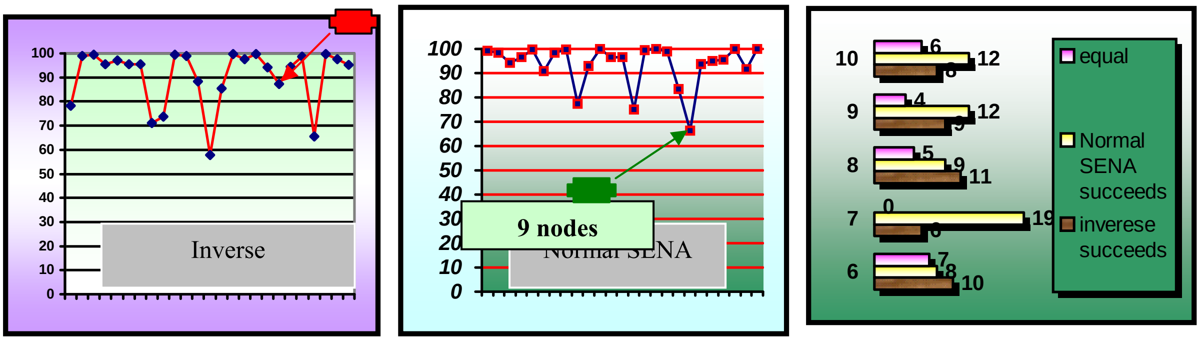


Figure_14, 15, 16

9 CYCLIC NETWORK

Algorithm for cyclic network is given below.
1. Take node input from user.
2. Start and goal destinations are the same.
3. For programming purpose, develop a pseudo goal node.
4. Resistance between all nodes and pseudo goal node will be of the order of the resistance value between start node and all other corresponding nodes.
5.Now same logic is implemented as that of acyclic network to find the path.
6.At last, pseudo goal node is removed from the obtained path and in that place; start node is kept for the final path.

Since the trend shown by the results of cyclic graphs match that of acyclic graph optimization pattern, therefore results over nine nodes are avoided here.

9.1 RESULTS

Trials	Node 6	Node 7	Node 8	Node 9
1	94.958	90.542	98.591	89.107
2	78.992	96.801	92.102	98.646
3	81.513	89.847	82.616	99.807
4	65.546	71.21	97.5	99.529
5	86.555	99.722	99.861	99.747
6	99.16	99.583	93.848	99.571
7	**100**	**100**	97.321	99.722
8	82.353	97.497	99.623	99.893
9	99.16	97.775	98.809	84.824
10	**100**	91.099	99.782	99.606
11	90.756	98.609	99.861	99.98
12	**100**	**100**	92.499	99.931
13	98.319	97.357	95.872	98.79
14	**100**	99.166	99.008	98.676
15	99.16	99.305	**100**	99.576
16	98.319	83.032	97.023	**100**

Trials	Node 6	Node 7	Node 8	Node 9
17	98.319	98.609	99.028	98.971
18	91.597	96.801	95.138	93.219
19	83.193	94.159	94.245	97.433
20	91.597	86.787	96.229	92.52
21	**100**	99.583	99.544	92.969
22	82.353	91.377	98.234	97.133
23	**100**	91.377	58.821	99.812
24	99.16	69.68	90.375	99.955
25	94.118	99.722	95.197	99.762

Figure_17 % Optimization Achieved for Number of Nodes in Graphs

Nodes	Optimization%	DEVIATION σ
6	92.605	9.1
7	93.586	8.37
8	94.845	8.48
9	97.567	3.95

Figure_18: Overall Analysis of Results

9.2 PROBABILITY OF ACHIEVING OPTIMIZATION

Nodes	$\overline{X}$	μ Interval	Variance σ	Trials n	F (Z) for μ = 90	F (Z) for μ = 100	P= F (Z) $_{90}$ – F (Z) $_{100}$
6	92.605	90 - 100	9.1	25	0.9236	0.00003	**0.9236**
7	93.586	90 –100	8.37	25	0.9838	0.00001	**0.9838**
8	94.845	90 –100	8.48	25	0.9979	0.0012	**0.9967**
9	97.567	90 –100	3.95	25	1	0.001	**≈ 1**

Figure_19 Probability Estimation

The above-tabulated results show that as number of nodes increase, the probability of getting optimization between the said regions is approaching 1. It is faster than what was observed in the acyclic graphs.

10 CONCLUSION

It has been shown in the results that the proposed SENA-method is capable of returning a near optimal solution for shortest path finding type of problems. Thus it can handle hard optimization for the problems that have complexity of fully connected graphs where the number of possible paths increase in proportion to factorial of number of nodes in the graph. Also this technique is applicable to Euclidean or non-Euclidean cases equally well as there are no constraints of Euclidean geometry assumed in the formation of graph instances. Also it is established that its elimination capacity for paths which will be close to optimal, from all possibilities is reaching to ~99% on average basis for higher number of nodes where other techniques starts reducing their efficiency, in contrast, its in fact monotonously showing better results. The statistical analysis shows using probabilistic estimate that as number of nodes increases; the near complete optimization can be achieved. The cases where SENA fails to achieve required optimization; inverse technique can be used as an efficient tool. So conclusively it's quite effective for determining a close to optimal heuristic. Further, it is shown that the technique is equally applicable to cyclic graphs also. *Therefore the SENA algorithm is capable of returning result which is nearly the best optimization in graphs of varying nature: cyclic, acyclic, Euclidean, non-Euclidean and of the highest order of complexity i.e. fully connected wherein the total paths are increasing in proportion to factorial of node number.*

11 REFERENCES

1. Cormen, Leiserson and Rivest, Introduction to Algorithms, McGraw-Hill, 1994.
2. N.Christofides, Worst-case analysis of a new heuristic for the traveling salesman problem, Symposium on New Directions and Recent Results in Algorithms and Complexity, page 441, New York, NY, 1976. Academic Press.
3. E. L. Lawler, J. K. Lenstra, A. H. G. Rinnooy Kan, and D. B. Shmoys. The Travelling Salesman Problem: A Guided Tour of Combinatorial Optimization, Wiley, New York, 1992.
4. Arkin, E. M., and Hassin, R. ``Approximation algorithms for the geometric covering salesman problem'', *Disc. Appl. Math.* **55**, 197-218. (1994),
5. Anily, S., and Hassin, R. ``The swapping problem'', *Networks* **22**, 419-433. (1992),
6. Goemans, M. X., and Williamson, D. P. ``A general approximation technique for constrained forest problems'', *SIAM J. Comp.* **24**, 296-317. (1995a),
7. Blum, A., Chalasani, P., Coppersmith, D., Pulleyblank, B., Raghavan, P., and Sudan, M, ``The minimum latency problem'', *Proc. 26th Ann. ACM Symp. on Theory of Comp.*, ACM, 163-171 (1994).
8. Michel, C., Schroeter, H., and Srivastav, A., ``TSP and matching in printed circuit board assembly'', *European Symposium on Operations Research*. (1995).
9. Hoogeveen, J. A. ``Analysis of Christofides heuristic: Some paths are more difficult than cycles'', *Oper. Res. Lett.* **10**, 178-193 (1978).

RISK ASSESSMENT AND OPTIMIZATION OF ROAD TUNNELS

Milan Holický
●
Klokner Institute, CTU in Prague, Šolínova 7,
166 08 Praha 6, Czech Republic, Tel: +420 224 310 208,
Fax: +420 224 355 232,
holicky@klok.cvut.cz

Abstract

Probabilistic methods of risk optimization are applied to specify the most effective arrangements of road tunnels. The total consequences of alternative arrangements are assessed using Bayesian networks supplemented by decision and utility nodes. It appears that the optimization may provide valuable information for a rational decision concerning number of escape routes. Discount rate seems to affect the total consequences and the optimum arrangements of the tunnels more significantly than number of escape routes.

Key words

Risk assessment, social risks, economic consequences, road tunnels, Bayessian network, optimization, escape routes, discount rate, expected life time

INTRODUCTION

Tunnel structures usually represent complex technical systems that may be exposed to hazard situations leading to unfavourable events with serious consequences. Minimum safety requirements for tunnels in the trans-European road network are provided in the Directive of the European Parliament and of the Council 2004/54/ES [1]. The Directive also gives recommendations concerning risk management, risk assessment and analysis.

Methods of risk assessment and analysis are more and more frequently applied in various technical systems [2,3] including road tunnels [4]. This is a consequence of recent tragic events in various tunnels and of an increasing effort to take into account social, economic and ecological consequences of unfavourable events [2,3,4]. Available national and international documents [5] to [10] try to harmonise general methodical principles and terminology that can be also applied in the risk assessment of road tunnels. The submitted contribution, based on previous studies [11] to [17] and recent PIARC working documents, attempts to apply methods of probabilistic risk optimization using Bayesian networks supplemented by decision and utility nodes [18]. It appears that Bayesian networks provide an extremely effective tool for investigating the safety of road tunnels.

GENERAL PROCEDURE OF RISK ASSESSMENT

The main components of the whole risk management consist of risk assessment and risk control. The risk control is outside the scope of this paper. The risk assessment consists of risk analysis and risk evaluation. A general procedure of risk assessment is shown in Figure 1 indicating a flowchart of the main steps. The flowchart is adopted from ISO document [9] and from recent working materials of PIARC/C3.3/WG2. The contents of individual steps are mostly obvious from the relevant key words used for description of the flowchart. Two key steps of the risk analysis, probability analysis and risk estimation are shortly described below.

PROBABILITY ANALYSIS

Probabilistic methods of risk analysis are based on the concept of conditional probabilities $P_{fi} = \mathrm{P}\{F|H_i\}$ of the event F providing a situation H_i occurs [1, 3]. In general this probability can be found using statistical data, experience or theoretical analysis of the situation H_i.
If the situation H_i occurs with the probability $\mathrm{P}(H_i)$ and the event F during the situation H_i occurs with the probability $\mathrm{P}(F|H_i)$, then the total probability P_F of the event F is given as

$$P_F = \sum_i \mathrm{P}(F \mid H_i)\mathrm{P}(H_i) \qquad (1)$$

Equation (1) makes it possible to harmonize partial probabilities $\mathrm{P}(F|H_i)$ $\mathrm{P}(H_i)$ related to the situation H_i.

The main disadvantage of the purely probabilistic approach is the fact that possible consequences of the events F related to the situation H_i are not considered. Equation (1) can be, however, modified to take the consequences into account.

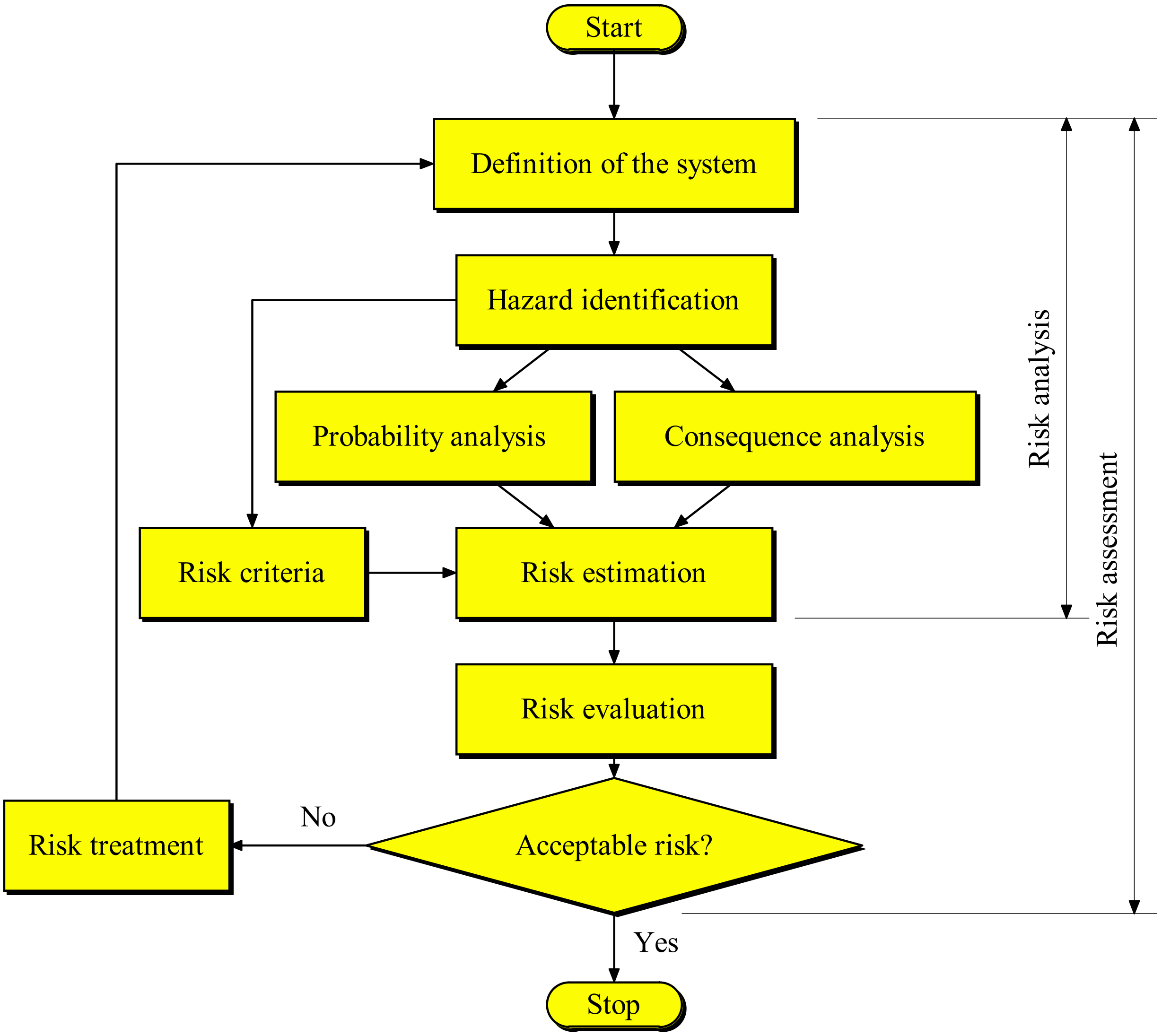


Figure 1. Flowchart of iterative procedure for the risk assessment (adopted from [9])

RISK ESTIMATION

A given situation H_i may lead to a set of events E_{ij} (for example fully developed fire, explosion), which may have social consequences R_{ij} or economic consequences C_{ij}. It is assumed that the consequences R_{ij} and C_{ij} are unambiguously assigned to events E_{ij}. If the consequences include only social components R_{ij}, then the total expected risk R is given as [11]

$$R = \sum_{ij} R_{ij} \mathrm{P}(E_{ij} \mid H_i)\mathrm{P}(H_i) \tag{2}$$

If the consequences include only economic consequences $\mathrm{C_{ij}}$, then the total expected consequences C are given as

$$C = \sum_{ij} C_{ij} \mathrm{P}(E_{ij} \mid H_i)\mathrm{P}(H_i) \tag{3}$$

If criteria R_d and C_d are specified, then acceptable total consequences should satisfy the conditions

$$R < R_d \text{ and } C < C_d \tag{4}$$

that supplement the traditional probabilistic condition $P_f < P_{fd}$.

When the criteria are not satisfied, then it may be possible to apply a procedure of risk treatment as indicated in Figure 1. For example additional escape routes may be provided. Such measures might, however, require considerable costs, which should be considered when deciding about the optimum measures.

PRINCIPLES OF RISK OPTIMIZATION

The total consequences $C_{tot}(k,p,n)$ relevant to the construction and performance of the tunnel are generally expressed as a function of the decisive parameter k (for example of the number k of escape routes), discount rate p (commonly about $p \approx 0{,}03$) and life time n (commonly n = 100 let). The decisive parameter k usually represents a one-dimensional or multidimensional quantity significantly affecting tunnel safety.

The fundamental model of the total consequences may be written as a sum of partial consequences as

$$C_{tot}(k,p,n) = R(k,p,n) + C_0 + \Delta C(k) \tag{5}$$

In equation (5) $R(k,p,n)$ denotes expected social risk that is dependent on the parameter k, discount rate p and life time n. C_0 denotes the basic of construction cost independent of k, and $\Delta C(k)$ additional expenses dependent on k. Equation (5) represents, however, only a simplified model that does not reflect all possible expenses including economic consequences of different unfavourable events and maintenance costs.

The social risk $R(k,p,n)$ may be estimated using the following formulae

$$R(k,p,n) = N(k)\, Z_1\, Q(p,n), \quad Q(p,n) = \frac{1 - 1/(1+p)^n}{1 - 1/(1+p)} \tag{6}$$

In equation (6) $N(k)$ denotes number of expected fatalities per one year (dependent on k), Z_1 denotes acceptable expenses for averting one fatality, and p the discount rate (commonly within the interval from 0 to 5 %). The quotient q of the geometric row is given by the fraction $q = 1/(1+p)$. The discount coefficient $Q(p,n)$ makes it possible to express the actual expenses Z_1 during a considered life time n in current cost considered in (5). In other words, expenses Z_1 in a year i correspond to the current cost $Z_1\, q^i$. The sum of the expenses during n years is given by the coefficient $Q(p,n)$.

A necessary condition for the minimum of the total consequences (5) is given by the vanishing of the first derivative with respect to k that may be written as

$$\frac{\partial N(k)}{\partial k} Z_1 Q(p,n) = -\frac{\partial \Delta C(k)}{\partial k} \tag{7}$$

In some cases this condition may not lead to a practical solution, in particular when the discount rate p is small (a corresponding discount coefficient $Q(p,n)$ is large) and there is a limited number of escape routes k that can not be arbitrary increased.

STANDARDIZED CONSEQUENCES

The total consequences given by equation (5) may be in some cases simplified to a dimensionless standardized form and the whole procedure of optimization may be generalized. Consider as an example the optimization of the number k of escape routes. It is assumed that involved additional costs $\Delta C(k)$ due to k may be expressed as the product $k\, C_1$, where C_1 denotes cost of one escape route. If C_1 is approximately equal to expenses Z_1 (assumed also in [14]), equation (5) may be written as

$$C_{tot}(k,p,n) = N(k)\, C_1\, Q(p,n) + C_0 + k\, C_1 \tag{8}$$

This function can be standardized as follows

$$\kappa(k,p,n) = \frac{C_{tot}(k,p,n) - C_0}{C_1} = N(k) Q(p,n) + k \tag{9}$$

Obviously both variables $C_{\rm tot}(k,p,n)$ and $\kappa(k,p,n)$ are mutually uniquely dependent and have the extremes (if exist) for the same number of escape routes k. A necessary condition for the extremes follows from (7) as

$$\frac{\partial N(k)}{\partial k} = -\frac{1}{Q(p,n)} = -\frac{1-1/(1+p)}{1-1/(1+p)^n} \qquad (10)$$

An advantage of standardized consequences is the fact that it is independent of C_0 and C_1. It is only assumed that $C_1 \approx Z_1$ is a time invariant unit of the total consequences.

MODEL OF A TUNNEL

A road tunnel considered here (Figure 2) is partly adopted from a recent study [14]. It is assumed that the tunnel has the length of 4000 m and two traffic lanes in one direction are used by heavy goods vehicles HGV, dangers goods vehicles DGV and Cars.

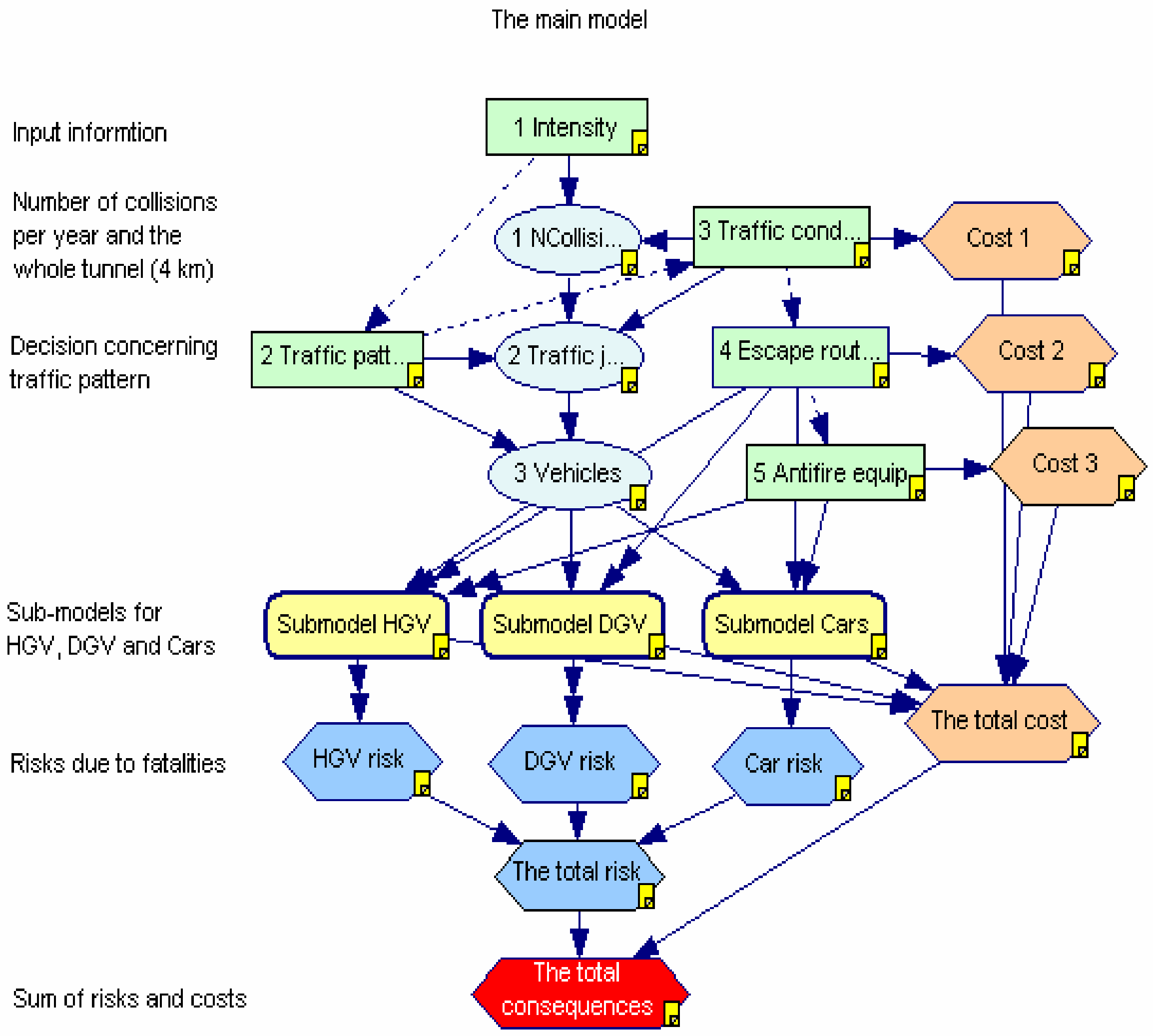


Figure 2. Main model of the tunnel

The total traffic intensity in one direction is 20×10^6 vehicles per year (27 400 vehicles in one lane per day). The number of individual types of vehicles is assumed to be HGV:DGV:Cars = 0,15:0,01:0,84. The frequency of series accidents for basic traffic conditions (that might be possibly improved) is considered as 1 $\times10^{-7}$ per one vehicle and one km [14], thus 8 accidents in the tunnel per year.

The main model of the tunnel shown in Figure 2 includes three sub-models for HGV, DGV and Cars, which describe individual hazard scenarios. The Bayesian networks used here need a number of other input data. Some of them are adopted from the study [14] (based on event tree diagram), the other are estimated or specified using expert judgement. Detailed description of the model is outside the scope of this contribution.

RISK OPTIMIZATION

Risk optimization of the above described tunnel is indicated for selected input data in Figure 3, Figure 4 and 5. Figure 3 shows variation of the components of standardized total consequences $\kappa(k,p,n)$ with number of escape routes k for a common value of the discount rate p = 0,03 and assumed life time n = 100 years.

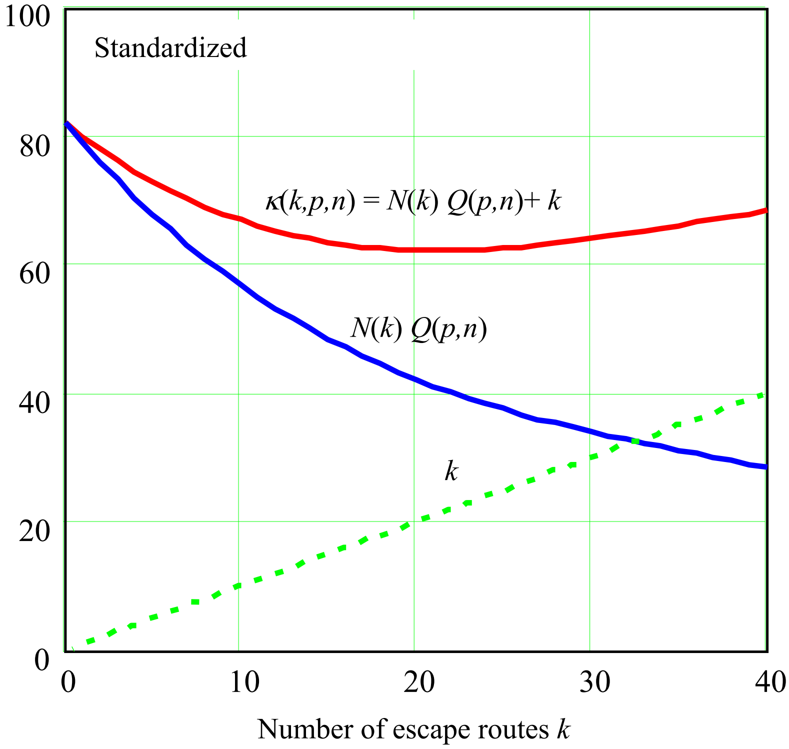


Figure 3. Variation of the components of standardized total consequences κ(k,p,n) with k for the discount rate p = 0,03 and life time n = 100 years

Figure 4 shows variation of the standardized total consequences $\kappa(k,p,n)$ with k for selected discount rate p life time n = 50 years only, Figure 5 shows similar curves as Figure 4 but for expected life time n = 100 years (common value). Both Figures 4 and 5 clearly indicate that the discount rate p and life time n affect the total consequences more significantly than the number of escape routes k. It appears that the total consequences considerably increase with increasing n. For small discount rates $p \leq 0.01$ and life time n = 100 years the total consequences decrease monotonously with increasing k and for $k \leq 39$ (the distance of escape routes up to 100 m) do not reach its minimum. Therefore, in this case condition (10) does not lead to a practical solution.

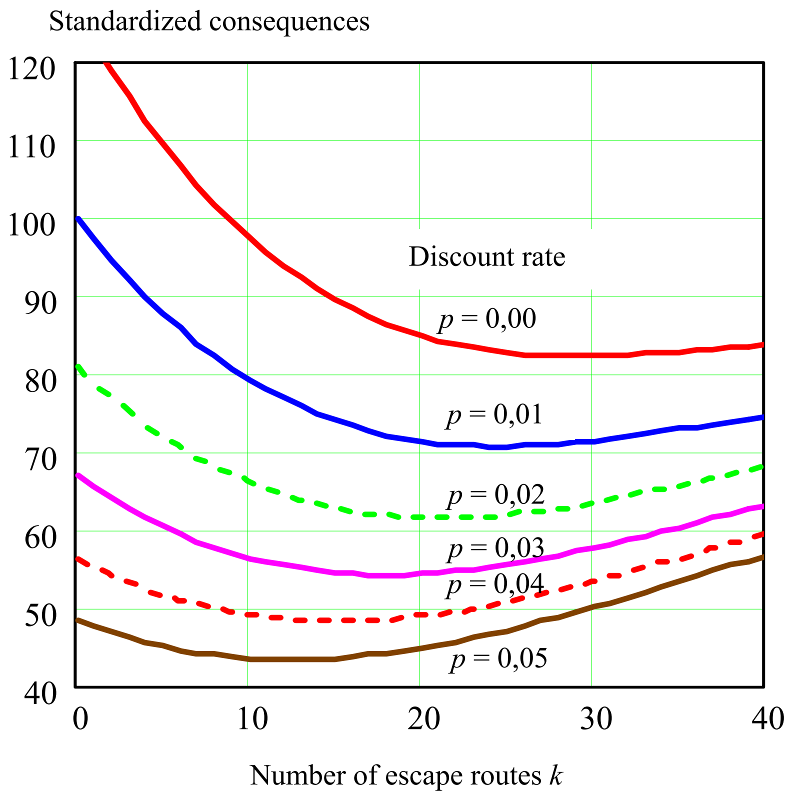


Figure 4. Variation of the standardized total consequences κ(k,p,n) with k for selected discount rate p life time n = 50 years

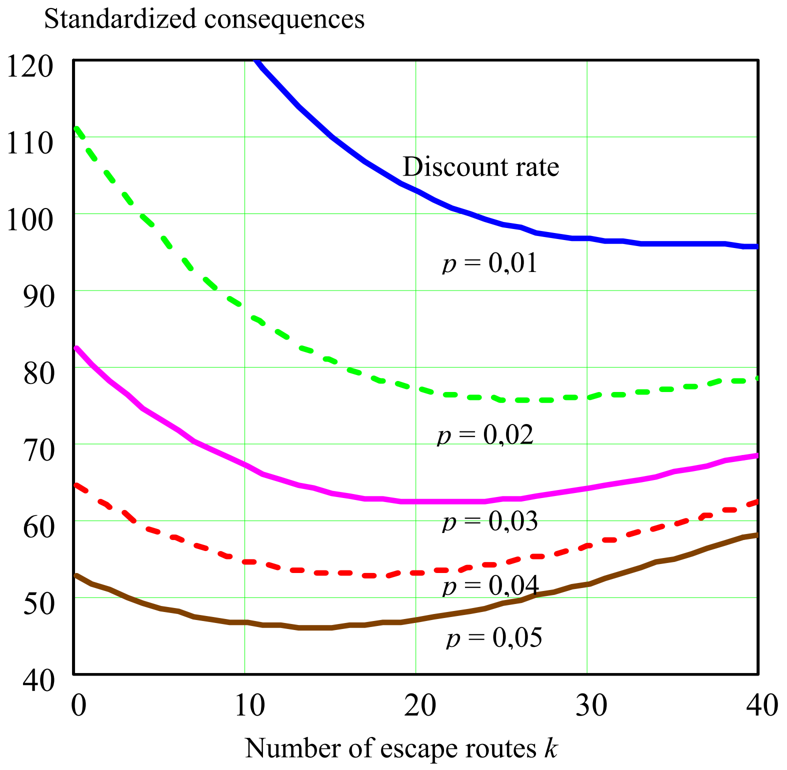


Figure 5. Variation of the standardized total consequences κ(k,p,n) with k for selected discount rate p life time n = 100 years

Figure 6 shows variation of the total consequences $\kappa(k,p,n)$ with number of escape routes k and discount rate p assuming again expected life n = 100 years.

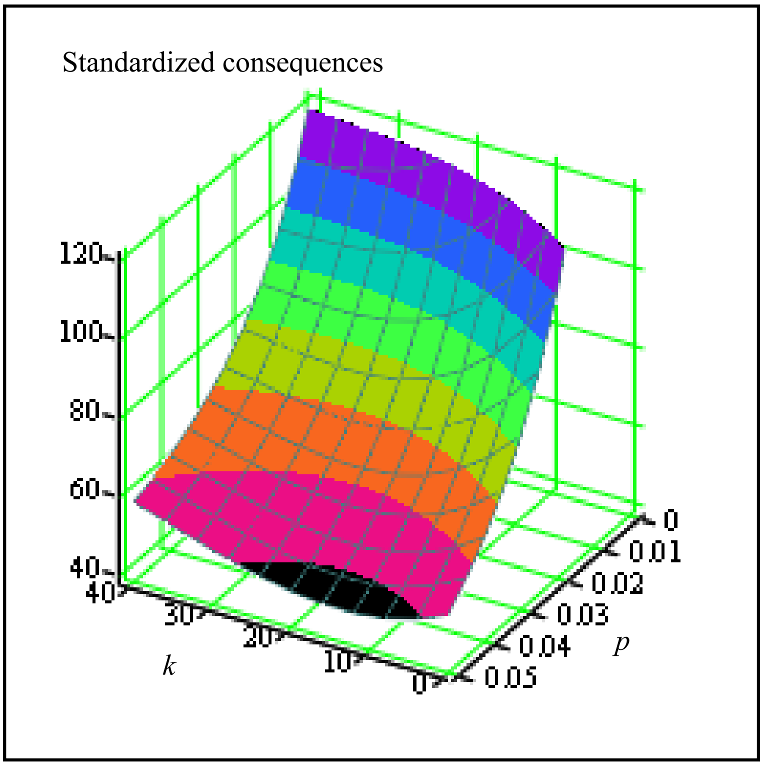


Figure 6. Variation of the standardized total consequences κ(k,p,n) with k for selected discount rate p life time n = 100 years

Figure 6 clearly illustrates previous finding that the discount rate p affects the total consequences $\kappa(k,p,n)$ more significantly than the number of escape routes k.

CONCLUSIONS

Similarly as in case of other technical systems the risk assessment of road tunnels commonly includes

- definition of the system
- hazard identification
- probability and consequences analysis
- risk evaluation and possible risk treatment

Two kinds of criteria commonly applied in the risk assessment of road tunnels relate to:

- expected individual risk
- cumulative social risk (fN curves)

Probabilistic risk optimization based on the comparison of social and economic consequences may provide background information valuable for a rational decision concerning effective safety measures of road tunnels. It appears that the discount rate and assumed life time may affect the total consequences and the optimum arrangements of the tunnels more significantly than the number of escape routes. However, further investigations of relevant input data concerning social and economic consequences are needed.

ACKNOWLEDGEMENT

This study is a part of the project Isprofond 5006210025 „SAFETY OF ROAD TUNNELS IN ACCORDANCE WITH DIRECTIVES OF EU“ supported by National fond of traffic infrastructures.

REFERENCES

[1] Directive 2004/54/EC of the European Parliament and of the Council of 29 April 2004 on minimum safety requirements for tunnels in the trans-European road network. Official Journal of the European Union L 201/56 of 7 June 2004.

[2] Melchers R.E. *Structural reliability analysis and prediction.* John Wiley & Sons, Chichester, 1999, 437 p.

[3] Steward M.S. & Melchers R.E. *Probabilistic risk assessment of engineering system.* Chapman & Hall, London, 1997, 274 p.

[4] Holický M., Šajtar L. Risk Assessment of road tunnels based on Bayesian network. *Advances in Safety and Reliability, ESREL 2005*. Taylor & Francis Group, London, 2005, pp. 873-879.

[5] NS 5814, Requirements for risk analysis. 1991.

[6] CAN/CSA-Q634-91 Risk analysis requirements and guidelines. 1991.

[7] ISO 2394 General principles on reliability for structures. 1998.

[8] ISO/IEC Guide 73: 2002, Risk management – Vocabulary - Guidelines for use in standards.

[9] ISO/IEC Guide 51: 1999, Safety aspects – Guidelines for their inclusion in standards.

[10] ISO 9000: 2000, Quality management systems – Fundamentals and vocabulary.

[11] Vrouwenvelder A., M. Holický, C.P. Tanner, D.R. Lovegrove, E.G. Canisius: CIB Report. Publication 259. Risk assessment and risk communication in civil engineering. CIB, 2001.

[12] Worm E.W. Safety concept of Westershelde tunnel, rukopis článku poskytnutý firmou SATRA v březnu 2002.

[13] Brussaard L.A., M.M. Kruiskamp and M.P. Oude Essink. The Dutch model for the quantitative risk analysis of road tunnels. ESREL 2004, Berlin, June 2004.

[14] Vrouwenvelder A.C.W.M. and Krom A.H.M. Hazard and the Consequences for Tunnels Structures and Human Life. 1st International Symposium Safe and Reliable Tunnels in Prague, CUR, Gouda, The Netherlands, 2004.

[15] Weger D. de, M.M. Kruiskamp and J. Hoeksma. Road Tunnel Risk Assessment in the Netherlands - TUNprim: A Spreadsheet Model for the Calculation of the Risks in Road Tunnels. ESREL 2001.

[16] Ruffin E., P. Cassini P. and H. Knoflacher. Transport of hazardous goods. See chapter 17 of Beard A and Carvel R (2005). The Handbook of Tunnel Fire Safety. Thomas Telford Ltd, London, 2005.

[17] Knoflacher H. and P.C. Pfaffenbichler. A comparative risk analysis for selected Austrian tunnels. 2nd International Conference Tunnel Safety and Ventilation, Graz, 2004.

[18] Jensen Finn.V. Introduction to Bayesian networks. Aalborg University, Denmark, 1996.

MODEL OF A RELIABILITY FOR STRUCTURAL - COMPLICATED SYSTEMS, INCLUDING MULTYSTATE ELEMENTS

Melnikov V.A.

Abstract. The problem of development of Boolean models of a reliability for systems, including elements with many states is considered on the basis of multivalued logic, algebra of trains, algebra of groups of incompatible events and classical logistic-probabilistic method (LPM). The inexpediency of development of Boolean models of a reliability on the basis of multivalued logic is displayed. The numerical examples demonstrating serviceability of LPM and their new possibilities are demonstrated. The perspective of development of methods of an evaluation of effectiveness of operation at different levels of operation rate by formulation of a set of different tasks, solved by the same LPM is underlined.

Key words: monotone Boolean model of a reliability, Boolean two-state algebra of logic, K-valued algebra of logic, multivalued logic, algebra of trains, algebra of logic of groups of incompatible events, LPM.

In § 3.1 of [1], written by Kurt Rainshke, is declared, that s alternative "all or nothing " for investigations of a reliability of complicated engineering systems and devices appears too rough. Therefore engineers in the practice would like to use more fine division of levels of operation rate. In this fact the experts in a reliability theory in many countries are working in the field of creation of models of reliability in which monotone Boolean models of reliability are developed on a multivalued case. The reference to nine publications of the foreign authors using idea of multivalued logic is given.

In the foreword of the book [2] the authors, answering Kurt Rainshke on his criticism [3] of so-called "shortages" of LPM, considering only two states of elements of the system, underlined, that they are irrelevant with basic singularities of these methods, but reflected only rate of the development of the theory and practice of those years [4].

The classical Boolean algebra of logic envelops all binary - discrete world, in which the arguments X_i of Boolean functions of algebra of logic (FAL) accept values from a two-element set {0, 1}. L.Kroneker (1823-1891) wrote: " A Zero and unit from the god, remaining matter of human hands ".

So by the "hands" of Y.Lukasevich (1878-1956) was created the first continual algebra of logic - multivalued logic, in which the limitation of Xi {0, 1} is taken off. On a interval
[X_{min}, X_{max}] of a numerical axes the minimax operations of a conjunction - min (X1, X2), disjunction-max (X1,X2) and refusing - = **K-X**□, where under K. K.Rainshke understands a maximum level of operation rate [$\underset{x}{1}$, p.110], and Mc-Notan - [5] 2X0 =2 [0,5 [X_{min}+ X_{max}].

Without objecting the indicated generalizations and attempts of " magnifications of a potency " of logistic-mathematical methods in information processes of analog area [5], there is no need to introduce into use more complicated mathematical methods, using continual algebra of logic in tasks of a reliability (safety) of structural - complicated systems. Examples in [1], demonstrate the absence of actuality of presented situations and practical impossibility of substitution of logic variables by probabilities.

So in an example 3.3., called to show multivalued model of a reliability of a system, consisting of four elements, (each element can be in three states), the restricted operation rate of the first level is the same that refusals of not designated other elements:

X1 (Cpu) - refusal in additional devices of real time;

X2 (interface) - refusal of the output channel;

X3 (extended memory) - refusal of a HDD;

X4 (controller) - refusal of the device of a parallel printing.

On our opinion this system consists not from four elements but from eight, each of which is described by Boolean algebra of logic.

In the fact that in real life there are elements with three and more states, we shall overview methods and history of a research of a reliability of such systems without use of multivalued logic.

In the book [6, p. 165] there is a brief analysis of eleven publications since 1956 about a problem of three incompatible states in a reliability theory, without considering of 14 publications B.Dillon himself. In these publications the means of a calculus, alarm graphs, polynomial expansions for systems with elementary structures was used. For a case of systems of a bridge type B.Dillon offered the method of transformation of the "triangle" by the "star". As a result of such transformation the bridge structure was substituted by a system with sequential and parallel junction of elements.

In [7, p. 173] the reliability of systems, consisting of elements with three states, is considered by means of classical LPM. That function of serviceability of a system (FSS), is formed with the help ***m*** parallel SPSO (shortest paths of successful operation), and everyone SPSO- is a series connection of ***n*** arguments X1. In this case the fundamental regularities possible, thanks to polynomial expansion $(R_i + Q_{i0} + Q_{i3})^n$ to all possible hypotheses, allow to evaluate separately a refusal of a system as "cut off" ($Q_{c.o}$) and as " short circuited" ($Q_{c.з}$).

The refusal of a system as "abruption" is evaluated by LPM by substitution of logic variables Xi by the appropriate probabilities by rules:

$$Xi = \begin{cases} 1, \text{ if i-element is efficient;} \\ 0, \text{ if i-element has given up as " cut off ".} \end{cases} \qquad (1)$$

The refusal of a system as " closure " is evaluated by same FSS, but the rules of substitution of a truth Xi will be inverted:

$$Xi = \begin{cases} 1, \text{ if i-element is “short circuited”;} \\ 0, \text{ if i-element is efficient.} \end{cases} \qquad (2)$$

The authors of [8] came to a conclusion, that the offered LPM based mode of account of a reliability of complicated systems, with elements, which can be in three incompatible states, allows to apply logistic-probability methods without using of the formulas:

$$R_{посл.} = \prod_{i=1}^{n}(1-Q_{io}) - \prod_{i=1}^{n}Q_{i3}, \qquad (3)$$

$$R_{пар} = \prod_{i=1}^{m}(1-Q_{i3}) - \prod_{i=1}^{m}Q_{io}, \qquad (4)$$

and transformations "star - triangle", and also complicated methods such as algebra of trains [9].

The author of algebra of trains (AT) B.Kulik. wrote [9, p. 19]: " Now, apparently, it is difficult to estimate the effect of AT in a solution of problems, connected with the reducing of complicated work while using of some LPM algorithms, but with the help АК it is possible to simplify statement and solution of such tasks, when the system consists of elements, for which the set of possible events is not restricted only by two states (for example, " the operation - refusal ") and supposes any additional set of intermediate states or events ". In the paragraph 4.4 [9] " Logistic-probability models at any number of states of elements " the numerical example for a bridge circuit from five elements is given. Four elements are submitted by three probabilities. They form a complete group of incompatible events. Taking into account difficulties, connected with small printing [9], and high value of a numerical example with the answer, we shall give FSS (5) and table of input data:

$$У(X_1,\ldots,X_5) = X_1X_3 \vee X_2X_4 \vee X_1X_5X_4 \vee \mathbf{X_2X_5X_3} \qquad (5)$$

X_1		X_2			X_3			X_4			X_5		
a_1	a_2	$в_1$	$в_2$	$в_3$	c_1	c_2	c_3	d_1	d_2	d_3	l_1	l_2	l_3
0,6	0,4	0,5	0,2	0,3	0,7	0,2	0,1	0,4	0,2	0,4	0,4	0,3	0,3

The answer will be Rc = P {y (X_1, …., X_5) = 1} = 0,86988.

We think, what not any expert in a reliability theory will manage to get the answer of this task, in spite of the fact that it seems very simple. The professor A.Mojaev has received this solution on the basis of general LPM (GLPM), combining classical algebra of logic and algebra of **logic of groups of incompatible events** (GIE) [10, p. 136].

In [10] in a conclusion the characteristics of some directions of the further development GLPM is given, where literally is told: " The methods of the account of GIE for the first time allowed to take off one of the most heavy limitation all LPM – requirement of dual representation of states of elements of investigated systems. Now, with the help GLPM, it is possible to get monotone and not monotone models of such systems, in which the elements can be not only in two, but also three, four and generally in any final set of incompatible states … ".

Agreeing with a conclusion in [8] about difficulties in understanding of algebra of trains by the engineers, it is necessary nevertheless to admire scientific value of this algebra [9]. For other case illustration of real difficulties in understanding and practical use of multivalued logic not only by engineers, but also by scientists, we shall notice publication [11], in which task from [8] was solved with the help of three-value logic. The comparison goes not for the benefit of three-value logic.

Standing up for the models, in which instead of one " of a level of operation rate " it would be possible to speak about "degrees of realization of the task", about "the effectiveness of operation " etc., in publication [1] the perspective declared only on the basis of multivalued logic. Rejecting this perspective in the beginning of paper as doubtful and more complicated, when the arisen problems solved on basis of LPM more simply and also more precisely, we shall illustrate now idea about an evaluation of the «effectivenesses of operation» at different « levels of operation rate».

If we have a real possibilities of formalization of these levels of operation rate, there is no reason to reject "«black-and-white" (« all or nothing ») model, and the set of such estimations will allow to see all gamma multicolor. In 1977 the professor I.A.Ryabinin wrote [12]: "Criticizing the "black-and-white" variant of a research (works - not works; yes - no; is true - is false) for it simplification, some experts stand up for multicolor model, in which ostensibly it is possible to take into account losses of even a part of percent of a system effectiveness. Frequently it appears, that such multicolored model cannot be checked up, it is not sufficiently determined and defined, and it seems that apparent multicolore is the same that irresponsibility ».

In [7] the example of a ship electric power system, consisting of 17 elements is given. 16 real tasks are formulated. FSS are placed in tables 31 and 32, all solutions at identical input data are given in the table 37.

Thus, the alternative «all or nothing », at its competent use, is not rough, and it is flexible enough, clear and responsible.

Most difficult and responsible in LPM is not an evaluation of probability function P {Y (X_1, …., X_5) = 1} for structural-complicated systems, but formalization of serviceability with the help of the shortest paths of successful (dangerous) operation -SPSO, (SPDO), which can be made by the highly experienced experts of the given subject field. "Automation" of the process of formalization with the help of raising in a degree of a matrix of nodal connections [13, p. 93], eliminations of intermediate knots [13, p. 96], solutions of a system of the logic equations [13, p. 100], is yet not less easy, than direct compiling of the list of SPSO, (SPDO) on the basis of common sense, which in the given context looks like engineering logic.

Thus, the **Boolean logic** helps **engineering logic** to formalize ideas about truth or false of our understanding of serviceability (danger) of structural-complicated systems .

The literature

1. K.Rainshke, I.A.Ushakov. A reliability estimation of systems with use of the graphs. M.: « Radio and communications », 1988, 208 p.

2. I.A.Ryabinin, G.NCherkesov. Logistic-probability methods of a research of a reliability of structural - complicated systems. M.: " «Radio and communications», 1981, 264 p.

3. K.Rainshke. Models of a reliability and responsivity of systems. M.: Mir, 1979, 452 p.

4. I.A.Ryabinin. Reliabitity of engineering systems: Principles and analysis. M: Mir, 1976, 532 p.

5. L.I.Volgin. Continual logistic-algebraic calculations as a basis of information process engineerings in analog area. // A radio electronics Engineering, Computer science, Control. Zaporozhye, 2000/2, p. 34-38.

6. B.Dillon, C. Singh. Engineering methods of security of a reliability. M.: Mir, 1984, 318 p.

7. I.A.Ryabinin. A reliability and safety of structural - complicated systems. Spb.: "Polytechnics", 2000, 248 p.

8. A.P.Kovalev, A.V. Spivakovsky. Application of logistic-probability methods for a reliability estimation of structural - complicated systems . // "Electricity", № 9/2000, p. 66-70.

9. B.A.Kulik. Logistic-probability methods and algebra of trains. // the Theory and information process engineering of simulation of safety of complicated systems .Spb.: IPMASH RAS. Issue 5, Preprint 123, 1995, p. 18-43.

10. A.S.Mojaev , V.N. Gromov. Theoretical basis of a general logistic-probability method of the automized simulation systems. // СПб. ВИТУ, 2000, 144 p.

11.A.S.Smirnov, D.O.Gaydamovich. The analysis of a reliability of the structural - complicated electrical circuits in view of two types of refusals. // "Electricity", № 2/2001, p.50-56.

12. I.A.Ryabinin. A reliability of a ship electric power supplying. // “The Sea collection ». №1. 1977, p. 79-82.

13. I.A.Ryabinin, Y.M.Parfenov. A reliability, survivance and safety of ship electric power systems. Spb., MNA, textbook, 1997, 430 p.

SCENARIO MANAGEMENT OF RISKS OF ACCIDENTS AND CATASTROPHES IN BUSINESS AND ENGINEERING

Solojentsev E.D.

●

Institute of Mechanical Engineering Problems of RAS,
sol@sapr.ipme.ru

Abstract. The stages of development of Management and Risk are described. The scenario management of risks of accidents and catastrophes in complex systems on the stages of designing, debugging and exploitation test and exploitation itself are considered. In the scenario management of accidents and catastrophes risks the personnel and the General designer are taken into account. The uniform approach to the modelling of risks in technical, economic and organisational systems is presented on the basis of substantial description of a SCENARIO of an accident or a catastrophe, and then the construction of models of the risk for the purpose of analysis and management. As the intellectual core for the risk quantitative evaluation and analysis and the scenario management of accidents and catastrophes risk, LP-methods and risk LP-models with groups of incompatible events are used.

Keywords: management, risk, accident, catastrophe, system, logic, probability, model, business, engineering

INTRODUCTION

The phenomenon of complexity of modern technical, economic, organisational and ecological systems, in our opinion, has not been cognised completely scientifically and has not been decided satisfactorily in applied sense. It makes us search for other approaches to the management of the accidents and catastrophes risk.

As the engineering discipline, the management of the accidents and catastrophes risks in complex systems is closely connected with the applied mathematics, as mathematics is those means, which help in most cases to make the adequate statement of a task, and also the precise formulation of conditions and assumptions, at which it is solved possible. However, the fundamental mathematics, based on proved theorems and strictly established laws, *does what it can, but as it must be done*. The applied mathematics, based on some hypotheses, experimental data and common sense, tries to decide problems *that are necessary, but as it can.* The era - w*hat is necessary and how it must be done* - for complex systems has not come yet [1].

Understanding the impossibility of an universal comprehension of the phenomenon of complexity of modern systems, we devote the book only to the questions of *scenario management* of the accidents and catastrophes risk in complex systems at all stages of their life cycle: at designing, debugging and exploitation tests and exploitation itself.

ACTUALITY

The risk consists of two components: the probability of an accident or a catastrophe and the damage. In risk tasks the probability and the damage are calculated by the models of different types: a probability model and an economic one. A risk decrease requires large expenses, and without these expenses, large losses (damage) are possible. Thus, the risk management includes the numerical estimation of a risk as the probability of an accident or a catastrophe and solving the optimisation task of the distribution of resources for the actions lowering the risk. We use the knowledge on the risk by two different ways: 1) passively, in the insurance against accidents fixing a price for the risk; 2) actively, in the management carrying out the actions lowering the risk of separate events.

The reason for accidents and catastrophes in complex systems, created and served by man, are mostly the mistakes at designing, testing and exploitation these systems. These mistakes are the consequence of both the imperfection of techniques and technologies of performance of certain types of work, and the limitation of resources allocated on this work. The problem of the maintenance of safe exploitation is aggravated by new task which has appeared everywhere: by estimation of the risk of prolongation of a resource of the worn out equipment by monitoring results. In complex systems existence various combinations of initial events-failures of separate elements is quite typical; the probability of each combination is small, but the sum of such "improbable" events is great.

The intellectual core of the scenario management of the risks of accidents and catastrophes in complex systems is made up by the logic and probabilistic methods (LPM) and the risk logic and probabilistic (LP) models with groups of incompatible events (GIE). These components form a different outlook of developers and users, induce them to consider a system as a whole and concentrate their efforts on the decision of paramount tasks, instead of to aiming the resources on minor needs and requirements. The ranking of the complex system elements according to their importance allows to increase the objectivity of distribution of the resources for decreasing the risk of accidents and catastrophes.

THE PRESENT STATE OF AFFAIRS AND HISTORY

Management and Risk have always existed since mankind appeared. Management was carried out taking into account the risk. It was based on intuition, experience and common sense and was empirical. Management provided the existence of man and the community. At later stages of the development of mankind, states appeared. Management was carried out by the Supreme governor of the country based on laws and religion. *The basis of such management in both society and engineering* has remained the same up to now. Later for increasing the effectiveness of management, people introduced some elements of the management mathematical theory and the optimisation mathematical theory into the practice of solving each separate task.

During the industrial revolution, *the classical theory of control* (regulation) by separate mechanisms, devices and processes, based on the description of dynamics of objects by differential equations, was created. The risk control was taken into account indirectly by the criteria of stability, possibility of resonant phenomena, destruction, etc. The success of the classical theory of control is enormous, for example, the control of the start and movement of a space ship.The scientists F.R.Bellman,H.Chestnut,R.W.Mayer,L.S.Pontryagin, Ya.Z.Tsipkin, etc made the most valuable contribution in creating the classical control theory.

In period 1939 - 1945 for the purposes of management such mathematical disciplines as Operations Research and Game Theory are appeared. The Operations Research is based on mathematical modelling of processes and phenomena and begins with the analysis of a criterion of efficiency of decision. The Game Theory is the theory of mathematical models of acceptance of optimum decisions in conditions of a conflict and uncertainty. The mathematical models had, as rule, discrete variables, were described by the system of the algebraic equations and had the economic contents.

During the Second World War for the purposes of management there appeared such mathematical discipline as *Operations Research* using the system approach to task statement and decision-making .Later this discipline almost completely became used only for the decision of optimisation tasks by methods of linear and non-linear programming. The methods for the decision of separate tasks of optimisation with the criteria of economic efficiency (the transport task, cutting the materials, etc.) were created.

Immediately after the Second World War *the control cybernetic theory* (Norbert Wiener) appeared. According to observable parameters at the input and output of an object the mathematical model of an object - "a black box" was constructed. Such control was used for the solving separate tasks of optimal control. The risk of such control was considered as the failure probability in the achievement of the purpose because of the non-adequacy of a model and the presence of white noise.

In 1952 years *the risk management science of investments* (Harry M. Markowitz [2]) appeared. For the first time the choice task of an optimum portofolio of valuable papers was formulated and decided. . “The models of averages and dispersions" were used. For each securities in a portofolio it took into account: the expected returns, as mathematical expectation, and risk, as standard deviation and uncertainty measure of expected returns. Such new concepts as the diversification, the indifference curves of the investor, achievable and effective sets of portofolioes were introduced. The contribution by Harry M. Markowitz was important and he got the Nobel premium for economy in 1990.

The portofolio optimization task of valuable papers had two criteria: expected returns and a portofolio risk. The normal law of distribution for expected returns of each paper and all portofolio was used.

The theory Harry M. Markowitz had such large success and development, that other approaches to risk research in business with discrete non-parametrical distributions of casual values were paralysed.

With the appearance of computers *information management* or management information automated systems came into being. These systems have: a well-structured database, an information technology with the window interface, software for solving a certain type of optimisation tasks, expert systems for decision-making, software for making reports and illustrations. It allows to give out any information for inquiries or use it for the decision of tasks quickly, to allocate the area of optimal allowable decisions, to choose the most effective decisions. The final decision is made by the main expert. In management information systems the tasks of a risk numerical estimation were not decided.

A new step in the development of the management theory was the appearance of *the situation management based* on logic-linguistic models ([3], D.Pospelov). It was shown, that the management of complex objects is absolutely impossible without using qualitative semantic sense information, which can not be expressed quantitatively. T*he logic, sets and logic connections of objects and events* were introduced into the theory and practice of management. The following was suggested: various ways of the description of observable situations, using the languages with developed semantics; various methods of the construction of knowledge models, allowing to reflect in these models qualitative ratio and laws; various procedures of decision-making in management on the basis of logic-linguistic models. A range of considered applications included the tasks of operatively-dispatching character for seaports, airports.The risk tasks in the systems of situation management were not decided.

Of much importance was the creation of *logic and probabilistic methods* ([1,4], I.A.Ryabinin) for quantitative modelling and analysis of reliability and safety of structurally complex technical systems. These logic and probabilistic methods (LPM) are a special section of mathematics, connected with logic-mathematical calculus . These methods allow to range the elements of a complex system according to their importance. These methods have passed approbation in the real projects of the naval fleet. They have become the intellectual core of management systems for reliability and safety in complex technical systems.

The development of logic and probabilistic methods was the creation of the generalised LPM ([5], A.C.Mojaev), which used all the logic connections (AND,OR,NOT) and introduced the schemes of the functional integrity. It has allowed to make the scenario for the successful or unsuccessful functioning of any technical or organisational system as the count using also the fictitious tops. It was created the program system for the for quantitative structure-logical analysis of the stability and effectiveness of structure-complex systems. It has successful used for years for training the students and solving the different applied tasks of analysis and management.

On the basis of logic and probabilistic approach *the risk LP-modelling and LP-analysis theory with groups of incompatible events* (GIE) ([6-8], E.Solojentsev) was created. It has allowed to model and analyse risks in the systems, which elements and the system itself have several states, and to apply LP-models with GIE for quantitative modelling and the analysis of the risk not only of technical, but also of economic, organisational and ecological systems. The states of elements in the systems were described both quantitatively and qualitatively, that is semantically.

In 1997 years the work on the strategy of risks management with he attraction of the new approaches from the area of fundamental sciences began. *The State Program: "Safety of Russia" has been* developed. In the book of the authors of this program "Management of risks" [9] special attention is paid to the problems of the risk management strategy. The authors' concept is the following. On the basis of accumulated experience a new science – the mathematical theory of safety and risk can be created. This theory must lie between the level of taking political decisions and strategic decisions as laws and the level of the development of concrete technical systems. As a methodical basis for creation of such a theory, they offer to use non-linear dynamics, the theory of bifurcation and chaos. The offered methodical base for the theory of risks is probably true for modelling of the earthquakes and the snow avalanches, but it is wrong for structure-complex human-machine systems.This statement is erroneous and it initialised the writing of the present book.

Scientific novelty of the curried out researches is in the following:

1. The uniform approach to the modelling and analysis of risks in technical, economic and organisational systems is discussed on the basis of substantial description of a SCENARIO of an accident or a catastrophe, and then the scenario-based construction of structural, logic and probabilistic models of the risk. It allows to realise the scenario management of risks in complex systems. The existence of different risk theories in various subject areas is not justified. Accidents and catastrophes, as a rule, are

caused by the combination of system defects (design, manufacture, organisation), exploitation defects (mistakes of the personnel).

2. As the intellectual core for the risk quantitative evaluation and analysis and the scenario management of accidents and catastrophes risk , LP-methods and risk LP-models with groups of incompatible events are suggested. The traditional "pure" mathematics and the non-linear mechanics try to explain everything by theorems. However still the greatest contemporary scientists J. Von Neumann and Norbert Wiener knew, that the complete confluence of theory and practice is indispensable; they could not be satisfied by cachetic concepts of mathematicians, lacking in practical value. Real complex technical, economic and social systems cannot be described with the help of differential equations.

3. In the scenario management of accidents and catastrophes risks the personnel and the General designer are taken into account. In such systems as nuclear power stations, starting rocket complexes and banks emergencies in general are caused by people. The role of the style, concepts and methods of the General designer in solving the risk tasks is quite great.. Out of a variety of new ideas he has to choose those which have the scientific and techno-logical substantiation and can be realised at given moment. He must not lose the general orientation and miss the most relevant details as well. Only by full-size tests and modelling one can became confident in the reliability of chosen decisions. The great Leonardo da Vinci (it is known) put forward a set of new ideas, including a helicopter, a parachute, and a ball-bearing. However the mankind needed five centuries to put them into practice.

4. The methods and models of scenario management of risk *at the stages of designing, debugging and exploitation tests and at exploitation* itself are described. The safety of complex systems, that are generally considered as complex human-machine systems and systems with various states of elements and the system itself, is formed at all the stages of the life cycle.

PRACTICAL SIGNIFICANCE

The appeal of scenario management of accidents and catastrophes risk for economists and engineers consists in its exclusive clearness, unambiguity and large opportunities for the analysis of influence of any element on the risk of accidents and catastrophes. The large number of examples is given, because examples teach sometimes better and faster than dry theory. Though complex mathematical methods are used in the book, it is intended for engineers and economists. All problems are discussed from the point of view of their practical application.

The work consists of several chapters, which have the following contents.

1. Technogenic accidence and catastrophes of the XX century. The data about great accidents and catastrophes, the most dangerous branches of industry, the amount of risk and damage of some objects are resulted. The accidents and catastrophes' sources are considered, the state safety program is presented, the methods of non-linear mechanics and LP-methods for modelling catastrophes are discussed. The stages of development of Management and Risk are described.

2. Men and risks. The frauds in business, the mistakes of attendants, asymmetrical actions of terrorists, hakker attacks on information networks, the position of personnel in the modern industrial civilisation are discussed.

3. Principles of risk management at designing. The style , concepts and methods of the general designer in describing the scenarios of accidents and catastrophes and risk management are discussed. General scientific knowledge, models and rules used in the area of risk are stated. Non-parametrical distributions of casual events are described; the essence of the Okkama Razor rule and the physical approach to the tasks of risk is explained; the scheme of risk management as a complex object is given, the task of minimisation of the number of decisions and the concept of acceptable risk is stated.

4. Risk management at debugging tests. The essence and the condition of debugging tests are described, the losses at debugging are given, the normative documents and the results of the debugging processes analysis are discussed, the principles of debugging management are stated, the debugging management scheme is described, the technology of debugging and its procedure is described, the scenarios of accidents and the example of the development of debugging tests program is supplied. A conclusion is made, that the mistakes one to the poor-quality of debugging tests and resources limitation are quite possible.

5. Risk management at exploitation tests. The technique of forecasting failures at exploitation tests is stated, the methodology of critical questions is described. The sources of critical questions are described, the methodology of using critical questions for working out the test program is described, the scenarios of failures and critical questions are given. The idea of the evolutional exploitation tests is described. The principles of the choice of working conditions for tests are presented.

6. Risk management at exploitation based on monitoring. The problem of destruction, deterioration and aging of the equipment in exploitation is considered. The role of monito-ring in society , economy, sports, medicine and engineering is discussed. The examples of the scenarios of failures and the monitoring of risk of functioning of a starting rocket complex and risk estimations of the worn out power equipment of resource prolongation are given.

7. The logic and probabilistic safety theory. The basic definitions of the safety theory are represented. The logic and probabilistic theory of the safety and the reliability is stated. The equations for all minimum paths of successful functioning system and for all minimum sections of system failures are presented. The methods of orthogonalization of logical functions are considered. The example of modelling and analysis of safety are given, the advantages of safety LP-models are considered.

8. The LP-risk theory with the groups of incompatible events. The scenarios of failure, events-signs and events–grades are considered, the basic equations are presented, the connection of the Bayes' formula and GIE is explained, the definition of the price for a risk is given, risk dynamic LP-models are described, estimations of accuracy and robustness of LP-risk models are made.

9. Identification of risk LP-models with GIE. The task of the LP-failure risk model identification based on statistical data is set. The methods of the LP-risk model optimisation/ identification are stated and investigated, the task of defining global extreme is set, the results of computer researches are discussed .

10. Analysis of risk. The purposes, methods and results of statistical and combinatorial risk analysis are given. The method of LP-risk analysis by the contributions of events-signs and events–grades to the object risk, to the average risk of a set of objects and to the accuracy of LP- risk model are presented. These methods of analysis are a base of risk management.

11. SOFTWARE for the risk analysis and management. The following is described here: intellectual AWS' for safety management, Software for identification and analysis of LP-risk models with GIE, Software for a structural-logic modelling of risks , Software for orthogonalization of L-functions based on cortege algebra.

12. Risk LP-models in business. A credit risk LP-model and the analysis results of the bank credit activity is given. The scenarios, the bribes LP-model and the frauds LP-models in business are considered. The management of the condition and development of a company according to the risk criterion is considered. The scenarios and the risk LP-models of the interaction of banks and companies and the risk LP-model of the loss of quality and the market are described.

13. Risk LP-models in engineering. The scenario and the LP-model of an explosion in a submarine are presented; the safety control system of a nuclear reactor is described; the task of the risk at the resource prolongation of the power equipment is discussed; the review of the known applications of risk LP-models in engineering is made.

14. Personnel and risks at a nuclear-dangerous plants. The unsolved problems are discussed; the approaches by Bernoulli and Columb to using the knowledge on risks are discussed. The peculiarities of financing the processes of the risk management are described, the parameter for the regulation of reliability of engineering and man is given, the account technique of natural and technogenic accidents is described, the risk of the poor-quality organisation of work is discussed.

The work is intended for experts and scientists working in the area of modelling, quantitative estimation and analysis of risk as well as the risk management in technical, economic and organisational systems *on the stages of designing, debugging and exploitation tests and exploitation itself.* It will also be useful for the students and post-graduate students of economic, financial and technical universities.

REFERENCES

1. Ryabinin I.A. *Relibility and safety of structure – complex systems*. Sankt-Petersburg, Politechnika, 2000, 360 p.
2. Harry M. Markowitz, «Portfolio Selection», Journal of Finance, 7, No1 (March 1952).
3. Pospelov D.A.. Logic and linguistic models in the systems of management . -

Moscow: Energoizdat, 1981.- 232 p.

4. Ryabinin I.A., Cherkessov G.N. Logic and probabilistic method of research of reliability of structure - complex systems . Moscow: Radio and Swyaz, 1981.
5. Mojaev A.S. Modern condition and some directions of development of logic-probabilistic methods of the analysis of systems // Theory and information technology of safety simulation for complex systems. Issue.1 -5 / Edited by I.Ryabinin and E.Solojentsev . St.Petersburg.: IPMASH RAS , 1994-95.
6. Solojentsev E.D., Karassev V.V., Solojentsev V.E. Logic and probabilistic models of risk in banks, business and quality / Edited by E.Solojentsev. St.Petersburg: Nauka,1999.-120p.
7. Solojentsev E.D., Karassev V.V. (2001) Risk logic and probabilistic models in business and identification of risk models. - Informatica 25 (2001) 49-55.
8. Solojentsev E.D., Karasev V.V. Identification of logic and probabilistic models of risk of structure-complex systems with groups of incompatible events. - Automation and Remote Control, No 3, 2002.
9. Risk Management (Stability development, Synergetic). Moscow: Nauka, 2000. – 431 p.
10. Kulik B.A. Representation of Logic Systems in a Probabilistic Space; in Term of Cortege Algebra. 1. Elements of Cortege Algebra.. 2. Measurable Logical Systems. – Automation and Remote Control, 1997, No1, p.126-136; 1997, No, p.169-179.

MANAGING AND MEASURING RISK IN TECHNOLOGICAL SYSTEMS

Romney B. Duffey
●
Atomic Energy of Canada Limited, Chalk River, ON, Canada
duffeyr@aecl.ca

John W. Saull
●
International Federation of Airworthiness, East Grinstead, UK
John@ifairworthy.fsnet.co.uk

Abstract
Safety Management is intended to create order out of disorder, to reduce the "information entropy", for the purpose of improved safety. Our purpose here and now is to try to introduce some predictability and insight into the risk or occurrence of apparently random events, where a general risk prediction we adopt a fundamental must be testable against the world's existing data. The risk management issues are clear, given the classic features of major human involvement and contribution to accidents, errors and outcomes occurring with modern technological systems. Prior incidents and prior knowledge and experience must be fully incorporated or learned from. If we do not know where we are on the learning curve, we also do not know the probability of such an event, and we have no objective measure of the "safety culture". Emphasis on defining and finding so-called "lack of safety culture" has resulted in an extensive and detailed study of the safety management and process safety of many global corporations. We utilize the concepts adopted in thermodynamics and Information Theory to establish the information entropy as a finite, physically based and useful measure of risk in technological systems. The results that we demonstrate show that the risk is dynamic, and can be utilized for management and predictive risk analysis purposes.

Keywords
Risk management; Information theory; Measurement; Safety culture; Prediction; Technological systems; Uncertainty; Probability

THE RISK PREDICTION PURPOSE

One simple worldview is that at least 90% of accidents, disasters and undesirable events (outcomes) are really due to management causes and issues, which we regard here as simply categorized as due to insufficient learning. Workers, organizations, corporations, investors and managers are all placed at risk from such events. To solve that problem, the attributes of a desired organizational "safety culture" have been defined and investigated in a number of ways, primarily based on structured surveys, interviews and questionnaires. The idea is to provide a qualitative measure or idea of how staff and management really feel and act about safety, which we regard here as some implied elimination of the error states. There are no equations and no theory: it is social science and psychometrics applied to safety.

Modern technological systems fail, sometimes with catastrophic consequences, sometimes just everyday injuries and deaths. The risk is given by the *probability* of failure, error or of any adverse outcome, and hence the *measure of risk is reflected in and by the uncertainty.* We have already examined the worldwide trends for outcomes (measured as accidents, errors and events) using data available for large complex technological systems with human involvement. We found and showed how all the data agreed with the learning theory when the accumulated experience is accounted for in Duffey and Saull (2002) [1]. Here, learning includes both positive and negative feedback, directly or indirectly, as a result of prior outcomes or experience gained, in both the organizational and individual contexts as in Ohlsson (1996) [2].

We introduce a measure of uncertainty to provide predictability and the needed insight into the risk or occurrence of these apparently random events. In seeking such a general risk measure, we adopt a fundamental theoretical approach that is and must be testable against the world's existing data.

THE MANAGEMENT CHALLENGE

Many solutions and major recommendations have been made for improving the safety management of process plants and their staff, e.g. BP Baker Panel Report [3]. Typically, the recommendations focus on how to minimize and manage the personal and business risk, paraphrasing and generalizing as follows:

- Corporate management must provide effective leadership on and establish appropriate goals for process safety;
- Establish an integrated and comprehensive process "safety management system" (SMS);
- Develop and implement a system to ensure that all management and managers, supervisors, workers, and contractors, possess an appropriate level of process safety knowledge and expertise;
- Develop a positive, trusting, and open process safety culture;
- Clearly define expectations and strengthen accountability for safety performance at all levels;
- Provide more effective and better coordinated process safety support for line management;
- Develop and implement, maintain an integrated set of leading and lagging performance indicators for more effectively monitoring safety performance;
- Establish and implement an effective system to audit process safety performance;
- Senior corporate officials should monitor the ongoing process safety performance;
- Use the lessons learned from past outcomes and events to become a recognized industry leader in process safety management.

Management generally wants to do what is right, and Regulators particularly seem to like this type of approach, as it attacks the management failings in a hopefully non-threatening and constructive way. Safety culture surveys are aimed at the attitudes, beliefs, practices and norms that hopefully characterize a pro-active approach to improving safety. But the adage "you can only manage what you can measure" means there must still be an *objective* measure of risk. What we present here is to enable such general recommendations to become a specific safety reality.

What we propose and develop are the validated means, tools and methods for management to use to: manage risks; prioritise work and recommendations; objectively measure and report the state of learning and the "culture"; and provide the company a rational approach to try to actively *predict* progress and outcomes. In that sense, what we propose is to move away from reliance on qualitative surveys to special emphasis on the quantitative measurement of learning using the knowledge gained from experience.

WHAT WE MUST PREDICT

We manage the risk, but only if we include the human element. We have shown how all outcomes develop in phases from a string or confluence of factors too complex to predict but always avoidable. We now know that a universal learning curve (ULC) exists and we can utilize that to predict outcome rates and track our progress as we improve, based on the known probability of an outcome. We have shown that [4, 5] the risk probability is given by the classic result:

$$p(\varepsilon) \equiv F(\varepsilon) = 1 - e^{-\int \lambda d\varepsilon} \qquad (1)$$

where, from the Learning Hypothesis [1] at a given experience, ε, the failure rate, λ(ε) naturally includes the human element as given by:

$$\lambda(\varepsilon) = \lambda_m + (\lambda_0 - \lambda_m) \exp - k(\varepsilon-\varepsilon_0) \qquad (2)$$

We suggest, at least for the present, that it is practically *impossible* to try to describe all the nuances, permutations and possibilities behind human decision-making. Instead, we treat the homo-technological system (HTS) as an integral system. We base our analysis on the Learning Hypothesis, invoking the inseparability of the human and the technological system. Using the data, we invoke and use experience as the correct measure of integrated learning and decision-making opportunity; and we demonstrate that the HTS reliability and outcome probabilities are dynamic, simply because of learning.

The basic and sole assumption that we make every time and everywhere is the "learning hypothesis" as a physical model for human behaviour when coupled to any system. Simply and directly, we postulate that humans learn from their mistakes (outcomes) as experience is gained. So, the rate of reduction of outcome rate (observed in the technology or activity as accidents, errors and events) is proportional to the rate of the outcomes that are occurring.

That learning occurs is implicitly obvious, and the reduction in risk must affect the outcome rate directly. To set the scene, let us make it clear that the probability of error is quite universal, and can affect anyone and everyone in a homo-technological system (HTS). There are clear examples of highly skilled well-trained operators, fully equipped with warning and automated systems still making fundamental errors as an inseparable part of the technological system.

THE RISK PROBABILITY AND THE RATE OF ERRORS

Given the outcome rate, now we need to determine the outcome (error) probability, or the chance of failure. The hazard function is equivalent to the *failure or outcome rate* at any experience, λ(ε), being the relative rate of change in the reliability, R, with experience, 1/R(ε) (dR(ε)/dε). The *CDF or outcome fraction*, F(ε), is just the observed frequency of prior outcomes, the ratio n/N, where we have recorded n, out of a total possible of N outcomes. The *frequency of prior outcomes* is identical to the observed *cumulative prior probability*, p(ε), and hence is the CDF, so F(ε) = p(ε) = (n/N) = 1 - R(ε), where R(ε) is the *reliability,* 1-n/N, a probability measure of how many outcomes or failures did *not* occur out of the total.

The *future (or Posterior) probability*, p(P) is proportional to the Prior probability, p(ε) times the Likelihood, p(L), of future outcomes. The chance of an outcome in any small observation interval, is the PDF f(ε), which is just the rate of change of the failure or outcome fraction with experience, dp(ε)/dε. The *Likelihood,* p(L) is the ratio, f(ε)/F(ε), being the probability that an outcome will occur in some interval of experience, the PDF, to the total probability of occurrence, the CDF; and we can write the PDF as related to the failure rate integrated between limits from the beginning with no experience up to any experience, ε, as in equation (1).

We can also determine the maximum and minimum risk likelihoods, which are useful to know, by differentiating the resulting probability expression. The result shows how the risk rate systematically varies with experience and that the most likely trend is indeed given by the learning curve. In other words, we learn as we gain experience, and then reach a region of essentially no decrease, in rate or in probability, and hence in likelihood. It is easy to obtain the first decrease in rates or probabilities but harder to proceed any lower. This is exactly what is observed in transport, manufacturing, medical, industrial and other accident, death and injury data [1].

From the analysis of many millions of data points that include human error in the outcomes, we have been able to derive the key quantities that dominate current technological systems. These now include commercial air, road, ship and rail transport accidents; near-misses and events; chemical, nuclear and industrial injuries; mining injuries and manufacturing defects; general aviation events; medical misadministration and misdiagnoses; pressure vessel and piping component failures; and office paperwork and quality management systems.

From all these data, and many more, we have estimated the minimum failure rate or error interval, initial rate, λ_0, of 1/ε, or a typical initial error interval, initial rate, at small experience of about one per 20,000 to 30,000 hours ($\lambda_0 \sim 5.10^{-5}$ per hour of experience); and a minimum attainable rate, λ_m, at large experience, ε, of about one per 100,000 to 200,000 hours ($\lambda_m \sim 5.10^{-6}$ per hour of experience);
The learning rate constant for the ULC, k ~ 3, is derived from the fit of a mass of available data worldwide for accidents, injuries, events, near misses and misadministrations. The following numerical dynamic form for the risk rate is our "best" available estimate from equation (2), adopting $\lambda_0 = (n/\varepsilon)$, with n = 1 for the initial outcome [1,2],

$$\lambda = 5.10^{-6} + (1/\varepsilon - 5.10^{-6})\, e^{-3\varepsilon} \qquad (3)$$

The risk rate, λ, can be evaluated numerically, as well as the probability, p(ε), and the differential PDF, f(ε). The result of these calculations is shown in Figure 1, where ε ≡ τ units in order to represent the accumulated experience scale.

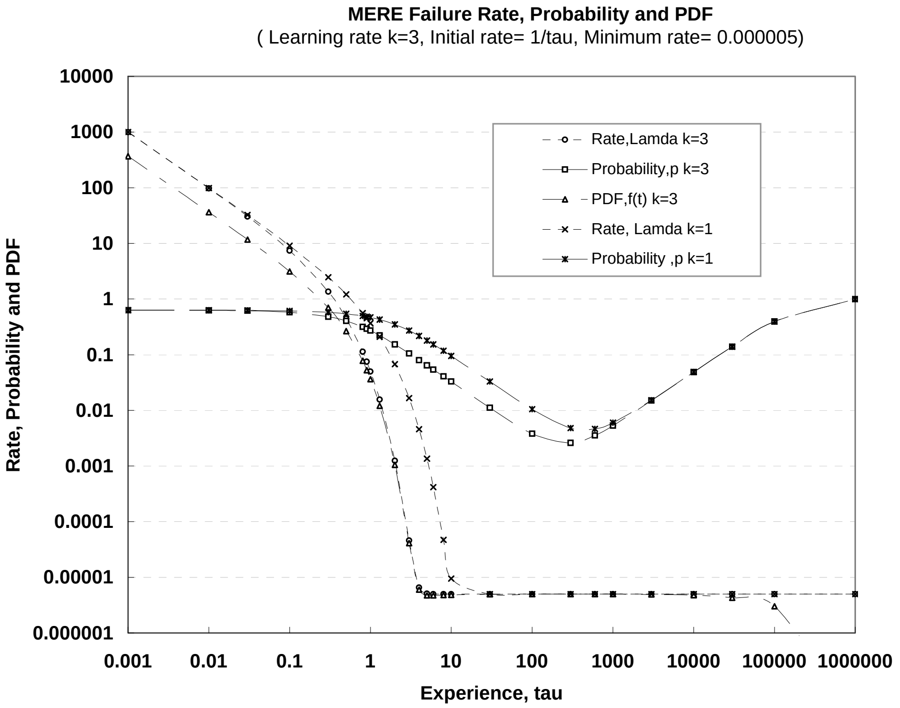


Figure 1. The best risk estimate with learning.

It is evident that for k>0 the probability is a classic "bathtub" shape, being just under near unity at the start (Figure 1), and then falling with the lowering of error rates with increasing experience. After falling to a low of about one in a hundred "chance" due to learning, it rises when the experience is ε > 1000 tau units, and becomes a near certainty again by a million tau units of experience as failures re-accumulate, since λ_m~5.10^{-6} per experience tau unit. The importance of learning is evident, since without learning there is no achievable minimum, which is the goal of management.

Our maximum risk is dominated by our inexperience at first, and then by lack of learning, and decreasing our risk rate largely depends on attaining experience. Our most likely risk rate is extremely sensitive to our learning rate, or k value, for a given experience.

So, as might be logically expected, the *maximum likelihood for outcomes occurs at or near the initial event rate when we are least experienced.* This is also a common sense check on our results: *we are most at risk at the very beginning.* Therefore, as could have been expected, the most likely and the least risks are reduced only by attaining increasing experience and with increased learning rates.
This approach to reduce and manage risk should come as no surprise to those in the education community, and in executive and line management positions. *A learning environment has the least risk.*

ORGANIZATIONAL LEARNING AND SAFETY CULTURE: THE "H-FACTOR"

Having examined the data and methodologies used to establish SMS, let us now return to the definition of "safety culture", which is where we started, and how it can be quantified.

Recall that the desiderata for the creation of a "safety culture", coupled to an organizational structure, places unending emphasis on safety at every level. But there is always a probability of error, a near- miss or of an event from which we must learn. We propose and prefer the use of the term and the objective of sustaining a "Learning Environment", where mistakes, outcomes and errors are used as learning vehicles to improve, and we can now define why that is true. We can manage and quantify safety effectively tracking and analyzing outcomes, using the trends to guide our needed organizational behaviors.

In the Statistical Error State Theory (SEST) [5] we found the variation in outcomes varied exponentially with depth of experience. Also the degree of order attained in a HTS was defined by "information entropy", or H-factor, the summation being a function of the probabilities of error state occupation.

The H-factor is well known in statistical mechanics where it is called the "uncertainty function", e.g., Greiner et al. (1997) [6], and in Information Theory where it is called the Shannon "information entropy", e.g., Pierce (1980) [7]. It has some key properties, namely: "as a fundamental measure of the predictability of a random event, which also enables intercomparison between different kinds of events". The H-factor is an objective measure of the *uncertainty*, and hence of the risk. This property is exactly what we would require to assess a SMS's effectiveness in reducing outcomes; and in assessing the risk in any given "safety culture".

In addition, the H-factor has the useful and necessary properties that for equally possible outcomes, $p(P) \sim 1/N$, and the (Laplace-Bernoulli) uniform prior presents the largest uncertainty, as we would expect. For a "sure thing" the H-factor is independent of the probability; and also satisfies the condition of additive probabilities for independent events. Its obvious application to safety management measurement is however *totally new* as presented here in Duffey and Saull [8], and arises quite naturally from the need for management to create order from disorder, and reduce uncertainty.

In terms of probabilities based on the frequency of microstate occupation, $n_i = p_i N_j$ and using Stirling's approximation we have the classic result for the Information Entropy:

$$H_j = - \Sigma\, p_i \ln p_i \qquad (4)$$

and the *maximum uncertainty value occurs for a uniform distribution of outcomes*. The corresponding probability of occupation as a function of experience:

$$p_i = p_0 \exp(\alpha - \beta\varepsilon_i) \qquad (5)$$

We note that since we have observed the outcomes, the usual normalization condition for all the N_j outcomes to exist is, summing the probabilities over all the j observation ranges, $\Sigma_j\, p_i = 1$. For the probability distribution of a continuous random variable, we can transform the sum to an integral. This normalization says simply that whatever outcomes happened must occur. *The risk always exists, somewhere in observational space*.

In practice, the probability of occupation according to the SEST is approximated by a fit to the available outcome data [8] given by:

$$p_i = p_0 \exp - aN^*, \qquad (6)$$

where, a, is a constant, and N*, the non-dimensional measure of the depth of experience, $\varepsilon/\varepsilon_M$. Thus, for our continuous probability function, we can evaluate the (so-called grand) partition function, and write the probability of error state occupancy as:

$$p_i = p_0 \exp(-aN^*)/\int_0^\infty p_0 \exp(-aN^*) \quad (7)$$

or,

$$p_i = (a/\varepsilon_M) \exp(-aN^*), \quad (8)$$

and hence the probability decreases as the learning rate and experience depth increases. Since the outcomes are represented by a continuous random variable learning curve, the Information Entropy, H, in any jth observation interval is also given by the integral:

$$\begin{aligned} H_j &= - \int p_i \ln p_i \, dp \\ &= p_i^2 (1/4 - ½ \ln p_i) \quad (9) \end{aligned}$$

So, substituting in the expression for the Information Entropy, H, which we term the *"H-factor"*:

$$H_j = ½ \{p_0 e^{-aN^*}\}^2 \{aN^* + ½) \quad (10)$$

where, on a relative basis, $p_0 = 1$, and then H → 0.25 as experience decreases as $N^* \rightarrow 0$. This parameter is a measure of the uncertainty, and hence of the risk.

As either the learning rate or depth of experience increases ($N^* \uparrow$or $a\uparrow$), or the zeroth order occupancy decreases ($p_0 \downarrow$), so does the value of the H-factor decline, meaning a more uniform distribution and increased order. We illustrate the variation in the relative information entropy, H, with non-dimensional experience depth, N*, in Figure 3, taking the zeroth probability as unity ($p_0 = 1$) for a range of learning rates. The range chosen varies around the "best" value of a = 3.5, which is as derived from the USA aircraft near-miss and Australian auto death data so that:

$$P_i = p_0 e^{-3.5N^*} \quad (11)$$

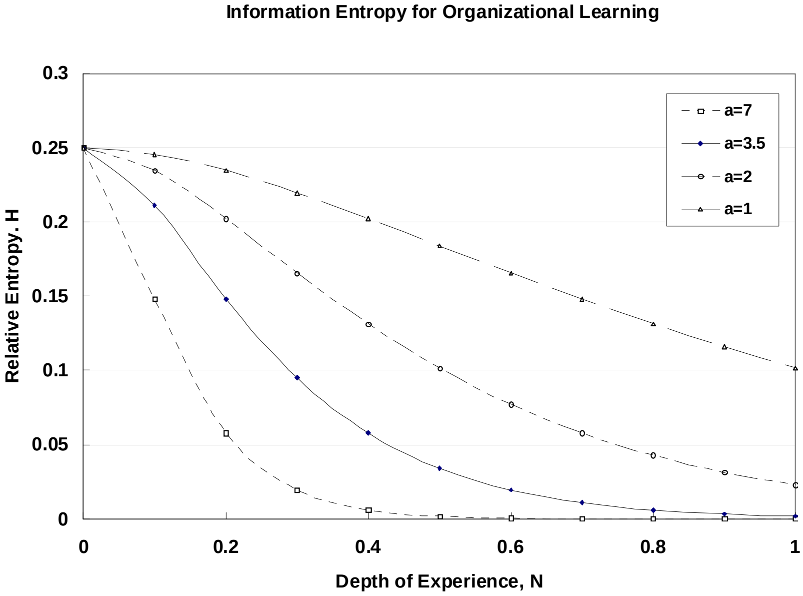


Figure 3: Organizational Learning and Experience

Clearly, the relative value of the information entropy H-factor at any experience depth is a direct measure of any cultural aspect of modern technologies called "organizational learning". This terminology is meant to describe the attributes of a HTS, and its ability to respond effectively to the demands for continuous improvement, as reflected in internal organizational and communication aspects.

The faster decline and decrease in the H-factor with increasing depth of experience and increasing learning constant is a reflection of increasing HTS organizational order. This is to be expected, and makes sense: it is exactly what safety management intends. This relational effect is also exactly what we mean by maintaining a Learning Environment, and has been derived from the new SEST analysis.

Before this discovery, all one could conduct were semi-empirical, qualitative and highly subjective comparative surveys of "organizational attitudes and beliefs". These would review opinions and attitudes to management, training and personnel systems without having a truly objective measure.

MANAGING INFORMATION AND SAFETY CULTURE

We can now argue that this purely theoretical concept of degree of order, the H factor, is actually a true practical and quantitative measure of the elusive quality termed "safety culture" by sociologists, human factors experts and industrial psychologists. Safety culture is therefore a reflection of the degree of order, and reduced uncertainty, attained in and by any HTS; and creating order is equivalent to reducing the risk probability of any outcome.

As we stated in the very beginning of this paper, it is management's expressed aim and intent in any technological system to create order from disorder, which it can only achieve by decreasing the information entropy. Unfortunately, most safety managers who are trained in traditional industrial safety methods, and corporate officers familiar to the world of business and accounting decisions and risks, would not recognize the concept of entropy, let alone information entropy, if they saw it. However, it is so simple to communicate the concept of the learning hypothesis and the impact on organizational learning, that it should be possible to obtain the management buy-in needed to adopt this approach to assess risk and safety.

Equally important to this quantification is the realization that this H-factor uses the actual outcomes as an explicit function of organizational learning and management effectiveness. We indeed do and can "manage what we can measure". This is simply common sense.

CONCLUSIONS: IMPLICATIONS FOR RISK MANAGEMENT AND PREDICTION

The implications of using this new approach for estimating risk are profound. This new probability estimate is based on the failure rate describing the ULC, which is derived from the Learning Hypothesis; and utilizes the validation from the outcome data of the world's homo-technological systems. For the first time, we are also able to make predictions of the probability of errors and outcomes for any assumed experience interval in any homo-technological system.

In addition the results implies a finite lower bound probability of based on the best calculations and all the available data. Analysis of failure rates due to human error and the rate of learning allow a new determination of the risk due to dynamic human error in technological systems, consistent with and derived from the available world data. The basis for the analysis is the "learning hypothesis" that humans learn from experience, and consequently the accumulated experience defines the failure rate.
The extension of the concept to "safety culture" shows this risk can be interpreted as uncertainty, and that uncertainty can be quantified by the Information Entropy, or H-factor. Management wish to emphasize "safety culture", which actually corresponds to a sustained Learning environment, managing risk creates order, reduces uncertainty and ensures predictability. Based on our theory and practical data, we have shown how to quantify order, reduce uncertainty and predict risk. The risk probability is based on experience.

The results demonstrate that the risk is dynamic, and that it may be predicted using the learning hypothesis and the minimum failure rate, and can be utilized for predictive risk management purposes.

REFERENCES

[1] Duffey, R.B. and Saull, J.W. (2002). Know the Risk, First Edition, Butterworth and Heinemann, Boston, USA.

[2] Ohlsson, S. (1996). Learning from Performance Errors, Psychological Review, Vol. 103, No. 2, 241-262.

[3] Report of the BP US Refineries Independent Safety Review Panel, (2007). The Baker Panel Report, available: www.bp.com.

[4] Duffey, R.B. and Saull, J.W. (2006). The Human Bathtub: Safety and Risk Predictions Including the Dynamic Probability of Operator Errors, Proceedings 14th International Conference on Nuclear Engineering (ICONE14), Paper No. 89476, Miami, Florida, July 17-20.

[5] Duffey, R.B. and Saull, J.W. (2005). The Probability of System Failure and Human Error and the Influence of Depth of Experience, Proceedings of International Topical Meeting on Probabilistic Safety Analysis (PSA'05), San Francisco, CA, September 11-15.

[6] Greiner, W., Neise, L. and Stocker, H. (1997). Thermodynamics and Statistical Mechanics, Springer, New York, pp. 150-151.

[7] Pierce, J.R. (1980). An Introduction to Information Theory Symbols, Signals and Noise, second Revised Edition, Dover Publications, Inc., N.Y.

[8] Duffey, R.B. and Saull, J.W. (unpublished). Managing Risk.

TRAFFIC ACCIDENTAS INFORMATION SYSTEM AND RISK CRASH EVALUATION

Jiri Stodola

●

University of Defence in Brno, Faculty of Military Technology,
PS 13, K 252, 612 00 Brno, Czech Republic
Tel/fax: +420 973 442 278, E-mail: jiri.stodola@unob.cz

ABSTRACT

This article analyses the traffic accident rate on roads and highways and possibilities of risk evaluation related to traffic accident occurrence based on factors that were the causes of accidents. A new term – risk of traffic accident occurrence is a product of probability of accident occurrence and its impacts. The results are presented by way of example that uses selected statistical data of the Czech Republic traffic accident rate between 1993 - 2001. The article provides a brief methodological procedure of evaluation of the traffic accident rate using the risk of traffic accident occurrence.

KEYWORDS

Traffic Accident, Risk Factor, Road Traffic Safety, Traffic Accident Consequence, Road Traffic Risk and Safety Evaluation

1. INTRODUCTION

Recently, the safety of road traffic has become a very serious problem for nearly all countries. A growing density of the traffic causes an increasing amount of accidents associated with heavy losses of property and injuries or fatalities. That is way national and international authorities pay an exceptional attention to this problem and try to mitigate the negative trends in the time development of safety of road traffic, for example by modifications of traffic regulations. To be rational and effective, the measures taken by the authorities have to result from a detailed analysis of the causes of accidents. For these reason usually the authorities in the developed countries maintain the national accident databases that gather the information on the consequences of each road accident.

2. TRAFFIC ACCIDENT STATISTICS

Data from accident databases enable to carry out the required analyses and establish the trends of the traffic safety. An analysis of time development of absolute or relative number of accidents and their consequences is the most common way of evaluation of the trends of the traffic safety. Examples of such a kind of evaluation for the Czech Republic are presented in tab.1 – 4, and fig.1 - 3. The time development of number of accidents with a certain cause or number of injuries, fatalities and amount of property damages associated with this certain kind of accident is beyond any despite useful indicator of development of safety, but sometimes the results of the above-mentioned analysis can be quite controversial. A certain weakness of this system is the fact that it employs absolute numbers that prevent comparison between the individual periods of month/day/year, causes, etc., and inaccuracy due to changes resulting from individual data changes. Substantial disadvantage of this system consists in a non-existence of a measure of severity, or acceptability of the traffic accident impacts. That is why it is not possible to determine whether the traffic accident is or is not socially acceptable, or, it is at least satisfactory.

3. RISK, AND DEGREE OF RISK

The given evaluation obviously lacks a common feature of risk level that consists in:

- Simultaneous consideration of features (risk factors) of each traffic accident,
- Probability of a certain traffic accident occurrence,
- Appropriate expression (evaluation) of traffic accident consequences,

- etc.

Table 1: Accidents and their consequences in the Czech Republic in the last 10 years

Year	Number of accidents	Number of fatalities	Number of severe injuries	Number of slight injuries	Damage (millions of €)
1993	152,157	1,355	5,629	26,821	93.4
1994	156,242	1,473	6,232	29,590	133.2
1995	175,520	1,384	6,298	30,866	152.4
1996	201,697	1,386	6,621	31,296	189.2
1997	198,431	1,411	6,632	30,155	186.9
1998	210,138	1,204	6,152	29,225	213.6
1999	225,690	1,322	6,093	28,747	223.4
2000	211,516	1,336	5,525	27,063	221.7
2001	185,666	1,219	5,493	28,297	257.6
2002	190,718	1,314	5,492	29,013	277.9

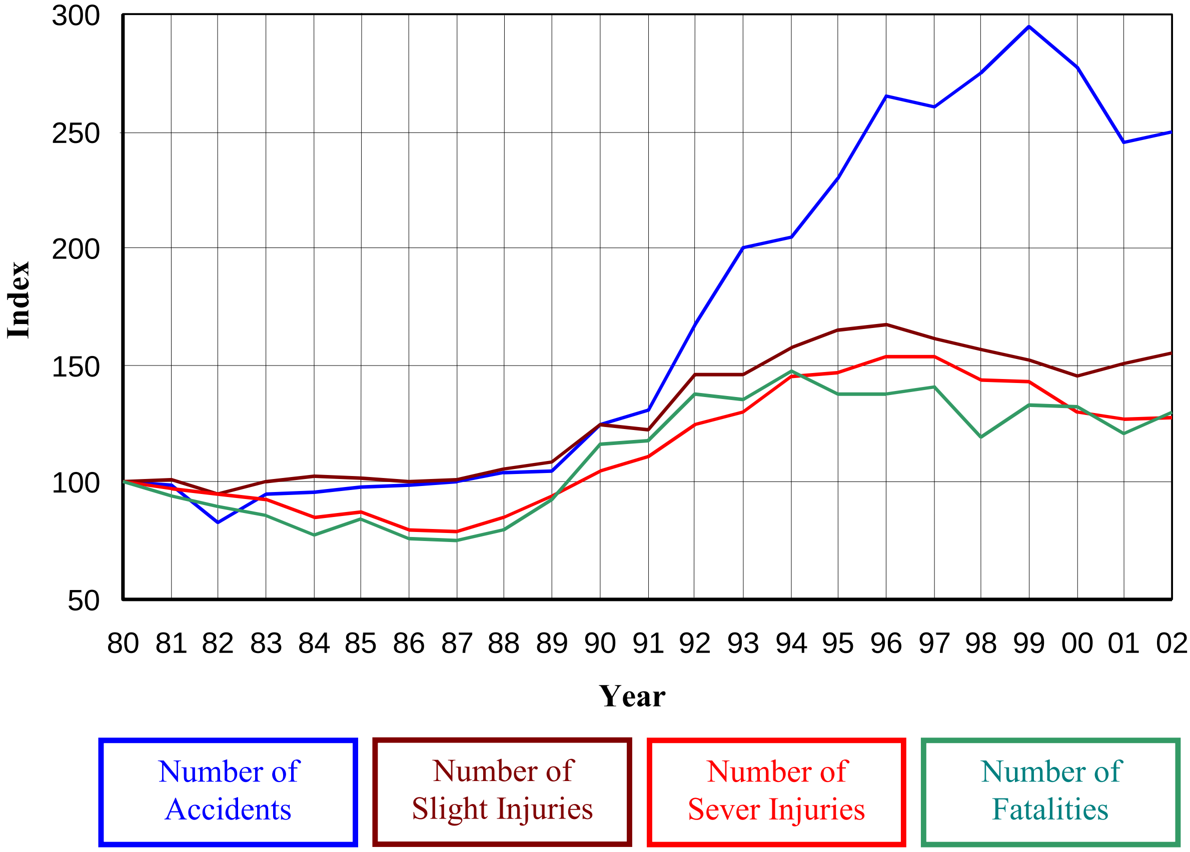


Figure 1: Statistics of accidents in the Czech Republic

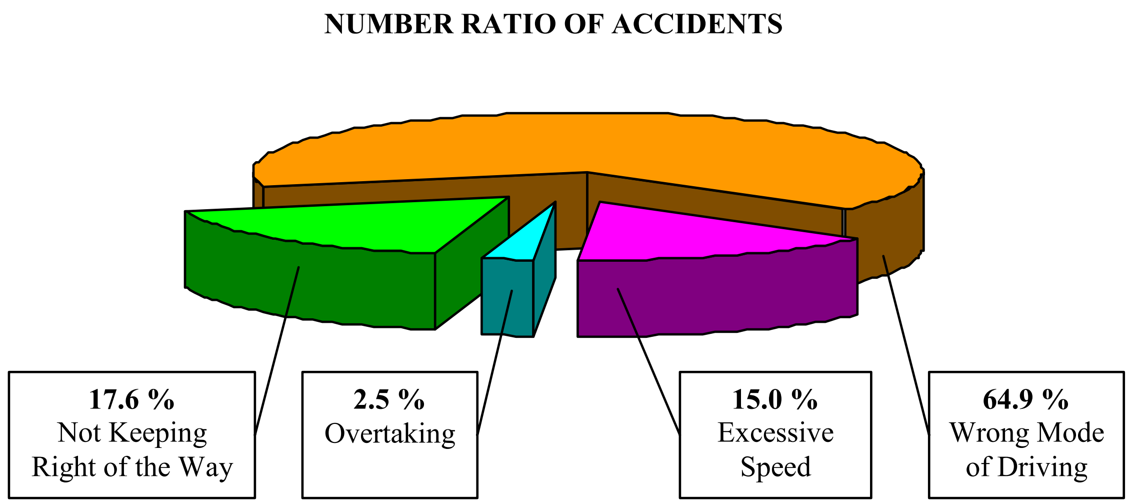


Figure 2: Main causes of accidents of drivers

Table 2: Magnitude of accidents caused by engine displacement

Car engine displacement year 2002	**Participation in number of accidents (in %)**	**Fatalities**		**Difference in number of fatalities (in comparison with year 2001)**	**Magnitude of accidents (Fatalities per 1,000 accidents)**
		Numb	%		
Up to 1 liter	4.8	49	5.6	-3	8.0
1.1 – 1.4 l	44.8	354	40.6	37	6.3
1.5 – 1.9 l	30.8	276	31.7	9	7.1
2 - 3 l	15.7	188	21.6	26	9.5
Over 3 l	0.7	4	0.5	-7	4.5
Undiscovered	3.2	0	0.0	-2	0.0

Table 3: Places of accident

Places of accidents year 2002	**Number of accidents**	**Number of fatalities**	**Number of severe injuries**	**Number of slight injuries**	**Damage (millions of €)**
In Municipality	139,345	501	2,886	17,689	171.5
Index (year 2000 = 100%)	103.0	110.1	102.9	103.7	108.2
Off Municipality	51,373	813	2,606	11,324	106.3
Index (year 2000 = 100%)	102.1	106.4	96.9	100.8	107.3
Motor Way	4,293	51	120	525	15.9
Index (year 2000 = 100%)	102.8	127.5	114.3	107.8	109.5

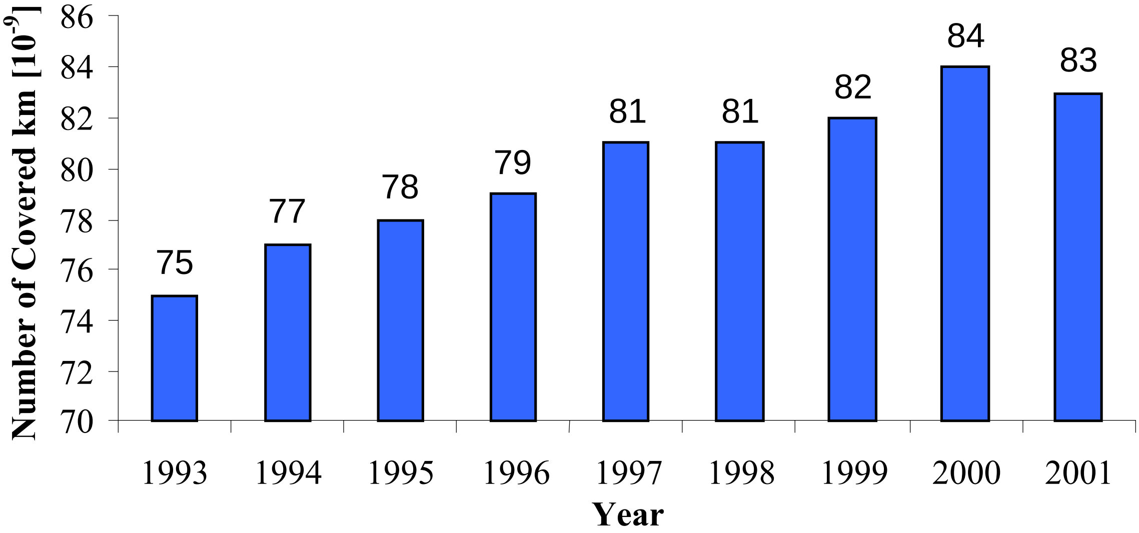


Figure 3: Annual number of kilometers covered by all vehicles in Czech Republic in one year

Table 4: Evolution of number of accidents, fatalities and injures in the last 10 years

Year	Number of accidents	Number of fatalities	Fatalities per 1 accident	Number of severe injures	Number of severe injures per 1 accident	Number of slight injures	Number of slight injures per 1 accident
			10^{-3}		10^{-3}		10^{-3}
2002	190,718	1,314	6.89	5,492	28.8	29,013	152.13
2001	185,666	1,219	6.57	5,493	29.59	28,297	152.41
2000	211,516	1,336	6.32	5,525	26.12	27,063	127.95
1999	225,690	1,322	5.86	6,093	27.00	28,747	127.37
1998	210,138	1,202	5.72	6,152	29.28	29,225	139.08
1997	198,431	1,411	7.11	6,632	33.42	30,155	151.97
1996	201,697	1,386	6.87	6,621	32.83	31,296	155.16
1995	175,520	1,384	7.89	6,298	35.88	30,866	175.85
1994	156,242	1,473	9.43	6,232	39.89	29,540	189.07
1993	152,157	1,355	8.91	5,629	36.99	26,821	176.27

It is evident that there exists a more complex evaluation using the institute of risk R in the following form:

$$R = P \times C,$$

where R … risk of traffic accident, P … probability of traffic accident occurrence, C … consequence of traffic accident.

A risk defined in this way is a non-dimensional parameter and it provides mutual comparison of various groups of causes of traffic accidents, their characteristics and it also enables mutual comparison of individual types of traffic accidents. To enumerate the risk of traffic accident according to this equation it is necessary to quantify probability of accident occurrence P and consequences of traffic accident C. Possibilities of that are presented in paragraph 4 and 5.

The second way in which it is possible to evaluate risks associated with traffic accident is usage of rate of accidents that represent probability of accident per one kilometer with respect to fatalities, severe injuries, and slight injuries. In the case of evaluation of damage caused by accident it is suitable to use co called specific damage that represents average damage per one kilometer. Equations for rate of accident and specific damage enumeration are presented in table 5.

Table 5: Equations for evaluation of rate of accidents and specific damage

Rate of accidents	**Rate of fatalities**	**Rate of severe injuries**	**Rate of slight injuries**	**Specific damage**
$R_A = \frac{N_\Sigma}{D_C}$	$R_F = \frac{N_F}{D_C}$	$R_{SEI} = \frac{N_{IS}}{D_C}$	$R_{SLI} = \frac{N_{ISL}}{D_C}$	$R_D = \frac{N_D}{D_C}$

where D_C … distance covered in the Czech Republic in calendar year, N_F … number of fatalities in calendar year, N_{IS} … number of severely injured people in calendar year, N_{ISL} … number of slightly injured people in calendar year, N_D … sum of damages.

Next possibility of risk evaluation is usage of so called degree of risk D_R that can be expressed by the following equation:

$$D_R = \frac{C_{Ai} N_\Sigma}{C_{A\Sigma} N_i} ,$$

where C_{Ai} … number of consequences by given cause of accident, $C_{A\Sigma}$ … number of consequences by all accidents, N_Σ … number of all accidents, N_i … number of accidents by given cause of accident.

Degree of risk D_R indicates how many times the given cause of accident is more risky than statistically significant average cause of an accident.

4. PROBABILITY OF TRAFFIC ACCIDENT OCCURRENCE

Probability of traffic accident occurrence encompasses a complete system of phenomena and, using a classical definition, equals to the probability share of frequency of specific type of traffic accident and total amount of traffic accidents in the period under survey. Probability of traffic accident P can be expressed in the following equation:

$$P = \frac{N_i}{N_\Sigma} ,$$

where N_i … number of accidents of evaluated i-th type in calendar year , N_Σ … total number of accidents in calendar year.

To determine this probability we can use sufficient credible data in the statistics of the traffic accident rate. Classical probability defined in this way shall be valid exactly in two-status model and its constraints rest in a necessity or assumption of similar possibilities of occurrence of random events – e.g. types of traffic accident. In practice, it may often happen that random event– type of traffic accident is not definite and may not happen anyway. There are possibilities of more generally approach to a probability, in practice - an axiomatic, or, in our case - statistic approaches are used.

5. IMPACT OF TRAFFIC ACCIDENT

The impact can be considered a measure of the traffic accident severity. It is a significant part of magnitude of risk. Here exists a general inversion principle based on the fact that an accident with a high level of probability of occurrence, but with non-serious impacts has also a low level of risk rate. And vice versa, an accident even very improbable but with serious impacts is considered as highly risky. To date, no transport standards provide a unique method of evaluation of the impacts of traffic accidents. In general, the impact of traffic accident can be established using two methods:

1) Use of expert methods when a severity level can be attributed to each accident as a relative value of accident impact with a meaning of weight, e.g., within the range of values from the interval: $0 \leq C \leq 1$, with possible interpretation: with no impacts $C\,min \rightarrow 0$, catastrophic impacts $C\,max \rightarrow 1$.

2) Use of international standards when severity of single categories of accidents is established by a scale - Minor, Major, Critical, Catastrophic, with exact definition of severity of individual categories. In some domains (e.g. aviation, etc.) for each category there exists a maximum value of socially acceptable probability of accident occurrence (Table 6).

3) Expressed impact is a tool with similar meaning as probability; to assess the impact of traffic accident it is possible to use a probability when the traffic accident impact is expressed, for example, by the number of persons killed at the type of traffic accident examined against the total number of persons killed in all accidents in the period under survey. Thus, a severity of a given type of traffic accident if „weighted" relative to other accidents by the weight of number of persons killed, or by other „weight", e.g., a property damage as a proportion of the magnitude of resulting property damage of the participants of the accident at the type of traffic accident relative to the total property damage of the participants of all traffic accidents in the period under survey.

Table 6: Hazard Severity Categories

Description	Category	Definition
Catastrophic	I	Death and/or vehicle loss.
Critical	II	Severe injury, and/or major vehicle damage.
Marginal (major)	III	Minor injury, and/or minor vehicle damage.
Negligible (minor)	IV	Less than minor injury, and/or vehicle damage.

6. EXAMPLE OF APPLICATION

Some results of calculation are given in table 7 and figure 4. Figure 5 show forecast of number of fatalities in the course of accidents per one million inhabitants and compare the Czech Republic and Great Britain, the Netherlands and Sweden.

Resulting from statistical data the rate of accident, fatalities, severe and slight injuries, specific damages, and degree of risk were evaluated. From results presented it is evident that the most risky factors in the Czech Republic are as follows:

- hitting the oncoming vehicle during overtaking,
- riding a motorcycle,
- pedestrian on the road,
- excessive speed.

Resulting from the analysis there can be stated that are the following most hazardous factors: wrong overtaking, higher than permissible driving speed, riding a motorcycle, and pedestrian behavior. Analyses showed that low level of alcohol in blood does not significantly increase the traffic accident risk.

Table 7: Rate of accidents, fatalities, injuries, and specific damage

Year	Rate of				Specific damage
	Accidents	Fatalities	Severe injuries	Slight injuries	
	$10^{-6} \cdot km^{-1}$	$10^{-8} \cdot km^{-1}$	$10^{-8} \cdot km^{-1}$	$10^{-7} \cdot km^{-1}$	$€ \cdot km^{-1}$
2002	2.3	1.6	6.6	3.5	0.0034
2001	2.2	1.5	6.6	3.4	0.0031
2000	2.5	1.6	6.6	3.2	0.0026
1999	2.8	1.6	7.4	3.5	0.0027
1998	2.6	1.5	7.6	3.6	0.0026
1997	2.4	1.7	8.2	3.7	0.0023
1996	2.6	1.8	8.4	4.0	0.0024
1995	2.3	1.8	8.1	4.0	0.0019
1994	2.0	1.9	8.1	3.8	0.0018
1993	2.0	1.8	7.5	3.6	0.0012

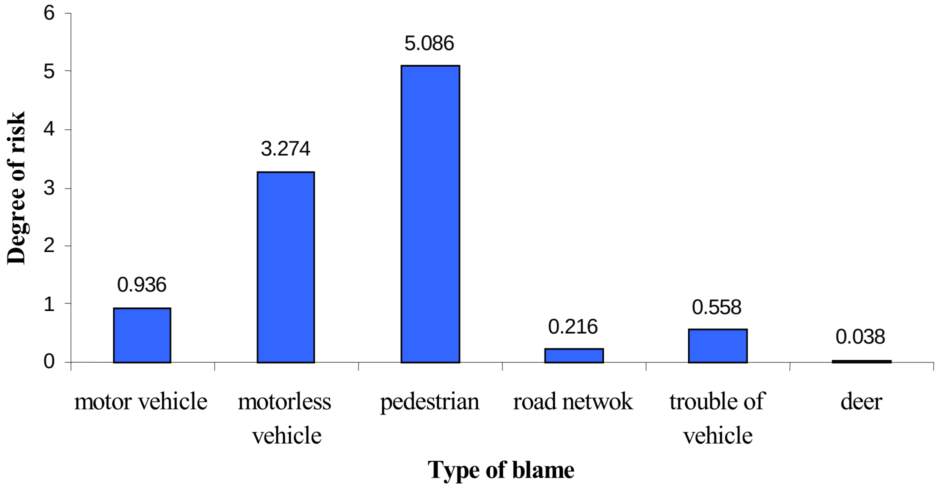


Fig. 4: Degree of risk due to blame

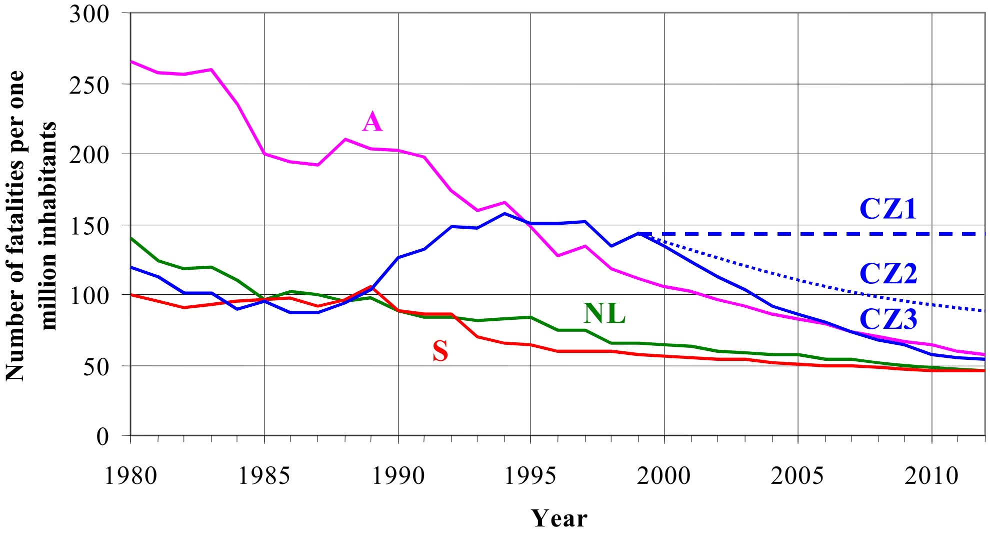


Figure 5: Forecast of number of fatalities in the course of accidents per one million inhabitants

Description of Figure 5:

A - Great Britain,
NL - the Netherlands
S - Sweden
CZ1 - Zero variant is the extrapolation of accident frequency development as in progress in the last seven years
CZ3 - Desirable variant, i. e., the accident level in the Czech Republic striving for the situation in the developed EU countries in the real-time horizon.
CZ2 - Hypothetical variant expresses the compromise between two above mentioned variants.

7. CONCLUSION

The Police of the Czech Republic maintain annual detailed statistics of the traffic accidents in the form of summary numbers and figure surveys divided by various criteria. This system provides important information that may serve as the grounds for creation of new and effective preventive measures. However, this information system does not use the institute of risk in the road traffic. And at the same time it is evident that trends in development of risk provide relatively objective and complex information to solve these traffic accidents as serious all-society phenomena. It refers mainly to the causes and consequences of traffic accidents and influence of various factors that determine the traffic accident rate.

Described methodology defines selected terms as objective tools for the systems analysis of causes and impacts of the traffic accidents. Risk of traffic accident rate is a non-dimensional parameter that can enable comparison of various effects and circumstances otherwise incommensurable. The advantage is that we can use existing statistical surveys and alternatively evaluate the safety of the road traffic. A certain disadvantage is the fact that in road traffic field there are no generally binding criteria of social or individual acceptability of the magnitude of risk related to the traffic accident. That is why the information system of the traffic accident rate cannot be used to establish whether the Czech Republic traffic accident rate is at an acceptable level, or whether it is necessary to reduce it.

REFERENCES

[1] CHUDOBA, J.: Determining of Probability of Car Accident by Conveyance of Car. Project. University of Liberec. 31 pages . 2002
[2] HAJEK, M.: Solution of problems verbal communication between participants of road traffic by risk situations in traffic. Technical university Ostrava, 2003 (76 pages)
[3] Czech Traffic Police Force.: Statistics of the Traffic Accident Rate in the CR, Set of Documents from 1993 to 2002.1, 1, Ministry of Interior, Praha, 1993 – 2002
[4] VINTR, Z.-HOLUB, R.-VALA, M.: A Risk-based Evaluation of Safety Development in Road Traffic. In: Probabilistic Safety Assessment and Management (PSAM 6) – Proceedings of 6th International Conference on Probabilistic Safety Assessment and Management. Oxford: Elsevier Science, p 493 – 498, 2002. ISBN 0-08-044122-X
[5] STODOLA, J.: Risk of Traffic Accident and Possibilities of its Evaluation. Žilina, ISBN 80-8070-121-0 (pages 127 – 130) 2004
[6] STODOLA, J.: Safety of Traffic of Motor Vehicles. Military Technical Magazine. Nr 6. Praha, 1982.s (in Czech)
[7] STODOLA, J.: For Active Safety of Automobiles. Military Technical Magazine. Nr 4. Praha, 1987 (in Czech)
[8] STODOLA, J.-VINTR, Z: Traffic Accident Information System and Possibilities of Risk Crash Evaluation. Book of Abstracts of World Automotive Congress FISITA. STA Barcelona Depósito legal: B-26597-04, 2004. (Page 412)

THE SEVESO II DIRECTIVE AND DANISH ACTIVITIES SUPPORTING ITS APPLICATION IN SOME EASTERN EUROPEAN COUNTRIES

Kozine, N.J. Duijm

●

Systems Analysis Department, Risoe National Laboratory, Denmark

H. Hagen (Consultant)

INTRODUCTION

The Council of the European Communities adopted the first so-called Seveso Directive, Directive 82/501/EEC in 1982. This Directive aimed at controlling major-accident hazards of industrial activities in the member states of the European Community, following the serious accidents at Flixborough, United Kingdom in 1974 (an hydrocarbon explosion in an refinery) and at Seveso, Italy in 1976 (release of dioxin following a runaway reaction in a chemical plant). This Directive was replaced by Directive 96/82/EC [1], called the Seveso II-Directive, on the control of major accident hazards, and adopted by the Council of the European Union in 1996.

The aim of the Seveso Directives is the prevention of major accidents involving dangerous substances and limitation of the consequences of such accidents, however limited to establishments having dangerous substances in excess of fixed threshold quantities. The Seveso II Directive is a legislative document that all the members of the EU must fulfil through implementation of national legislation

The Directive in addition to the European Union is also adopted by Norway, Iceland and Switzerland and countries intending to join the EU in addition as a condition have to introduce Legislation fulfilling the Directive.

At present, following the changes of the political scene in Europe and growing concern of the public about prevention of the consequences of major accidents, not least transboundary accidents such as the resent Baia Mare accident (Rumania) has lead to a proposals on broadening the scope of the Directive and the UNECE Convention on the Transboundary Effects of Industrial Accidents.The Seveso II Directive is based on the experiences accumulated during the implementation of Seveso I, in particular lessons learnt from accidents, which have occurred within the European Union since the adoption of Seveso I. The main changes are:

- The scope of Seveso II has been broadened and simplified, referring to the presence of dangerous substances at establishments in excess of threshold quantities, while Seveso I referred either to substances in connection with certain industrial activities or to separate storage of substances.

- The measures to be taken by the Operators of the Establishments to prevent and limit the consequences of major-accidents have been redefined and now include the setting up of a "Major-Accident Prevention Policy". The intention is to emphasise the commitment of the Operators of Establishments and the setting up of safety management systems as important elements to promote high levels of protection throughout the Community in an effective and consistent manner.

- Increased emphasis on measures to minimise environmental impacts of major-accidents including emergency preparedness and land-use planning, identification of possible domino effects, information to the public and where relevant to neighbouring countries (UNECE Convention on the Transboundary Effects of Industrial Accidents).

- To obtain uniform levels of protection throughout the European Union, the Member States are required to ensure that the Competent Authorities assess the Safety Reports and in particular are required to organise a system of ongoing inspections.

- Based on the Rome Treaty, the purpose of the Directive is the prevention of major accidents and to harmonise the efforts in this field within the EU to avoid that disparity in measures to prevent major accidents should affect the functioning of the common market.

- Seveso II is related to the new EU legislation on the protection of safety and health of workers, the Directive 89/391//EEC [2] in particular, which have come into force since Seveso I was adopted.

OBLIGATIONS ACCORDING TO THE SEVESO II DIRECTIVE

Operators of establishments, where substances in excess of certain threshold quantities given in the Seveso II Directive are present, are required to produce a Safety Report, demonstrating that:

- A major accident prevention policy and a safety management system for implementing it are in effect.
- Major accident hazards have been identified and necessary measures have been taken to prevent such accidents and limit their consequences for man and the environment.
- Adequate safety and reliability have been incorporated into the design, construction, operation and maintenance linked to major accident hazards.
- Internal emergency plans have been drawn up and information has been supplied to the Authorities enabling an external emergency plan to be drawn up.

To fulfil these obligations the Operators shall adopt and implement procedures for systematic identification of major hazards arising from normal and abnormal operations and to assess their likelihood and severity. It is important to carry in mind, that hazard identification and risk assessment are more or less universally required in other EU Directives such as the Machinery Directive, the Framework Directive on worker protection and the Directive on equipment and protective systems intended for use in potentially explosive atmospheres. The requirements on risk assessment included in these Directives may be limited to the safety or safe use of machines, explosion prevention and protection, while the Seveso II Directive has a wider scope covering the protection of man and the environment as a whole. The outcome of risk assessments as required by these Directives may however be useful in support of the risk assessment to be carried out by the Operator to demonstrate the adequacy of the measures taken to prevent major accidents - not least to avoid duplication of work.

Risk assessment always includes a final judgement, by the Operators as well as the Authorities, whether the measures taken are adequate or additional measures have to be taken. This judgement may in most cases be based on technical and managerial expertise, supported by comparison with the results of quantitative or qualitative risk analysis, use of recognised Standards, Codes of Practices and lessons learnt from accidents. It is important to note that no commonly agreed acceptance criteria have been laid down in support of these judgements at Community level.

The Seveso II Directive emphasise the responsibility of the Operators to take all necessary measures to prevent major accidents and limit the consequences if such accidents should occur. The obligations to provide the persons working on the site with information, training and equipment in order to ensure their safety, which were included in Seveso I, are now covered through other Directives.

The Competent Authorities are obliged as a minimum to receive and assess the Safety Reports and communicate the conclusions of the examinations to the Operator. The examination of the Safety Reports and the conclusions drawn must be seen in context with the requirements for setting up an inspection system.

It is important to have in mind, that the role of the Authorities in the Member States may be increased significantly, when the Seveso II Directive is fully implemented. Hopefully this may lead to a constructive dialog with the Operators and the employees to achieve the aim of the Directive and not result in more or less useless bureaucracy.

ACTIVITIES BY THE EUROPEAN COMMISSION TO SUPPORT THE IMPLEMENTATION OF THE SEVESO DIRECTIVES

To support uniform implementation of the Directive the Commission organises periodic meetings of the Competent Authorities covering interpretation of the content of the directive and exchange of information on the implementation.

In addition the Major Accident Hazards Bureau (MAHB) has been established at the Joint Research Centre in Ispra Italy. One of the tasks of MAHB is to collect, classify and distribute relevant information on the prevention of major accidents such as lessons learnt from accidents, Safety Reports and Codes of Practice.

In addition, the Commission and the MAHB have developed guidance documents to support the implementation of Seveso II, comprising:

- *Guidance on the preparation of a Safety Report* [3]
- *Guidelines on a Major Accident Prevention Policy and Safety Management System* [4]
- *Explanations and Guidelines on harmonised criteria for dispensations* [5]
- *Guidance on Land-use Planning* [6]
- *General Guidance for the content of information to the public* [7]
- *Guidance on Inspections* [8]

COURSES IN EASTERN EUROPE

The Danish Ministry of Labour in collaboration with the Danish Ministry of the Environment supports the adoption of Major Hazard Legislation in Eastern-European countries. In addition to other activities, it has funded projects to promote the use of hazard identification and risk assessment in connection with the implementation of Seveso II. These projects emphasise the promotion of the collaboration between the Competent Authorities, which is necessary due to the broad scope of the Directive.

As part of these projects, training courses were held for Polish, Czech, Slovak, and Estonian experts representing authorities, research institutes as well as the industry. These courses addressed the obligations of the Seveso-II Directive, the contents of the safety report, and a variety of risk analysis methods, from hazard identification to risk communication. Participants were actively involved in the course by analysing and discussing case studies. During all courses, there were vivid discussions about the national implementation (at different stages in the different countries) of the Directive, especially between authority and industry representatives.

COACHING IN SELECTED EXERCISES

In Poland and Estonia, as part of these projects, Danish experts were/are involved in coaching the formulation of safety reports for two selected companies in Poland (concluded) and one in Estonia (ongoing). The coaching aims at transferring practical experience in the preparation of Safety Reports and the assessment by the Competent Authorities.

The selected companies provides a first draft of a risk analysis study and a safety report, which is discussed and improvements are suggested a first meeting between local staff and Danish experts. This leads to a final report, established by the company and an assessment report by the Competent Authorities.

The Safety Report and the assessment report are in turn presented and discussed at a Workshop, with representatives from other relevant companies and authorities.

Finally, following the workshop a Polish group of experts visited Denmark to gain information on Danish major hazard installations, in particular related to the introduction of additional safety measures adopted as a consequence of lessons learnt and technical developments since the Directive was adopted.

CONCLUSION

In the atmosphere of growing mutual understanding of the necessity of preventing major industrial accidents in Europe including Russia, it is important to share experiences in providing safe process and procedures in handling hazardous substances and not least to establish networks for future exchange of information.

Major Hazard Legislation fulfilling the requirements of the Seveso II Directive are implemented in Eastern-European countries that wish to join the European Union and the necessary co-operation between the authorities has been initiated

The Seveso II Directive put obligations to industry, employees and authorities in order to control the risks related to the handling and storage of hazardous substances. It is required that the industry operators establish and implement a safety management policy, identify hazards at their plants, document that safety aspects are duly included in the design and operation of their plants and that emergency plans are in place.

The authorities are to verify that this has been done by assessment of safety reports and inspection, and to set up external emergency plans.

Danish experts have for several years supported the implementation of the Seveso II Directive in Eastern Europe through exchange of information and experience at meetings, courses, coaching and visits to Denmark. These activities have met a wish from these countries to transfer experience and practical methods necessary to live up to the requirements of Seveso II.

REFERENCES

1. Council Directive 96/82/EC on the control of major accident hazards involving dangerous substances.
2. Council Directive 89/391/EEC on the introduction of Measures to Encourage Improvements in the Health and Safety at Work.
3. Papadakis G. A. and A. Amendola (Eds.) (1997): Guidance on the preparation of a Safety Report to meet the requirements of Council Directive 96/82/EC (Seveso II). EUR 17690 EN. JRC Ispra.
4. N. Mitchison and Sam Porter (Eds.) (1998): Guidelines on a major accident prevention policy and safety management system, as required by Council Directive 96/82/EC (Seveso II). EUR 18123 EN. JRC Ispra.
5. J. Wettig, N. Mitchison (Eds.) (1999): Explanation and guidelines for the application of the dispensation rule of article 9(6) of Council Directive 96/82/EC (Seveso II), Report EUR 18124, Office for publications for the EC, Luxembourg.
6. Christou, M.D., Porter, S. (1999) Guidance on land use planning as required by council directive 96/82/EC (Seveso II), Report EUR 18695, Office for publications for the EC, L-2985 Luxembourg.
7. B. De Marchi, S. Funtowicz (1994): General Guidelines for Content of Information to the Public. Directive 82/501/EEC - Annex VII. EUR 15946.
8. Papadakis G. A and S.Porter (Eds.) (1999): Guidance on Inspections as Required by Article 18 of the Council Directive 96/82/EC (Seveso II). EUR 18692. JRC Ispra.Community Documentation Centre on Industriial Risk: Comparison of selected LPG related codes and Standards. EUR 14636, JRC Ispra . 1992.
9. Community Documentation Centre on Industrial Risk: National Approaches to the Safety Report – a Comparison. SP-I.91.07. JRC Ispra, 1991 .
10. Community Documentation Centre on Industrial Risk: Review of accident involving ammonia. EUR 14633, JRC Ispra 1992.

EU ADVANCES IN IDENTIFYING SOURCES OF UNCERTAINTY IN RISK ANALYSES

K. Lauridsen and I. Kozine
●
Systems Analysis Department, Risoe National Laboratory, Roskilde, Denmark

A. Amendola and M. Fiori
●
EC-Joint Research Centre, Ispra, Italy

INTRODUCTION

Quantitative Risk Assessment (QRA) aims at the modelling of stochastic uncertainties associated with the occurrence and circumstances of a major accident. But the process itself of carrying out a QRA implies several uncertainties. For the implementation of the risk assessment procedure a variety of techniques and models must be used, and uncertainties are introduced due to imperfect knowledge and expert judgement. Because QRA is used as input in many decisions related to the control of major accident hazards and the need for accuracy in the results increases, the adequate management of these uncertainties gains increased importance.

This paper presents the scope and some main results of a European project on the ASSessment of Uncertainties in Risk ANalysis of Chemical Establishments (ASSURANCE). The project aims at identifying the uncertainties associated with risk analysis of major industrial hazards and assessing the way these uncertainties can affect the final outcome of risk studies and of the relevant decisions based on that outcome. In order to achieve this goal, a number of benchmark exercises/case studies have been performed by the partners and the results were analysed in a modular and structured way. A reference plant served as the basis for a realistic description of these case studies. For this particular project an ammonia storage plant was selected, consisting of cryogenic and pressurised storage tanks, together with import loading/unloading facilities and the relevant piping. This installation was analysed independently by each partner, using common input data and boundary conditions, but different methods, tools and assumptions. The results were then compared and discrepancies identified, discussed and explained.

In order to permit the step-wise comparison of the results and to assess the contribution of each factor and each phase of Risk Assessment to the overall discrepancy, the analysis was divided in the various phases (Hazard identification, frequency estimation, consequence assessment, and overall risk assessment), and the results of each phase were compared. Moreover, detailed exercises addressing particular issues within each phase (e.g. source-term definition, dispersion modelling, vulnerability modelling, etc.) were performed, in order to give more insights on the factors affecting the overall discrepancies in the results.

Concerning the quantification of risk, a structured procedure was followed in reporting and comparison of results. This procedure required not only the assessment of the risk profile, i.e. estimate of the level of risk at each point in the area around the plant, expressed in the form of isorisk curves for individual risk and F-N curves for societal risk, but also the assessment of intermediate results. These included:

- Assessment of the frequencies associated to accident scenarios,
- calculation of release/evaporation rates and conditions,
- modelling of dispersion (including detailed results of selected scenarios), and
- dose-response calculations.

FINDINGS CONCERNING HAZARD IDENTIFICATION

Comparison of the approaches to hazard identification showed that the partners had used many different methods. As a matter of fact, no two partners had used exactly the same method, although some of the methods, of course, are of similar nature. The methods used were:

- HAZard and OPerability analysis (HAZOP)
- Master Logic Diagram (MLD)
- Structured What-IF Technique (SWIFT)
- Hazard Identification by Area Audit (HIAA)
- Function analysis and Hazard and Consequences Analysis
- HAZardous SCenario ANalysis (HAZSCAN)
- Use of (national) standard checklists based on accumulated experience from past accidents and past (detailed) studies.

These methods can be grouped into three general types of approach:

- Methods based on a top-down analysis, mainly represented by the Master Logic Diagram, which has a form similar to Fault Trees, starting from a top event and going down to combinations of basic events that can initiate an accident
- Methods based on a bottom-up analysis, like HAZOP, SWIFT and HAZSCAN, which investigate whether deviations of the process variables and failures of individual devices can initiate an accident
- Methods based on the systematic use of standard checklists, after division of the plant into areas. Here, the accumulated experience from past accidents and studies is combined with systematic rules to identify the areas that deserve a more detailed analysis.

Even though the partners had used different methods for the hazard identification they had all identified the accident scenarios that must be considered the most severe. But, due partly to the use of different methods, each partner had some scenarios that other partners did not have in their list of selected scenarios. Therefore, in order to have a common basis for comparison of the methods used in the following quantified risk analysis phase, 11 reference scenarios were agreed for further analysis by everybody together with possible additional scenarios identified by the individual participants. Examples of the reference scenarios are:

- Full section rupture of an 8" import pipeline
- Full section rupture of a specific 4" pipeline
- Full section rupture or disconnection of the loading/unloading arm to a ship
- Catastrophic rupture of the cryogenic tank
- Catastrophic rupture of a pressurised tank

FINDINGS CONCERNING THE QUANTIFICATION OF RISK

One of the ways the participants were requested to present the results of their quantitative risk analyses was by means of iso-risk curves for the individual risk, i.e. curves on a map where the risk is the same for all points on the curve. The individual risk is the probability that a person staying unprotected in the same location around the clock during one year will die as a consequence of an accident in the facility considered. The two curves in the example shown in Figure 1 are the maximum and minimum distances found by the participants for an annual fatality risk of 10^{-5}. As can be seen, there are rather large differences; the diameter of the outer curve is roughly 2 km.

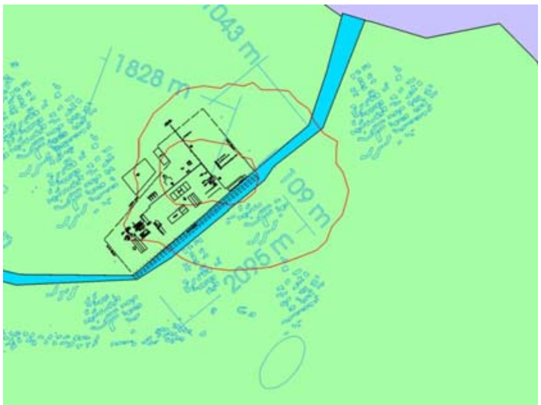


Figure 1 Iso-risk curves for annual individual risk of 10^{-5}

Although care was taken to specify reference scenarios well, major differences were seen in the partners' risk results for these scenarios. Some causes that could easily be identified were:

- some remaining misunderstandings concerning plant data and specification of the reference scenarios
- differences in data used for failure probabilities
- different assumptions used concerning release duration.

The uncertainty assessment for the quantitative analysis phase was carried out separately for the frequency assessments and consequence modelling of hazardous scenarios. This separation allowed the identification of root causes of the deviation in risk assessments and their range among the research teams.

Frequency assessment

Uncertainty analysis in frequency assessment was based on the understanding that there were three different types of uncertainty: modelling, completeness and parameter uncertainty. Modelling uncertainty results from the use of different analysis models (fault trees, event trees and simplified generic-based models). Completeness uncertainty is due to differences in the number and nomenclature of basic/intermediate and initiating events included in the modelling. And parameter uncertainty is due to different numerical inputs (basic/initiating event frequencies, in particular) used to assess the net frequency of a hazard.

Not all types of uncertainty can be analysed quantitatively. Thus, modelling uncertainty in principle allows quantification only if the input data are the same. Having the same inputs for modelling would reveal the variability in outputs. If inputs cannot be made the same, it is impossible to distinguish whether the variability in final assessments is caused by different inputs or the models themselves. After having analysed all the approaches for frequency assessments it became obvious that the inputs could not be made the same because the sets of basic and initiating events are very different and overlap only to some extent. It was concluded that for risk analysis studies it is hardly possible to split up completeness and modelling uncertainty analysis and to perform their quantitative categorisation. Yet, these two kinds of uncertainty can be analysed in a descriptive way through the review of different approaches and basic and intermediate events included in the analyses.

Figure 2 illustrates the deviation in frequencies found by six research teams for one scenario. As can be seen, there are deviations of more than one order of magnitude between some teams in this case. An investigation into the root causes of these deviations revealed differences in data sources, in the use of data from the same source and in the interpretation of plant data.

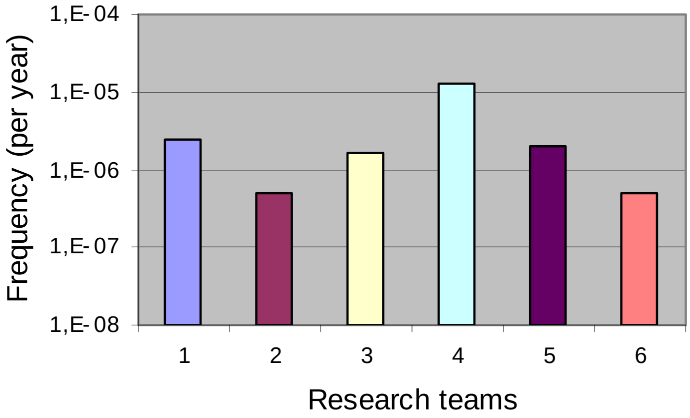


Figure 2 Frequencies of the rupture of a pressurised ammonia tank

Consequence assessment

The consequence assessment included four phases: (i) outflow calculations; (ii) pool formation and evaporation (whenever applicable); (iii) dispersion; and (iv) dose/response modelling. Since the assessment of the implications of differences in dose/response (vulnerability) models is straightforward, the project focused on the first three phases, and the final results reported were in most cases concentration- and dose endpoints. In particular, three concentration endpoints were used, corresponding to concentration levels of 6200, 1000 and 500 ppm, and three dose endpoints, corresponding to dose levels equivalent to 30 minutes exposure to a constant concentration of ammonia equal to the above values.

In addition to these endpoints, a number of intermediate results were calculated and reported (e.g. outflow rates and conditions, pool dimension and characteristics, percentage of droplets, evaporation rates and behaviour of the cloud).

The comparison of the calculated endpoints for the reference scenarios revealed again noteworthy discrepancies. In general, the sources of uncertainty in the consequence assessment can be divided in the following categories:

- Scenario completeness and correctness
- Uncertainty in definition of scenarios/ambiguity
- Modelling uncertainty, including the description of physical phenomena and the detailed model characteristics, constants and parameters
- Input assumptions/boundary conditions/interface between models
- Simplifications made throughout the analysis
- Overall level of “conservatism” of the analyst.

CONCLUSIONS AND RECOMMENDATIONS

The findings mentioned above and other observations during the project led to the identification of the following types of uncertainty and variability that influence the results of risk analyses and for which an effort is needed in order to minimise their influence:

- Misunderstandings or lack of knowledge about plant layout and operation
- Completeness of hazard identification
- Modelling uncertainty (failure modelling and consequence modelling)
- Data uncertainty
- Variability in, for instance, weather conditions or plant operational state

- Variability due to the fact that probabilities are not fixed numbers but distributions.

The first one, lack of knowledge/misunderstandings about plant layout and operation, may have had a particularly great influence in this project because the interaction between the individual risk analyst and the plant staff was not typical for a normal risk analysis. Due to the geographical diversity of the consortium it was not possible to visit the plant more than once, and in order to give all partners equal conditions all communication with the plant staff went through the co-ordinators of the exercise.

The differences experienced between the partners' results suggest the need for:

- Recommendations concerning the use of standardised approaches to risk analysis in Europe
- Recommendations for common data sources for failure rates of components.

In any case it will be necessary to ascertain that risk analyses performed in different EU countries by different analysts are comparable and lead to similar results.

After the project has finished results from it will be made available at the JRC's MAHB server (http://mahbsrv.jrc.it/) in order to secure their availability to the international community of risk analysts.

ACKNOWLEDGEMENTS

ASSURANCE was a benchmark exercise and the outcomes of the project are the joint effort of many project participants in addition to the authors of the present paper. O. Aneziris, D.A. Carter, S. Coulter, P. Crossthwaite, M. Fiori, S. Gadd, I.L. Hirst, D. Hourtolou, C.M.A. Jansen, Y. Malmén, M. Molag, A. Montesi, Z. Nivolianitou, I.A. Papazoglou, G. Spadoni and J. Tiihonen are all gratefully acknowledged for their work and for providing the background data. Special thanks are expressed to the anonymous company and its safety engineers. Without the availability of the plant and their co-operation and help to resolve the numerous questions of the partners, this project would never have been realised. Last but not least, the funding from the European Commission under the ENVIRONMENT programme is kindly acknowledged.

ADAPTATION, LEARNING AND INHERENT SAFETY OF 2ND GENERATION AIRSHIPS

Henry K Moskatov
●
Central Shipbuilding Res. Inst. «CENTRE»
123242, MOSCOW, D-242 The Russian Federation
e-mail: gkm@mail.sbnet.ru

Abstract. Inherent safety of the new generation airships, based on some fundamental laws of Space, is discussed in some detail. An algorithm is proposed to analyze risks, resulting from hazards not compensated by "inherent safety". Then a thoroughly verified statistical model of learning is used to evaluate results of airship flight testing-the probability of mission success and its confidence limit. The results can be used as a part of evidence for airship airworthiness certification.

Keywords: Adaptation, airship, airworthiness, degenerative feedback, laws of Space, statistical model of learning, inherent safety.

It looks like as if helium airships were on the threshold of the Renaissance. For the new generation airships niches have been found, where they are, and in the predictable future will remain, second to none. Their unique features and capabilities will find increasing application in the future particularly as fossil fuels become less acceptable in the coming decades [1,4-6,11,12].

Among these traces airships' *inherent safety* is of paramount importance. As we've shown in [11,12] it is a consequence of three fundamental laws of Space, acting simultaneously:

- the Law of Archimedes;
- the Principle of Le Chatelier-Brown (the Principle of Adaptation);
- the Principle of Rhythm and Periodicity.

2nd generation airships are safer than heavier-than-air vehicles of any type due to:

- using helium as lifting gas,
- independence of aerostatic lift of flight velocity and, hence-of the power plant used;
- non-zero metacentric height.

These three factors determine the main contribution to airships' inherent safety.

A helium airship with an autopilot, parametric and topological redundancy, embraced with degenerative feedback loops, becomes a system of mobile, dynamic equilibrium, where Le Chatelier- Brown Principle is valid. Yet the Principle has its feasibility limits: it holds true unless destabilizing factor exceeds a certain predetermined threshold. Its knowledge is essential for airships' safe operation, especially with respect to risk factors, associated with vehicles' statics and dynamics.

Any object possessing a certain amount of stored energy of any kind isn't absolutely safe. And helium airships aren't an exception. Inherent safety doesn't imply the vehicle's immunity from all possible hazards.

An appropriate algorithm to analyze risks, resulting from hazards, not covered by the feature, called "inherent safety", is given in figure below. Special safety assurance facilities are inevitable in this case.

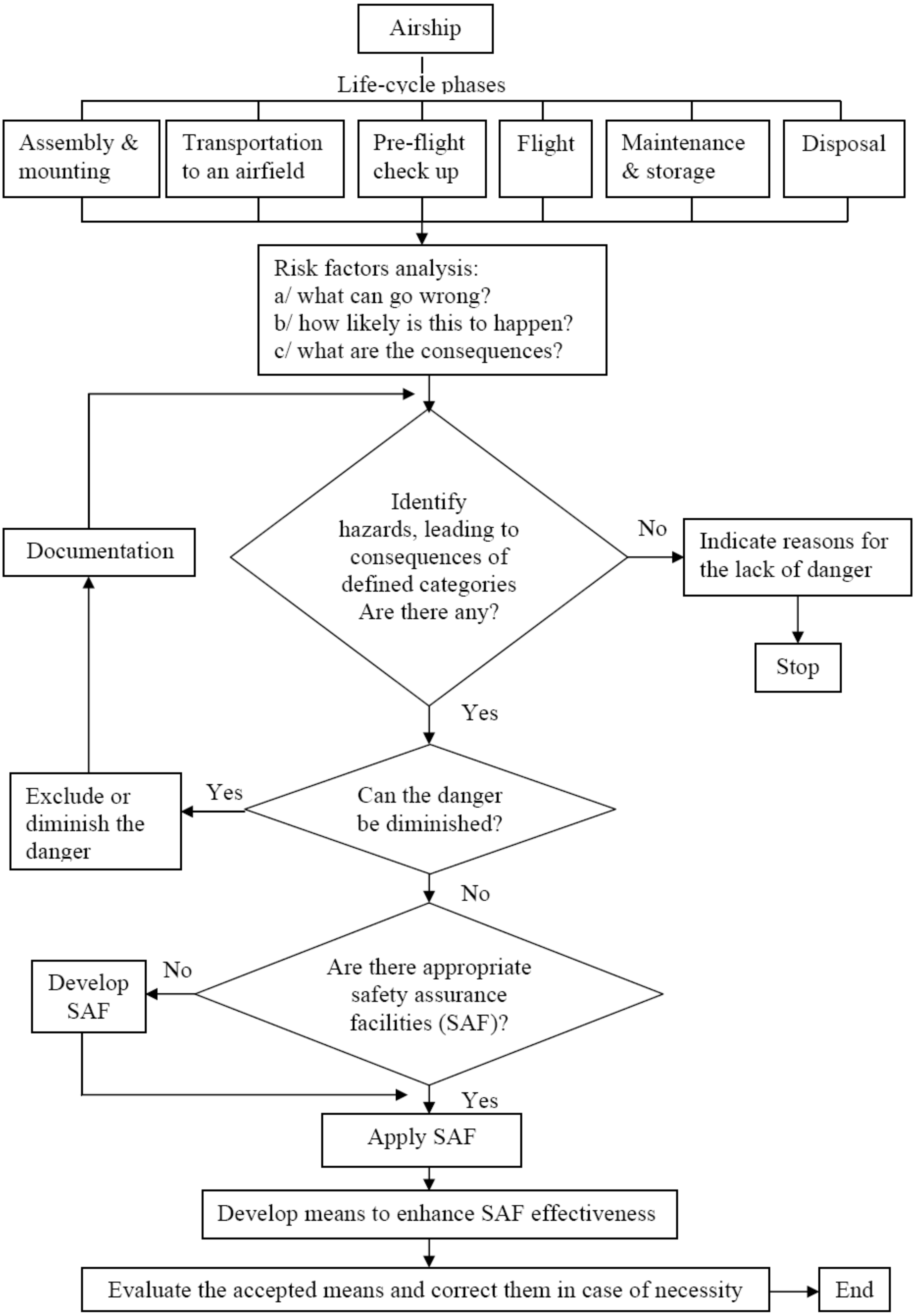

Airship
Life-cycle phases
Assembly & mounting
Transportation to an airfield
Pre-flight check up
Flight
Maintenance & storage
Disposal
Risk factors analysis:
a/ what can go wrong?
b/ how likely is this to happen?
c/ what are the consequences?
Identify hazards, leading to consequences of defined categories Are there any?
No
Indicate reasons for the lack of danger
Stop
Documentation
Yes
Can the danger be diminished?
Yes
Exclude or diminish the danger
No
Are there appropriate safety assurance facilities (SAF)?
No
Develop SAF
Yes
Apply SAF
Develop means to enhance SAF effectiveness
Evaluate the accepted means and correct them in case of necessity
End

Figure. Block-diagram of risk analysis & safety assurance algorithm.
Hazards consequence categories: minor, severe, major, catastrophic [17].

Some comments don't seem senseless in this connection. Risks analysis has to answer three questions: what can go wrong? (by hazard identification; how likely is this to happen? (by frequency analysis); what are the consequences? (by consequence analysis).

"The moment of truth" of an airship in any way comes during ground & flight tests.

For self-evident reasons vehicle testing is accompanied with corrective action The latter can produce one of the 3 issues:

-the vehicle's reliability is enhanced;

-it remains unchanged;

-it is deteriorated because of erroneous actions.

For a vehicle's reliability estimate the following indices can be used:

a) non-stationary probability of mission success;
b) non-stationary mean-time-between failures (MTBF);
c) non-stationary failure rate [3,7-9,12].

Practically for all reliability growth models of the (a) class the following mathematical structure is characteristic:

$$\hat{P}_n = \hat{P}_0 + \varphi(n, \hat{P}_0) \tag{1}$$

where

$\hat{P}_n$ - a statistical estimate of probability of mission success after the n-th test (or test phase);

$\hat{P}_0$ - a probability estimate before the testing began;

$\varphi(n, \hat{P}_0)$ - a learning nucleus - a cumulative term, characterizing mission success probability increment at the n-th test phase, $0 \le \varphi(n, \hat{P}_o), < 1$

A model design is to concretize $\varphi(n, \hat{P}_0)$, while model application is to define its unknown parameters on the basis of non-homogeneous test record. In the simplest case

$$\varphi(n, \hat{P}_0) = \sum_{i=1}^{n} \Delta P_i \ , \tag{2}$$

where ΔP_i - reliability increment as a result of i-th corrective action; $\Delta P_i >< 0$; $\Delta P_i = 0$.

Over 20 statistical models of learning and reliability growth were built.

Efforts to construct new ones can still be observed. Analytical surveys of the models have been published in USA, USSR and elsewhere.

The most successful representative of the second and third classes is the Duane model [3]. It was constructed on the basis of comprehensive testing experience of aerospace digital and analog systems. The Duane relationship is:

$$\Lambda_\Sigma = \frac{d(T)}{T} = KT^{-\alpha} \ , \tag{3}$$

where Λ_Σ – an accumulated failure rate,

d(T) – a number of failures in all tests;

T – general number of hours (cycles) of testing;

K – an experimentally determined constant value,

α – index of reliability growth speed.

If the perfection (reliability growth) is not attained, *α* = 0 and in this case

$$\Lambda_\Sigma = \frac{d(T)}{T} = K \ , \tag{4}$$

From (3) one can obtain, as a result of trivial transformations, instantaneous MTBF

$$T_0 = [(1-\alpha)KT^{-\alpha}]^{-1}, \tag{5}$$

If a vehicle is not perfected during testing, then α = 0, and MTBF becomes a constant value, independent of the time of learning, i.e. $T_0 = {}^{1}\!/_{K} = {}^{1}\!/_{\Lambda_\Sigma}$.

If the reliability criterion is non-stationary probability of mission success a growth model with degenerative feedback is recommended

$$\hat{P}_n = \hat{b} - {}^{\hat{c}}\!/_{\hat{a}} \exp(-n/\hat{a}), \tag{6}$$

where $\hat{a}, \hat{b}, \hat{c}$ - estimates of unknown parametres a, b, c; the latter being functions of two independent variables: i - the number of trials, K_i – accumulated number of successes; $\hat{b} = \lim_{n\to\infty} \hat{P}_n$; ${}^{1}\!/_{\hat{a}}$ characterizes the effectiveness of feedback [13].

The model deals with catastrophic failures (both in hardware and software); trials, proved unsuccessful due to human operator errors, are not taken into account.

The algorithm for estimating the unknown parameters:

$$\hat{b} = \frac{S4S2 - S7S5}{(S4)^2 - S8S5}, \tag{7}$$

$$\hat{c} = \frac{S8S2 - S4S7}{(S4)^2 - S8S5}. \tag{8}$$

Parametr $\hat{a}$ is the root value – a nontrivial solution of an equation $f(a) = 0$, where

$$f(a) = S1(S7S5 - S2S4) + (S8S2 - S7S4)S6 + S3((S4)^2 - S5S8). \tag{9}$$

Here

$$\left.\begin{array}{l} S1 = \sum_{i=1}^{n} i^2 y; \quad S2 = \sum_{i=1}^{n} K_i y; \quad S3 = \sum_{i=1}^{n} iK_i y; \\ S4 = \sum_{i=1}^{n} iy; \quad S5 = \sum_{i=1}^{n} y^2; \quad S6 = \sum_{i=1}^{n} iy^2; \\ S7 = \sum_{i=1}^{n} iK_i; \quad S8 = \sum_{i=1}^{n} i^2; \\ y = 1 - \exp(-i/\hat{a}). \end{array}\right\} \tag{10}$$

When $\hat{b} \geq 1$ $\hat{b}$ is assumed to equal 1, and $\hat{a}$, $\hat{c}$ are calculated according to

$$\hat{c} = (S4 - S2)/S5,$$

$$f(a) = S5(S7 - S3) + (S4 - S6)(S4 - S2) - S5(S8\text{-S1}) \tag{11}$$

The root $\hat{a}$ of the transcendental equation $f(a)$ = 0 is calculated by introducing test values a_j to (9),(11), where

$$a_{j+1} = a_j + 1, j = 0, 1, 2, \dots .$$

Formulas (7–11) were derived by using least squares method and provide likelihood function maximum.

Comprehensive numerical experiments conducted with the growth model assert that the test values a_j belong to a closed interval [0.001, …, 25] for a number of trials not exceeding 100.

In the course of calculations the signs of $f(a_j)$ and $f(a_{j+1})$ are compared. When the signs change for the first time the corresponding values of a_j, a_{j+1}, $f(a_j)$, $f(a_{j+1})$ are stored and used in the interpolation formula

$$\hat{a} = \frac{a_{j+1} f(a_j) - a_j f(a_{j+1})}{f(a_j) - f(a_{j+1})}.$$

The found value of $\hat{a}$ is used in (7-9) to determine $\hat{b}$, $\hat{c}$, P_n.

The confidence interval for $\hat{P}_n$ estimate is determined by using the Clopper – Pearson formula.

Reliability growth is established by checking up the validity of the following inequality

$$K_n / n > S7/S8.$$

If $K_n / n \le S7/S8$, learning (perfection) of the vehicle was not attained and

$$P_n = K_n / n \text{ [13].}$$

And now we'll consider a concrete example of an airship flight test record.

There were altogether one hundred trials, alas not all of them proved successful.

№ of a trial, *i*	1 – 10, 11, 12 – 21, 22, 23 - 100
An accumulated № of successes K_i	98
№ of test stages *M*	3

The first ten trials were succesful; then a failure in the empennage occurred. The reasons were investigated, found and the rudder was perfected to meet specifications. It was decided to carry on testing, partially to get convinced that the corrective action was sufficient. The next 10 flights were OK, but the 22-nd landing wasn't all right. We omit listing physical reasons of that failure. They were scrutinized and eventually compensated. The following 78 trials proved successful.

This airship flight test experience with corrective action was used in an original PC-program to estimate the nonstationary probability of mission success and the corresponding confidence interval for 90% confidence level

$$0,99313 < 0,99981 < 1. \qquad (12)$$

Hadn't we taken into account corrective actions (i.e. neglected reliability growth and regarded the sequence of trials as homogeneous) the result would have been different, namely

$$P_{100} = 98/100 .$$

In the case of independent homogeneous tests (the Bernoulli distribution) 2303 trials, all of them successful, would be required to confirm the probability of success as in (12) for the same confidence level.

The model is *sensitive to the stage the failure occurs*. Common sense prompts that an earlier failure detection and compensation are more advantageous than those conducted at a later flight test phase.

The result of other flight test records (not given here for brevity) demonstrated a remarkable coincidence of common sense and the mathematical model sensitivity.

THE SUMMING UP

The recently proclaimed "airships inherent safety" is based on three fundamental laws of Space - the Law of Archimedes, the Principle of Le Chatelier-Brown and Principle of Rhythm and Periodicity, acting together. Still it should be made clear that nothing and nobody liberate us, anyone involved in aeronautics, from responsibility for our own conduct. The notion implies not only the quality of the vehicle itself but also meteorologists', ground crews', airpilots' ability to predict and adequately meet challenges of our evolving Universe.

Global atmospheric warming and other solar-terrestrial links provoked hazards, to mention just a few, continue to produce more stormy conditions [4-6].

Airships flight schedule and envelope should cohere to rhythms of the Earth and the Solar system to retain significant variations of geophysical fields within admissible limits.

We must confess at last that we can no longer permit ourselves flying whenever and wherever we want. For our safety sake our plans, likes and dislikes, our partialities should be made compatible with fundamental laws of Space. At least, with those we know and understand.

Inherent safety doesn't imply the vehicle's immunity from all possible hazards. Hence special safety assurance facilities should be provided for hazards, not covered by airships inherent safety feature.

In fact all that has been lost, missed and/or misunderstood inevitably reveals itself during tests. If the aim of airship testing assessment is probability of mission success, the thoroughly verified statistical model of learning (6) is recommended for use [13].

If the aim is different – evaluating MTBF, the Duane growth model should be preferred, the latter being the core of the latest IEC International Standard [14].

Airship reliability and hence – safety growth is a natural product of degenerative feedback – corrective actions undertaken at all stages: R&D, production and testing.

In any way application of the growth models considerably reduces the number of trials (duration of testing), required to confirm specifications and consequently results in a notable saving of time, space and money.

The experimentally confirmed probability of mission success, as well as other results outlined, can be used as a part of evidence for airship airworthiness certification.

REFERENCES

1. Arie M. The airship accidents and catastrophes. Proc. 3-rd Intern. Airship Conv. and Exhibition. The Airship Association Ltd., 2000.
2. Barkovsky E. Catastrophes on the sea and under water: myths and reality. Morskoy Sbornik, 2001, №10, p.31-41 (in Russian).
3. Duane J.T. Learning curve approach to reliability monitoring. IEEE Trans. on aerospace, 1964, № 4, p.563-566.
4. Harris M. The use of weather satellite pictures in airship operations worldwide. Proc. 3-rd Intern. Airship Conv. and Exhibition. The Airship Association Ltd., 2000.
5. Harris M. The significance of understanding the weather hazards involved in flying airships in and around mountainous regions. Proc. 14th AIAA LTA Technical Committee Convention. Akron, Ohio, July 2001.
6. Harris M. The practical application of local meteorology for airship operations in city areas. Proc. 4th Intern. Airship Conv. and Exhibition. The Airship Association Ltd., 2000.
7. Maiorov A., Moskatov H., Shibanov G. Operating safety of automated objects. Mashinostroenie Publishers, M., 1988, 264 p (in Russian).
8. Moskatov H. Reliability of Adaptive Systems, Soviet Radio Publishers, Moscow, 1973, 103 p. (in Russian).
9. Moskatov H. System testing as learning with corrective action. In "Main problems of reliability theory and practice", Soviet Radio Publishers, Moscow, 1975, p.294-306 (in Russian).
10. Moskatov H. The Principle of Le Chatelier-Brown and inherently safe feedback systems' design. Safety Problems in Emergencies, 1992, #2, p.49-76 (in Russian).
11. Moskatov H., Kirilin A. Adaptation, redundancy and inherent safety of modern airships. AIRSHIP, № 118, December, 1997, p.13, 26-27.
12. Moskatov H. Airship flight testing as learning with corrective action. Proc.4th Intern. Airship Convention and Exhibition, Cambridge, UK. The Airship Association Ltd., 2002.

13. Pupkov K., Kostyuk G. Experiments planning and assessment. Mashinostroenie Publishers, M., 1977, 118 p (in Russian).

14. Reliability Growth – statistical test & estimation methods. Intern. Electrotechnical Commission Standard, 1-st Ed., Geneva, 1995.

15. Ryabinin I. Reliability and safety of structurally complex systems. Polytechnics Publishers, St. Petersburg, 2000,248 p (in Russian).

16. Ryabinin I., Parfenov Yu. Reliability, survivalability and safety of ship power systems. Naval Academy, St. Petersburg, 1997, 410 p (in Russian).

17. Risk Analysis. International Standard. International Electrotechnical Commission. Geneva, Suisse, 1995.

PLANNING OF INSPECTION PROGRAM OF FATIGUE-PRONE AIRFRAME

Yu. Paramonov, A. Kuznetsov

●

Aviation Institute, Riga Technical University, Riga, Latvia
1 Lomonosova Str., Riga, Latvia
e-mail: rauprm@junik.lv, andreyk@hotbox.ru

Abstract. To keep the fatigue ageing failure probability of an aircraft fleet on or below the certain level an inspection program is appointed to discover fatigue cracks before they decrease the residual strength of the airframe lower the level allowed by regulations. In this article the Minimax approach with the use one- and two-parametric Monte Carlo modelling for calculating failure probability in the interval between inspections is offered.

Keywords: failure probability, Minimax approach, inspection program, approval test

INTRODUCTION

Inspection program development should be made on the base of processing of approval lifetime test result, when we should make some redesign of the tested system if any requirement is not met. Here we consider some example of p-set function application to the problem of development and control of inspection program. We make assumption that some Structural Significant Item (SSI), the failure of which is the failure of the whole system, is characterized by a random vector (r.v.) (T_d , T_c), where T_c is critical lifetime (up to failure), T_d is service time, when some damage (fatigue crack) can be detected. So we have some time interval, such that if in this interval some inspection will be fulfilled, then we can eliminate the failure of the SSI. We suppose also that a required operational life of the system is limited by so-called Specified Life (SL), t_{SL}, when system is discarded from service. P-set function for random vector is a special statistical decision function, which is defined in following way. Let Z and X are random vectors of m and n dimensions and we suppose that it is known the class $\{P_\theta,\ \theta\in\ \Omega\}$ to which the probability distribution of the random vector W=(Z,X) is assumed to belong. The only thing we assume to be known about the parameter θ is that it lies in a certain set Ω, the parameter space. If $S_Z(x)=\bigcup_i S_{Z,i}(x)$ is such set of disjoint sets of z values as function of x that $\sup_\theta \sum_i P(Z\in S_{Z,i}(x))\le p$ then statistical decision function $S_Z(x)$ is p-set function for r.v. Z on the base of a sample $x=(x_1,...,x_n)$.

Later on the value x, observation of the vector X, would be interpreted as estimate $\hat{\theta}$ of parameter θ, Z would be interpreted as some random vector-characteristic of some SSI in service: $Z=(T_d,T_c)$.

FATIGUE CRACK GROWTH MODEL

The fatigue crack growth process flows in accordance with quite complicated rules, which depend on a big number of factors. An analytical approach describing that process could be considered as almost impossible. Nevertheless, it can be shown, that in general case crack growth process could be well enough approximated with the formula:

$$a(t)=\alpha\cdot e^{Qt}, \qquad 1.$$

where *a(t)* is a fatigue crack size at time *t* (the number of flight blocks); α is so called equivalent initial crack size (as if the airframe has been initially produced with the crack of such small size; α corresponds to the best fit of test data); and parameter *Q* defines the speed of growth of fatigue crack and depends on the loading mode (on the stress range in case of cycling loading).

For further needs, let us take a logarithm of both left and right sides of equation 1:

$$\ln a(t) = \ln \alpha + Qt \,. \qquad 2.$$

Thus,

$$t = \frac{\ln a(t) - \ln \alpha}{Q}, \qquad 3.$$

so the time when crack becomes detectable and the time when crack reaches its critical size can be calculated as

$$T_d = \frac{\ln a_{\det} - \ln \alpha}{Q} = C_d \big/ Q, \quad T_c = \frac{\ln a_{crit} - \ln \alpha}{Q} = C_c \big/ Q, \qquad 4.$$

where $a_{\det}$ is a crack size, at which chances to discover it tends to unit, a_{crit} is a crack size, which corresponds to the minimum residual strength of an aircraft component allowed by regulations, T_{d} is a time for crack to grow to its detectable size and T_{c} is a time for crack to grow to its critical size, C_{c} and C_{d} are appropriate constants.

Let us define *failure* as the situation, when we were unable to discover cracks with the size $a_{\det} \le a < a_{crit}$, or, in other words, if there are no inspections performed in [T_{d}; T_{c}] time interval.

It is clear, that varying the number of inspections $n_{inspections}$ in the service interval $[0, t_{SL}]$ we will discover a different number of cracks; therefore, the estimate of failure probability will vary as well. Unfortunately, we don’t know the real values of parameters, so we are using theirs estimates from a small number (one, seldom two) of available observations (fatigue cracks during fatigue test) instead.

USING MONTE CARLO MODELLING TO ESTIMATE FAILURE PROBABILITY

We use the Monte Carlo method to generate a set of cracks to be processed in accordance with procedure, described in Section 0. The parameters for modelling can be derived from the full-scale fatigue tests or from other real crack observations.

We can never know how the certain fatigue crack curve will look like. Thus, performing approximation of that fatigue crack curve with a certain model, the fatigue crack growth model parameters (FCGMP) – we have two FCGM parameters $X = \ln Q$ and $Y = \ln C_C$ – will vary as well, so they are random values, and these random values have theirs own parameters of distribution. To perform Monte Carlo modelling of the fatigue crack growth process we have to know FCGMPs’ distribution types and parameters, i.e. c.d.f. of each FCGMP. From the analysis of the fatigue test data it can be assumed, that the logarithm of time required the crack to grow to its critical size is distributed normally:

$$\ln T_c \sim N(\mu_{LT_c}, \sigma^2_{LT_c}) \,. \qquad 5.$$

From formulas 2 and 4 follows:

$$\ln T_c = \ln C_c - \ln Q \,. \qquad 6.$$

From additive property of normal distribution comes that $\ln T_c$ could be normally distributed either if both $\ln C_c$ and $\ln Q$ are normally distributed:

$$X = \ln Q \sim N(\mu_X, \sigma_X^2), \qquad Y = \ln C_c \sim N(\mu_Y, \sigma_Y^2), \qquad 7.$$

or if one of them is normally distributed while another one is a constant. Thus, the value of logarithm of our FCGM parameters is distributed normally or, on other words, FCGM parameters have a log-normal distribution.

To get estimates of FCGMP distribution parameters ($\hat{\mu}$ and $\hat{\sigma}$) we consider statistics of several crack observations. For each of those cracks we calculate estimates of distribution parameters ln$Q$ and lnC_{c}, and then gather all data together into the table with two columns: ln$Q$ and lnC_{c}. From that table we then derive estimates of mean value and standard deviation for each column, as well as estimate of correlation between ln$Q$ and lnC_{c}.

The Monte Carlo modelling in fact means the process of getting a big number of pairs [T_{d}; T_{c}] with upper mentioned specific distribution parameters. Having the array of [T_{d}; T_c] pairs we apply an inspection program looking for *failures* – situations, when both T_d and T_c are located between two consequent inspections. For each interval between inspections [t_{i-1};t_i] failure probability will be

$$P_{f_i} = P(t_{i-1} < T_d \le T_c < t_i), \qquad 8.$$

and for the entire inspection program

$$P_f = \sum_i P_{f_i} . \qquad 9.$$

MINIMAX DECISION MAKING APPROACH

As it was stated above, the goal is to develop an inspection program, defined by the vector of inspection time moments

$$\vec{t} = (t_1, t_2, \cdots t_{n_{TIP}}), \qquad 10.$$

i.e. to find a vector function $\vec{t}\,(\hat{\theta})$ ($\hat{\theta}$ is the estimate of FCGMP distribution parameters, n_{TIP} is the total number of inspections per inspection program, so $t_{n_{TIP}} = t_{SL}$) that limits aircraft failure probability at the required level $P_{f_{required}}$ with the minimum inspections n_{TIP} undertaken in service interval $[0, t_{SL}]$ ($t_{n_{TIP}} = t_{SL}$). In mathematical terms that can be presented as:

$$\sup_{\theta}\left(P_f\left(\theta, \vec{t}\,\right)\right) \le P_{f_{required}}, \qquad 11.$$

where

$$P_f\left(\theta, \vec{t}\,\right) = \sum_{i=1}^{n_{TIP}} P\left(T_{i-1} \le T_d \le T_c < T_i\right). \qquad 12.$$

In expression 12 $T_1, T_2, \ldots, T_{n_{TIP}}$ are time moments of inspections: random value

$$T = \left(T_1, T_2, \cdots, T_{n_{TIP}}\right) = \vec{t}\,(\hat{\theta}), \qquad 13.$$

where $T_0 = 0$, $T_{n_{TIP}} = t_{SL}$, and $n_{TIP} = 0,\ 1,\ 2,\ \ldots$. The expression $n_{TIP} = \infty$ symbolically means that the aircraft must be returned for redesign to the design office.

The inspection program definition vector $\vec{t}\left(\hat{\theta}\right)$ is a function, where both number of inspections during service interval n_{TIP} and disposition of inspection time moments $T_1, T_2, \ldots, T_{n_{TIP}}$ during $[0, t_{SL}]$ are to be chosen as a function of $\hat{\theta}$ and some limitations. It is clear, that there might be many ways how to position inspection time moments on $[0, t_{SL}]$ for a particular n_{TIP}. Let us apply the following inspection time moment disposition rule R_D: the time of the first inspection T_1 (T_1 is a random value because it is a function of $\hat{\theta}$) will be defined by procedure similar to the safe life approach (probability of failure without inspections is less than some small value P_{f1}), while all remaining inspections are distributed evenly in the interval $[T_1, t_{SL}]$. Of course, this rule R_D in general case does not minimise the total required number of inspections n_{TIP}; there are other rules that are more optimal, but our choice of rule R_D is caused by its simplicity for further applications; inspection programs created by this rule are currently used in practice for commercial jet aircrafts.

To apply a particular rule R_D we have to find the total required number of inspections n_{TIP}, which depends on the limiting value of the failure probability $P_{f_{required}} = 1 - R_{required}$, where required reliability $R_{required}$ is mandated, for example, by JAR regulations.

As it was shown above, the failure probability is a function of the number of inspections n and parameter θ; let us denote it as $P_f(\theta, n)$. We also suppose that $P_f(\theta, n)$ monotonically decreases when the number of inspections n increases (at least when n is large enough) and $\lim_{n\to\infty} P_f(\theta, n) = 0$ for all θ. Let n_{TIP} is a solution of the equation

$$P_f(\theta, n) = P_{f_{required}} . \quad 14.$$

Then let us denote

$$n_{TIP} = P_f^{-1}\left(\theta, P_{f_{required}}\right) = n\left(\theta, P_{f_{required}}\right) \quad 15.$$

as the minimal inspection number at which failure probability $P_f(\theta, n_{TIP}) \le P_{f_{required}}$. But the true value of the θ in unknown, so $\hat{n}_{TIP} = n\left(\hat{\theta}, P_{f_{required}}\right)$ and $\hat{P}_f = P_f(\theta, \hat{n}_{TIP})$ are random values. We suppose that we begin the commercial production and operation of aircrafts only if some specific requirements to reliability are met. For the simplest case there is a limitation for the maximum allowed number of inspections $n_{\max}$: we will return airframe project for redesign as unprofitable in case, when the required number of inspections in the inspection program n_{TIP} exceeds $n_{\max}$ (we need to inspect aircraft too often to ensure required reliability). It can be assumed, that the probability of failure for the returned projects is equal to zero, i.e.

$$\hat{P}_{f_{corrected}} = \begin{cases} P_f(\theta, \hat{n}_{TIP}) & , \hat{n}_{TIP} \le n_{\max} \\ 0 & , \hat{n}_{TIP} > n_{\max} \end{cases} . \quad 16.$$

In the more complex case there is a set of limitations. For example, in addition to limitation on the expected number of inspections $n_{calculated} = \hat{n}_{TIP}$ we will return airframe project for redesign if estimate of expectation value of T_c ($T_{c_{calculated}}$) is too small in comparison with t_{SL} (breaking minimum threshold $T_{c_{\min}}$); if estimate of time between two consequent inspections $\Delta t_{calculated}$ is smaller than a threshold $\Delta t_{\min}$; if estimate of initial equivalent crack size $\alpha_{calculated}$ exceeds crack detectable size $a_{\det}$ and so on. Let us denote the vector of calculated values of limiting values $\vec{d}_L = \vec{d}_L(\hat{\theta})$ as

$$\vec{d}_L = \begin{cases} n_{calculated} \\ \Delta t_{calculated} \\ T_{c_{calculated}} \\ \alpha_{calculated} \end{cases} , \quad 17.$$

and the set of its allowed values D_L as

$$D_L = \begin{cases} (0, n_{\max}] \\ [\Delta t_{\min}, \infty) \\ [T_{c_{\min}}, \infty) \\ [0, a_{\det}] \end{cases} . \quad 18.$$

Actually, the number of elements in $\vec{d}_L$ and, therefore, the number of dimensions in the set of the allowed values D_L may vary depending on modelling situation and specific requirements. For example, for inspection programs with the equal time between inspections in the whole service interval $[0, t_{SL}]$ the time between two consequent inspections $\Delta t = t_{SL} / n$, so it can be excluded from the set of limitations, but it is important in programs when the time between inspections may vary.

If vector of limiting values $\vec{d}_L$ does not match the set of its allowed values D_L, then the project is considered as unprofitably and is returned back for redesign in the design office. As we stated above, the probability of failure for returned projects is equal to zero, thus

$$\hat{P}_{f_{corrected}} = \begin{cases} P_f(\theta, \vec{d}_L) & , \vec{d}_L \in D_L \\ 0 & , \vec{d}_L \notin D_L \end{cases} . \quad 19.$$

The parameter θ, which defines the c.d.f. of vector (T_d, T_c), is a vector parameter. For considered case in this work, if both crack model parameters are random and have normal distribution, it consists of five components:

$$\theta = \left[\theta_{0_{\ln Cc}}, \theta_{1_{\ln Cc}}, \theta_{0_{\ln Q}}, \theta_{1_{\ln Q}}, r\right], \tag{20.}$$

where θ_0 stands for a location and θ_1 stands for a scale parameter of the appropriate crack growth model parameter $\ln C_c$ or $\ln Q$; r is a coefficient of correlation between $\ln C_c$ and $\ln Q$, and $\theta \in \Theta = \left\{ (-\infty,\infty);\ [0,\infty);\ (-\infty,\infty);\ [0,\infty);\ [0,1] \right\}$. As it was shown before, we don't know the real value of θ, thus we use its estimate $\hat{\theta}$. A part of elements of $\hat{\theta}$ may be assumed as known. For example, $\theta_{1_{\ln Cc}}$, $\theta_{1_{\ln Q}}$ and correlation coefficient *r* can be considered as constants, so processing fatigue crack growth data we should estimate only two remaining parameters $\theta_{0_{\ln Cc}}$ and $\theta_{0_{\ln Q}}$.

It can be shown that for considered decision making procedure random variable $\hat{P}_{f_{corrected}}$ has expectation value, which is a function of θ, and this function has a maximum value for $\theta \in \Theta$. To prove that let us fix one of two crack model parameters and look how the probability of failure depends on another one. Let us consider that the equivalent initial crack size is a constant: $\alpha = const$, i.e. $\theta_{0_\alpha} = \mu_\alpha = const$, $\theta_{1_\alpha} = \sigma_\alpha = 0$.

In accordance with upper defined rules the probability of failure tends to zero when the crack growth speed representing parameter $\theta_{0_{\ln Q}} = \mathrm{E}\{\ln Q\}$ tends to zero: this is a case when the item is extremely reliable and cracks are growing so slowly, that have no chance to grow up to a_{crit} in interval $[0, t_{SL}]$, thus there are no inspections required. The failure probability without inspections is defined by formula:

$$P_{f_{wi}} = P(T_c \le t_{SL}) = \Phi\left(\frac{\ln t_{SL} - \mu_{\ln T_c}}{\sigma_{\ln T_c}}\right), \tag{21.}$$

or, in terms of reliability,

$$R_{wi} = P(T_c > t_{SL}) = 1 - P(T_c \le t_{SL}) = 1 - \Phi\left(\frac{\ln t_{SL} - \mu_{\ln T_c}}{\sigma_{\ln T_c}}\right), \tag{22.}$$

where $\ln T_c$ is distributed normally as $\ln T_c \sim N(\mu_{\ln T_c}; \sigma^2_{\ln T_c})$.

From other side, if the $\theta_{0_{\ln Q}}$ is high, then the probability of failure tends to zero as well: with high probability we return for redesign all items due to the break of limiting rules, i.e. $\vec{d}_L \notin D_L$ (see formulas 17, 18 and 19). Between these zero values of $\mathrm{E}\left\{P_{f_{corrected}}\right\}$ there can be non-zero values somewhere in between, when the fatigue cracks maybe can reach theirs critical size during the time between inspections, maybe not, but there are no sufficient reasons to return project for redesign so far. Let us call a value of failure probability used for calculations (at the choice of the number of inspections required, or choosing vector-function $\vec{t}$) as $P_{f_{calc}}$. The following conclusion can be made from the upper mentioned: the dependence of the probability of failure as a function of θ is a function which has a maximum, the value of that maximal value is unknown, but somehow depends on the value of failure probability $P_{f_{calc}}$ used for calculations.

Let us call the value of expectation of failure probability for all allowable θ as $\mathrm{E}\left\{\hat{P}_{f_{corrected}}\right\}$. We have named it as "corrected" to distinguish it from $P_{f_{calc}}$, because we take into consideration some limitations. The goal is to find such a maximum value of failure probability for calculations $P^*_{f_{calc}}$ that the corrected value of failure probability $P_{f_{corrected}}$ does not exceed the required limiting value of failure probability $P_{f_{required}} = 1 - R_{required}$:

$$P^*_{f_{calc}} : \widetilde{P}\left(P_{f_{calc}}\right) \le P_{f_{required}} , \qquad 23.$$

where

$$\widetilde{P}\left(P_{f_{calc}}\right) = \max_{\theta}\left(\mathrm{E}_{\theta}\left\{ \ \hat{P}_{f_{corrected}} \ \right\}\right). \qquad 24.$$

Graphically this approach for a two-dimensional case (when either $\alpha = const$ or $Q = const$) is presented in Figure 3:

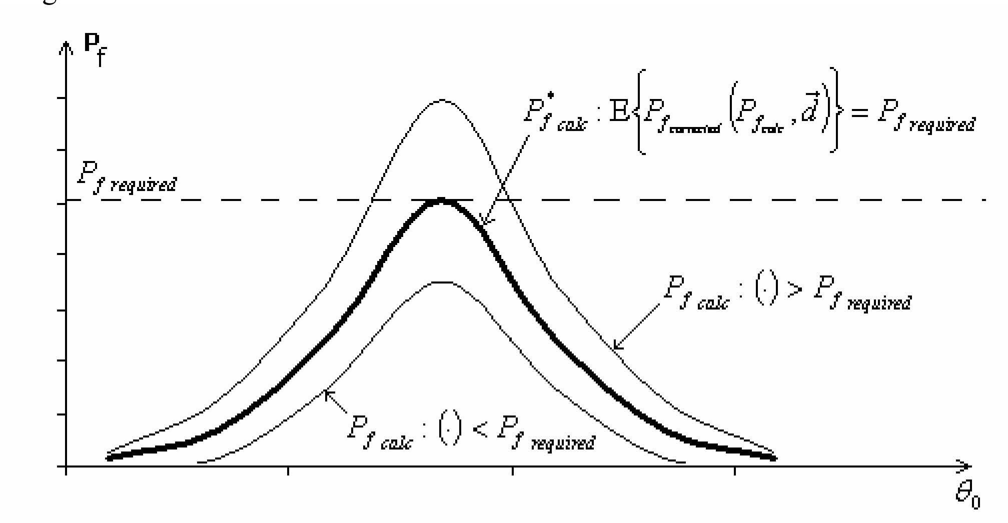

P_f
$P^*_{f\ calc} : \mathrm{E}\left\{P_{f_{corrected}}\left(P_{f_{calc}}, \vec{d}\right)\right\} = P_{f\ required}$
$P_{f\ required}$
$P_{f\ calc} : (\cdot) > P_{f\ required}$
$P_{f\ calc} : (\cdot) < P_{f\ required}$
θ_0

Figure 3. Minimax approach example ($\alpha = const$ or $\ln Q = const$)

For the more complex case we get a three-dimensional picture like in

Figure *4*:

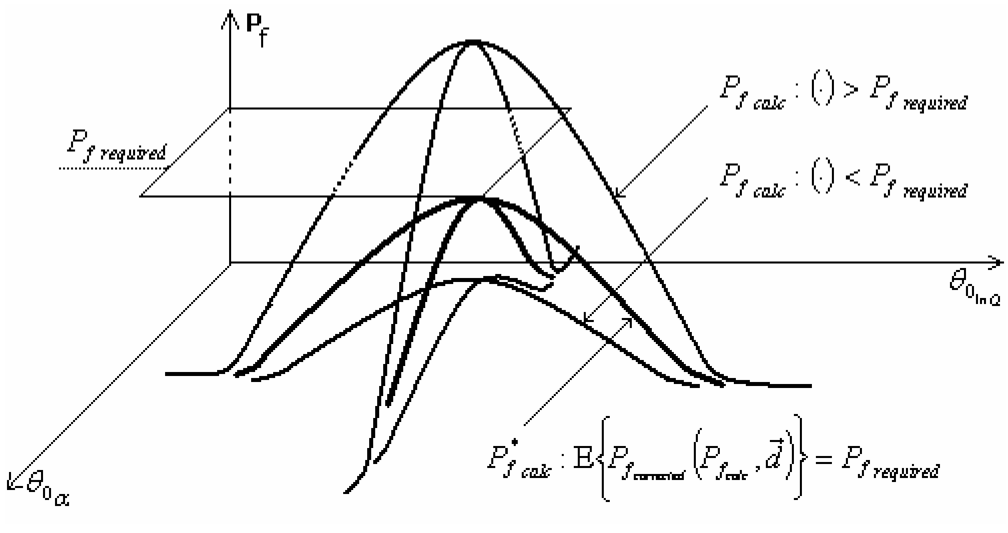

P_f
$P_{f\ required}$
$P_{f\ calc} : (\cdot) > P_{f\ required}$
$P_{f\ calc} : (\cdot) < P_{f\ required}$
$\theta_{0_{\ln Q}}$
$\theta_{0_{\alpha}}$
$P^*_{f\ calc} : \mathrm{E}\left\{P_{f_{corrected}}\left(P_{f_{calc}}, \vec{d}\right)\right\} = P_{f\ required}$

Figure 4. Minimax approach example (general case)

Depending on parameters the shapes of these two- or three-dimensional failure probability curves may vary, but this does not affect our conclusions.

NUMERICAL EXAMPLE AND CONCLUSION

The upper mentioned approach lets us to ensure reliability of the airframe on or above the required level by developing appropriate inspection program for the case of lack of the initial fatigue test data. There are examples of numerical modelling for one- and two-parametric models shown in Figure 5 and

Figure *6* below (please note: in pictures LQ=LN(Q)=$\theta_{0_{\ln Q}}$, LC=LN(Cc)=$\theta_{0_{Cc}}$).

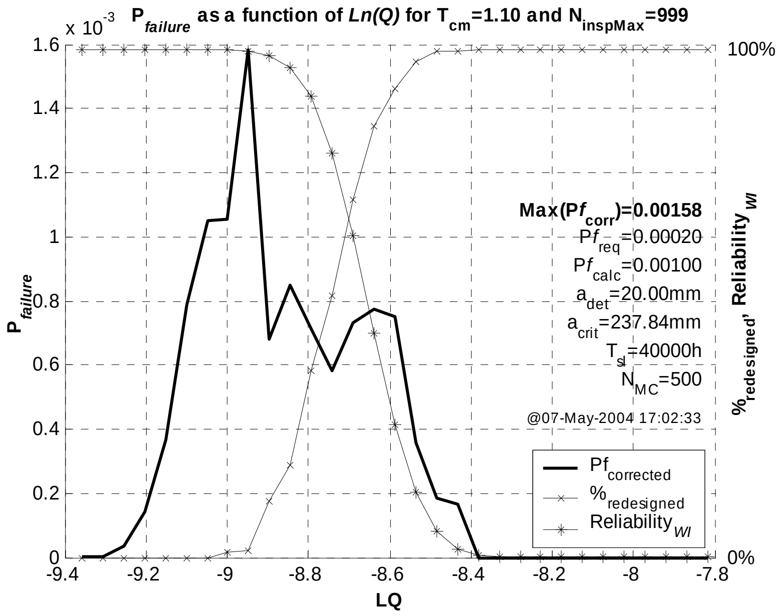


Figure 5. One-parametric numerical example ($\alpha = const$)

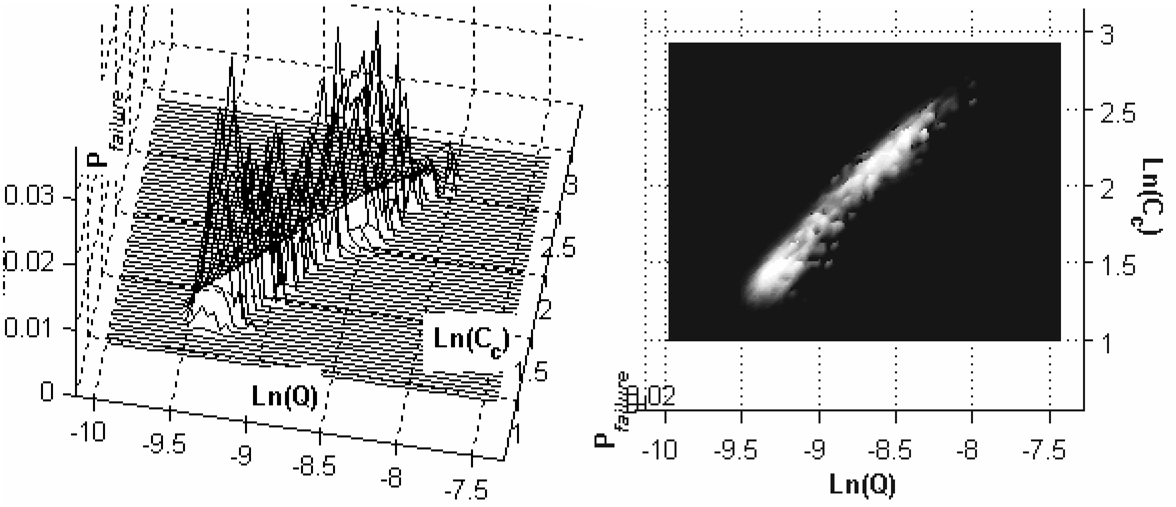


Figure 6. Two-parametric numerical example (3D and projection)

REFERENCES

1. MSG-3. Operator/Manufacturer Scheduled Maintenance Development. Revision 2003.1. // Air Transport Association of America, Inc. Washington, D.C., 2003.
2. Y.Paramonov, A.Kuznetsov. Fatigue crack growth parameter estimation by processing inspection results // In: Third International conference on Mathematical Methods in Reliability MMR2002 – Trondheim, Norway, 2002.
3. Y.Paramonov, A.Kuznetsov. Inspection data use for airframe inspection interval correction // In: Aviation – Issue #6 – Vilnius, Technika, 2002.
4. Y.Paramonov, A.Kuznetsov. Planning of inspections of fatigue-prone airframe // In: Proceedings of International conference "Diagnosis of technical systems, numerical and physical non-destructive quality testing – 2004" – Vilnius, April 23, 2004.
5. Y.Paramonov, A.Kuznetsov. Switching to doubled aircraft inspection frequency strategy analysis for exponential fatigue crack growth model // In: Proceedings of International conference on Longevity, Ageing and Degradation models in Reliability, Medicine and Biology (LAD2004), Volume 1, pp.143-154 – St. Petersburg, Russia, 2004.

STOCHASTIC APPROACH TO SAFETY AT SEA ASSESSMENT

Finkelstein M.S.
●
Department of Mathematical Statistics,University of the Free State
PO Box 339, 9300, Bloemfontein, Republic of South Africa,
e-mail: msf@wwg3.uovs.ac.za
and
CRSI "Elektropribor", St. Petersburg, Russia

Abstract. A general approach for analysing spatial survival in the plane is suggested. Two types of harmful random events are considered: points with fixed coordinates and moving points. A small normally or tangentially oriented interval is moving along a fixed route in the plane, crossing points of initial Poisson random processes. Each crossing leads to termination of the process with a given probability. The probability of passing the route without termination is derived. A safety at sea application is discussed.

Keywords: Spatial point process, Survival probability, Random field, Rate of occurrence.

1. INTRODUCTION: ONE-DIMENSIONAL CASE

A model of survival in the plane is presented in, based on the following simple reasoning used in the one-dimensional case. Consider a system subject to stochastic point influences (shocks). Each shock can lead with a given probability to a fatal failure of a system, resulting in termination of the process, and this will be called an "accident". The probability of performance without accidents in the time interval $(0,t]$ is of interest. It is natural to describe the situation in terms of stochastic point processes.

Denote by $h(t)$ the rate of occurrence or just the rate function of the corresponding point process of shocks $\{N(t);t>0\}$. It is well known ([2, p.31]) that for orderly processes, assuming the limits exist,

$$h(t)=\lim_{\Delta t\to 0}\frac{\Pr\{N(t,t+\Delta t)=1\}}{\Delta t}=\lim_{\Delta t\to 0}\frac{E[N(t,t+\Delta t)]}{\Delta t}. \tag{1}$$

Assume now that a shock occurring in $(t,t+dt]$ independently of the previous shocks leads to an accident with probability $\theta(t)$, and does not cause any changes in the system with probability $1-\theta(t)$. Denote by T_a a random time to an accident and by $F_a(t)=\Pr\{T_a\le t\}$ the corresponding distribution function (DF). If $F_a(t)$ is absolutely continuous, then

$$P(t)=1-F_a(t)=\exp\left\{-\int_0^t\lambda_a(x)dx\right\}, \tag{2}$$

where $\lambda_a(t)$ is a hazard rate, corresponding to $F_a(t)$ and $P(t)$ is the survival function: probability of performance without accidents in $(0,t]$. Assuming that $\{N(t);t>0\}$ is the nonhmogeneous Poisson processes:

$$\lambda_a(t)=\theta(t)h(t). \tag{3}$$

For the time-dependent case this result was proved in Block et al. [1]. Considering the Poisson point processes in the plane, we shall construct the corresponding hazard rate "along the fixed curve". An obvious application of this model is assessing the probability of a safe performance of a ship moving along a fixed route [4].

2. OBSTACLES WITH FIXED COORDINATES

Denote by $\{N(B)\}$ the nomhomogeneous Poisson point process in the plane (the random number of points in $B \subset \Re^2$, where B belongs to the Borel σ-algebra in $\Re^2$). We shall consider points as prospective point influences on our system (shallows for the ship, for instance). Similar to the one-dimensional definition (1) the rate $h_f(\xi)$ can be formally defined as

$$h_f(\xi) = \lim_{S(\delta(\xi))\to 0} \frac{E[N(\delta(\xi))]}{S(\delta(\xi))}, \tag{4}$$

where $B = \delta(\xi)$ is the neighborhood of ξ with area $S(\delta(\xi))$ and diameter tending to zero.

Assume for simplicity that $h_f(\xi)$ is a continuous function of ξ in an arbitrary closed circle in $\Re^2$. Let R_{ξ_1,ξ_2} be a fixed continuous curve connecting ξ_1 and ξ_2 - two distinct points in the plane. We shall call R_{ξ_1,ξ_2} a route. A point (a ship in our application) is moving in one direction along the route. Every time it "crosses the point" of process $\{N(B)\}$ an accident can happen with a given probability. We are interested in assessing probability of moving along R_{ξ_1,ξ_2} without accidents. Let r be the distance from ξ_1 to the current point of the route (coordinate) and $h_f(r)$ denote the rate in $(r, r+dr]$(a one-dimensional parameterization).

Let $(\gamma_n^+(r), \gamma_n^-(r))$ be a small interval of length $\gamma_n(r) = \gamma_n^+(r) + \gamma_n^-(r)$ in a normal to R_{ξ_1,ξ_2} in the point with coordinate r, where upper indexes denote the corresponding direction Let $\overline{R}$ be the length of $R_{\xi_1,\xi_2} : \overline{R} \equiv | R_{\xi_1\xi_2} |$ and assume that: $\overline{R} >> \gamma_n(r), \forall r \in [0, R]$. The interval $(\gamma_n^+(r), \gamma_n^-(r))$ is moving along R_{ξ_1,ξ_2}, crossing points of a random field. (For our application it is reasonable to assume the following model for the symmetrical $(\gamma_n^+(r) = \gamma_n^-(r))$ equivalent interval: $\gamma_n(r) = 2\delta_s + 2\delta_o(r)$, where $2\delta_s, 2\delta_o(r)$ are the diameters of a ship and of an obstacle, respectively, and for simplicity it is assumed that all obstacles have the same diameter. There can be other models as well). Using definition (4), the *equivalent rate* of occurrence of points, $h_{e,f}(r)$ along the route can be defined as

$$h_{ef}(r) = \lim_{\Delta r\to 0} \frac{E[N(B(r, \Delta r, \gamma_n(r))]}{\Delta r}, \tag{5}$$

where $N(B(r, \Delta r, \gamma_n(r))$ is the random number of points crossed by the interval $\gamma_n(r)$, moving from r to $r + \Delta r$.

It can be easily seen, as in Finkelstein [4], that for $\Delta r \to 0$ and $\gamma_n(r)$ sufficiently small:

$$E[N(B(r, \Delta r, \gamma_n(r))] = \int_{B(r,\Delta r,\gamma_n(r))} h_f(\xi) dS(\delta(\xi)) \overset{\Delta r\to 0}{=} \gamma_n(r) h_f(r) dr[1 + o(1)],$$

which leads to the expected relation for the equivalent rate of the corresponding one-dimensional point process (which is obviously also nonhomogeneous Poisson):

$$h_{ef}(r) = \gamma_n(r) h_f(r)[1 + o(1)]. \tag{6}$$

Hence, r-parameterization along the fixed route reduces the problem to the one-dimensional setting of Section 1.

As in the one-dimensional case, assume that crossing of a point with a coordinate r leads to an accident with probability $\theta_f(r)$ (and is survived with probability $\overline{\theta}_f(r) = 1 - \theta_f(r)$). Denote by R a random distance from the initial point of the route ξ_1 till a point on the route where an accident had occurred. Similar to (2)-(3), probability of passing the route R_{ξ_1,ξ_2} without accidents can be derived in the following way:

$$\Pr\{R > \overline{R}\} \equiv P(\overline{R}) = 1 - F_{af}(\overline{R}) = \exp\left\{-\int_0^{\overline{R}} \lambda_{af}(r) dr\right\} \tag{7}$$

$$\lambda_{af}(r) \equiv \theta_f(r)h_{ef}(r) \tag{8}$$

Assume that the hazard rate $\lambda_{af}(r)$ is now a stochastic process defined, for instance, as in Yashin and Manton [8] by an unobserved covariate stochastic process $Y = Y_r, r \ge 0$. Denote the corresponding hazard rate process by $\lambda_{af}(Y,r)$. It is well known (see, e.g. Kebir [7]) that under certain assumptions in this the following equation holds:

$$P(\overline{R}) = E\left[\exp\left\{-\int_0^{\overline{R}} \lambda_{af}(Y,r)dr\right\}\right], \tag{9}$$

which can be written via the conditional hazard rate process [8] as

$$P(\overline{R}) = \exp\left\{-\int_0^{\overline{R}} E\left[\lambda_{af}(Y,r) \mid R > r\right]dr\right\} = \exp\left\{-\int_0^{\overline{R}} \overline{\lambda}_{af}(r)dr\right\}, \tag{10}$$

where $\overline{\lambda}_{af}(r)$ is the corresponding equivalent or observed hazard rate:

$$\overline{\lambda}_{af}(r) = E\left[\lambda_{af}(Y,r) \mid R > r\right]. \tag{11}$$

As follows from (10), equation (11) can constitute a reasonable tool for obtaining $P(\overline{R})$, but the corresponding explicit derivations can be performed only in some simplest specific cases. On the other hand, it can help to analyze some important properties. Assume, for instance, that probability $\theta_f(r)$ is indexed by a parameter $Y : \theta_f(Y,r)$. Let Y be interpreted as a non-negative random variable with support in $[0,\infty)$ and the probability density function $\pi(y)$. In the sea safety application this randomization can be due to the unknown characteristics of the navigation (or (and) collision avoidance) onboard system, for instance (we are pooling from the population of ships). There can be other interpretations as well. Thus, the specific case, when Y in relations (9) and (10) is a random variable, is considered. The observed failure rate $\overline{\lambda}_{af}(r)$ then is the corresponding mixture failure rate:

$$\overline{\lambda}_{af}(r) = \int_0^\infty \lambda_{af}(y,r)\pi(y \mid r)dy, \tag{12}$$

where $\pi(y \mid r)$ is the conditional probability density function of Y given that $R > r$ [5]

$$\pi(y \mid r) = \frac{\pi(y)P(y,r)}{\int_0^\infty P(y,r)\pi(y)dy}, \tag{13}$$

and $P(y,r)$ is defined similar to (7), where $\lambda_{af}(r)$ is substituted by $\theta_f(y,r)h_{ef}(r)$.

Relations (12) and (13) constitute a tool for analyzing the shape of the observed failure rate $\overline{\lambda}_{af}(r)$. As shown in [5,6], the shape of $\overline{\lambda}_{af}(r)$ can differ dramatically from the shape of the conditional failure rate $\lambda_{af}(y,r)$ and this fact should be taken into consideration in applications. Assume, for example, a specific multiplicative form of parameterization:

$$\theta_f(Y,r)h_{ef}(r) = Y\theta_f(r)h_{ef}(r).$$

It is well known that, if $\theta_f(r)h_f(r)$ is constant in this case, than the observed failure rate is decreasing. But it turns out that even, if $\theta_f(r)h_{ef}(r)$ is sharply increasing, $\overline{\lambda}_{af}(r)$ can still decrease at least for sufficiently large r! [6]. Thus, the random parameter changes the aging properties of the corresponding distribution functions.

For the "highly reliable systems" when, for instance, $h_f(r) \to 0$ uniformly in $r \in [0, \overline{R}]$, one can easily obtain obvious approximations. On the other hand, applying Jensen's inequality to the right hand side of (9), a simple lower bound for $P(\overline{R})$ can be also derived:

$$P(\overline{R}) \geq \exp\left\{E\left[-\int_0^{\overline{R}} \lambda_{af}(Y,r)dr\right]\right\} = \exp\left\{-\int_0^{\overline{R}} E[\lambda_{af}(Y,r)]dr\right\}. \tag{14}$$

3. CROSSING THE LINE PROCESS

Consider a random process of continuous curves in the plane to be called paths. We shall keep in mind an application when ships' routes on a chart represent paths, while the rate of the stochastic processes to be defined represents the intensity of navigation in the given sea area. The specific case of stationary random lines in the plane is called a *stationary line process.*

It is convenient to characterize a line in the plane by its (ρ,ψ) coordinates, where ρ is the perpendicular distance from the line to a fixed origin, and ψ is the angle between this perpendicular line and a fixed reference direction. A random process of undirected lines can be defined as a point process on the cylinder $\Re_+ \times S$, where $\Re_+ = (0,\infty)$ and S denote both the circle group and its representations as $(0,2\pi]$. Thus each point on the cylinder is equivalent to the line in $\Re^2$ and for the finite case the point process (and associated stationary line process) can be described. The following result is stated in Daley and Vere-Jones [5, p.389]. Let V be a fixed line in $\Re^2$ with coordinates (ρ_v,α) and let N_V be the point process on V *generated by its intersections with the stationary line process.* Then N_V is a stationary point process on V with rate h_V given by

$$h_V = h\int_S |\cos(\psi-\alpha)| P(d\psi), \tag{15}$$

where h is the constant rate of the stationary line process and $P(d\psi)$ is the probability that an arbitrary line has orientation ψ (first order directional rose on S). If the line process is isotropic, then $h_V = 2h/\pi$. The rate h is induced by the random measure defined by the total length of lines inside any closed bounded convex set in $\Re^2$. Assume that the line process is (homogeneous) Poisson in the sense that the point process N_V generated by its intersections with an arbitrary V is a Poisson point process.

Consider now a stationary temporal Poisson line process in the plane. Similar to N_V, the Poisson point process $\{N_V(t); t>0\}$ of its intersections with V in time can be defined. The constant rate of this process, $h_V(1)$, as usual, defines the probability of intersection (by a line from a temporal line process) of an interval of a unit length in V and in a unit interval of time given these units are substantially small.

Let V_{ξ_1,ξ_2} be a finite line route, connecting ξ_1 and ξ_2 in $\Re^2$ and r, as in the previous section, is the distance from ξ_1 to the current point of V_{ξ_1,ξ_2}. Then $h_V(1)drdt$ is the probability of intersecting V_{ξ_1,ξ_2} by the temporal line process in $(r,r+dr)\times(t,t+dt); \forall r \in (0,\overline{R}),\ t>0$.

A point (a ship) starts moving along V_{ξ_1,ξ_2} at $\xi_1, t=0$ with a given speed $v(t)$. We assume that an accident happens with a given probability when "it intersects" the line from the (initial) temporal line process. A regularization procedure, involving dimensions (of a ship, in particular) can be performed in the following way: an attraction interval

$$(r-\gamma_{ta}^-, r+\gamma_{ta}^+) \subset V_{\xi_1,\xi_2},\ \ \gamma_{ta}^+,\gamma_{ta}^- \geq 0,\ \ \gamma_{ta}(r) = \gamma_{ta}^+(r) + \gamma_{ta}^-(r) << \overline{R},$$

where the subscript "ta" stands for tangential, is introduced. The attraction interval (which can be defined by the ship's dimensions) is moving along the route, attached to the point itself with changing in time coordinate:

$$r(t) = \int_0^t v(s)ds, \quad t \leq t_{\overline{R}}\ , \tag{16}$$

where $t_{\overline{R}}$ is the total time on the route. Similar to (5), we can construct the *equivalent rate of intersections*, $h_{e,m}(r)$, assuming for simplicity constant speed $v(t) = v_0$ and γ_{ta}:

$$E[N_V((r, r+\Delta r), \Delta t)] = h_V(1)\Delta r\Delta t\ . \tag{17}$$

Thus the equivalent rate is also constant

$$h_{em} = \Delta t\, h_V(1) = \frac{\gamma_{ta}}{v_0} h_V(1), \tag{18}$$

where $\Delta t = \dfrac{\gamma_{ta}}{v_0}$ is the time needed for the moving attraction interval to pass the interval $(r, r+\Delta r)$ as $\Delta r \to 0$. As assumed earlier, the intersection can lead to an accident. Let the corresponding probability of an accident θ_m be also constant. Then, using results of sections 1 and 2, the probability of moving along the route V_{ξ_1,ξ_2} without accidents is:

$$P(\overline{R}) = \exp\{-\theta_m h_{e,m}\overline{R}\}\ , \tag{19}$$

The non-linear generalization is rather straightforward. The line route V_{ξ_1,ξ_2} turns into the continuous curve R_{ξ_1,ξ_2} and lines of the stochastic line process turn also into continuous curves. Eventually

$$P(\overline{R}) = \exp\left\{-\int_0^{\overline{R}} \theta_m(r) h_{em}(r) dr\right\} \tag{20}$$

and assuming independence of fixed and moving obstacles, relations (7) and (20) can be combined in an obvious way.

REFERENCES

1. Block, H.W., Savits, T.H., and Borges, W. (1985). Age dependent minimal repair. *J. Appl.* Prob. **22**, 370-386.
2. Cox, D.R., and Isham, V. (1980) Point processes, London, Chapman and Hall.
3. Daley D.J., and Vere-Jones D. (1988) An introduction to the theory of point processes, New York, Springer –Verlag.
4. Finkelstein, M.S. (1998). A point process stochastic model of safety at sea. *Reliability Engineering and System Safety*, **60**, 227-233.
5. Finkelstein, M. S., and Esaulova, V. (2001). On inverse problem in mixture failure rate modeling. *Applied Stochastic Models in Business and Industry*, **17**, 221-229.
6. Finkelstein, M. S., and Esaulova, V. (2001). Modeling a failure rate for a mixture of distribution functions. *Probabaility in the Engineering and Informational Sciences*, **15**, 383-400.
7. Kebir, Y. (1991). On hazard rate processes. *Naval Research Logistics*, **38,** 865-876.
8. Yashin, A.I., and Manton, K.G. (1997). Effects of unobserved and partially observed covariate processes on system failure: a review of models and estimation strategies. *Statistical Science*, **12**, 20-34.

MEASURING RISK

Novosyolov A.
●
Institute of computational modeling SB RAS
Krasnoyarsk, Russia, anov@ksc.krasn.ru

Abstract. Problem of representation of human preferences among uncertain outcomes by functionals (risk measures) is being considered in the paper. Some known risk measures are presented: expected utility, distorted probability and value-at-risk. Properties of the measures are stated and interrelations between them are established. A number of methods for obtaining new risk measures from known ones are also proposed: calculating mixtures and extremal values over given families of risk measures.

Keywords: risk, risk measure, preference, expected utility, distorted probability, value-at-risk, mixture transform, extremal transforms.

INTRODUCTION

Quantifying risk is one of central problems in risk theory [1,2]. Risk measures are commonly used for the purpose. As of now there is a vast amount of different risk measures, including simple probability of an adverse event, second order measures (variance or standard deviation, beta) [3], quantile measures such as value-at-risk and its derivatives: conditional value-at-risk [4], expected shortfall [5] and some modifications.

A classic expected utility [6] is used more and more intensively in financial markets and other decision-making applications. Expected utility measures do form a wide class of risk measures, thus providing a flexible tool for decision-making under uncertainty. However any measure in this class is linear with respect to mixture of probability distributions, that may be undesirable in some cases.

Another wide class of risk measures was introdiced in [7]; these are the so called distorted probability measures. Fortunately they turned out to be nonlinear with respect to mixtures, so that they can represent human preferences in a more reliable fashion. One more attractive feature of risk measures in this class is that they are closely connected with other measures by some natural transforms. The latter provokes rigorous studying of the measures as possibly most appropriate tool for risk theory applications.

In the present paper we briefly describe properties of risk measures mentioned above and point out to interrelations among them. The following notation will be used throughout the paper. $(\Omega, \mathbf{A}, P)$ denotes a probability space with σ-algebra $\mathbf{A}$ and probability measure P; the latter may vary in some cases. *Risks* are represented by random variables, that is, measurable mappings from Ω to the measurable space $(\mathbf{R}, \mathbf{B})$, where $\mathbf{R}$ is the set of real numbers and $\mathbf{B}$ is a σ-algebra of its Borel subsets. Risks will be denoted by $X, Y, \ldots$ while their distribution functions by F_X, F_Y, etc. Denote also $\mathbf{X}$ the set of all risks and $\mathbf{F}$ the set of all distribution functions. *Risk measure* is any functional on $\mathbf{F}$. Introduce also partial orderings on $\mathbf{F}$ known as stochastic dominance of different orders. For a distribution function $F \in \mathbf{F}$ let $F^1 = F$ and $F^{(k)}$ for $k = 2,3,...$ are defined iteratively by

$$F^{(k)}(x) = \int_{-\infty}^{x} F^{(k-1)}(t)dt, \ x \in \mathbf{R}.$$

For $F, G \in \mathbf{F}$ we say that F preceeds G in the sense of stochastic dominance of order k ($F \leq_k G$) if $F^{(k)}(x) \geq G^{(k)}(x)$, $x \in \mathbf{R}$. For future reference denote W_a a degenerate (at a point $a \in \mathbf{R}$) distribution function and B_p a Bernoulli (with parameter $p \in (0,1)$) distribution function. Let also $\mathbf{W} = \{W_a, a \in \mathbf{R}\}$ be the class of all degenerate distribution functions, and $\mathbf{Be} = \{B_p, p \in (0,1)\}$ be the class of all Bernoulli distribution functions.

RISK MEASURES AND PREFERENCE

Let $\prec$ be a preference relation on $\mathbf{F}$, that is, a complete and transitive binary relation. We say that risk measure $\mu : \mathbf{F} \to \mathbf{R}$ represents the preference relation if for $F, G \in \mathbf{F}$

$$F \prec G \Leftrightarrow \mu(F) \le \mu(G) \tag{1}$$

A perfect risk measure should represent preferences of specific individual or decision-maker. Since it is very unlikely that preference relation is known completely, representation theorems are usually based on reasonable assumptions that restrict the collection of available preferences to a rather narrow class, and provide an analytical form for risk measures representing preferences from that class.

EXPECTED UTILITY MEASURE

Perhaps the first representation theorem of the sort is due to von Neumann and Morgenstern [6]. It states that under some assumptions (that actually mean linearity of preference with respect to mixture of distributions) there exists the unique (up to positive affine transforms) risk measure representing the relation. The resulting risk measure turns out to be the so called expected utility, that was well known for about 3 hundred years already. It has the form

$$\rho(F) = \int_{-\infty}^{\infty} U(x)dF(x),\ F \in \mathbf{F}, \tag{2}$$

where U stands for utility function. Different preferences correspond to different utility functions. A disadvantage of risk measure (2) is that it is always linear with respect to mixture of distributions, that is, for any $F, G \in \mathbf{F}$ and any $\alpha \in [0,1]$ the following is always true:

$$\rho(\alpha F + (1-\alpha)G) = \alpha\rho(F) + (1-\alpha)\rho(G).$$

Experiments (eg. [8]) show that in many cases human preferences do not possess linearity, so risk measure (2) might be a very rough approximation to what is actually needed.

Let us state some properties of expected utility here. The following theorem may be found e.g. in [9].

Theorem 1. Expected utility ρ is monotone with respect to stochastic dominance of the 1st order if and only if the utility function U is nondecreasing. Expected utility ρ is monotone with respect to stochastic dominance of the 2nd order if and only if the utility function U is nondecreasing and concave.

A question of great practical importance is: how much additional information do one need to identify the utility function U that generate expected utility representing a specific linear preference relation? In other words, what is the characteristic class $\mathbf{G} \subseteq \mathbf{F}$ of distribution functions such that there exists the unique continuation of expected utility from $\mathbf{G}$ to $\mathbf{F}$? The final answer is contained in the following

Theorem 2. Let $\prec$ be a linear preference relation on $\mathbf{F}$. To specify the utility function representing $\prec$ it is necessary and sufficient to know values of ρ for all degenerate distributions $W_a, a \in \mathbf{R}$, except for any two of them that may be chosen arbitrarily. This means that $\mathbf{W}$ is essentially the characteristic class for expected utility measure.

It is sometimes more convenient to use the so called certainty equivalent instead of expected utility, The new functional c on $\mathbf{F}$ is defined as a real number posessing the same utility as a distribution, that is:

$$c(F) = U^{-1}(\rho(F)), \quad F \in \mathbf{F}.$$

The functional is well defined if utility function is strictly monotone, that is often the case for preferences monotone with respect to stochastic dominance.

DISTORTED PROBABILITY MEASURE

Let $g:[0,1]\to[0,1]$ be a nondecreasing real function with $g(0)=0, g(1)=1$. A distorted probability measure

$$\pi(F)=\int_{-\infty}^{0}[g(1-F(x))-1]dx+\int_{0}^{\infty}g(1-F(x))dx\,,\quad F\in\mathbf{F} \tag{3}$$

was introduced in [7], [10]. A function g is called a distortion function. In the special case $g(x)=x, x\in[0,1]$ this measure coincides with expectation: $\pi(F)=E_F$, and in all other cases it is essentially nonlinear in distribution. Note that $\pi(W_a)=a$, $a\in\mathbf{R}$ for any parameter function g, and state the representation family theorem for the risk measure.

Theorem 3. Let $\prec$ be a preference relation on $\mathbf{F}$ corresponding to a distorted probability measure. To specify the parameter function g representing $\prec$ it is necessary and sufficient to know values of π for all nondegenerate Bernoulli distributions $B_p, p\in(0,1)$. This means that $\mathbf{Be}$ is essentially the characteristic class for distorted probability measure.

Note that distorted probability measure may be represented in the form

$$\pi(F)=-\int_{0}^{1}F^{-1}(v)dg(1-v)\,,\quad F\in\mathbf{F} \tag{4}$$

Simple consequences of (4) are monotonicity of distorted probability measure with respect to first order stochastic dominance, and the fact that Value-at-risk measure

$$\tau_\lambda(F)=F^{-1}(\lambda)\,,\ F\in\mathbf{F} \tag{5}$$

is a special case of (3) with

$$g_\lambda(v)=\begin{cases}0, & v<1-\lambda\\ 1, & v\geq 1-\lambda\end{cases}$$

where $\lambda\in(0,1)$ is a parameter.

FAMILY-GENERATED RISK MEASURES

Since risk measures are used to represent individual preference among probability distributions, they should catch attitude of an individual to risk. Constructing new risk measures may provide flexible tool for the purpose. In the present section several ways of obtaining new risk measures from given families are presented and studied to some extent.

Let Λ be a parameter set endowed with a structure of probability space $(\Lambda,\mathbf{C},Q)$. Next, let $\mathbf{\Lambda}=\{\mu_\lambda,\lambda\in\Lambda\}$ be a family of risk measures, id est, functionals $\mu_\lambda:\mathbf{F}\to\mathbf{R}$. Consider the following functionals generated using this family.

Mixture risk measure $\mathbf{M}_\Lambda:\mathbf{F}\to\mathbf{R}$.

$$\mathbf{M}_\Lambda(F)=\int_\Lambda \mu_\lambda(F)dQ(\lambda)\,,\quad F\in\mathbf{F}. \tag{6}$$

Maximal risk measure $M^\Lambda:\mathbf{F}\to\mathbf{R}$.

$$M^{\Lambda}(F)=\sup_{\lambda\in\Lambda}\mu_{\lambda}(F), \qquad F\in\mathbf{F}. \tag{7}$$

Minimal risk measure $M_{\Lambda}:\mathbf{F}\to\mathbf{R}$.

$$M_{\Lambda}(F)=\inf_{\lambda\in\Lambda}\mu_{\lambda}(F), \qquad F\in\mathbf{F}. \tag{8}$$

Now let us state some results for these derivative measures.

Theorem 4. Let $\mathbf{\Lambda}$ be a family of risk measures such that each $\mu_{\lambda},\lambda\in\Lambda$ is expected utility measure with utility function U_{λ}. Then mixture risk measure (6) is also an expected utility measure with utility function $U(x)=\int_{\Lambda}U_{\lambda}(x)dQ(\lambda), x\in\mathbf{R}$.

However extremal measures (7) and (8) for a family of expected utilities do not in general constitute an expected utility. Informally the class of expected utilities is closed with respect to mixtures and is not closed with respect to taking exprema.

Theorem 5. Let $\mathbf{\Lambda}$ be a family of risk measures such that each $\mu_{\lambda},\lambda\in\Lambda$ is distorted probability measure with distortion function g_{λ}. Then mixture risk measure (6) is also a distorted probability measure with distortion function $g(v)=\int_{\Lambda}g_{\lambda}(v)dQ(\lambda), v\in[0,1]$.

So the class of distorted probability measures of risk is also closed with respect to mixtures. The following fact is somewhat surprising: any distorted probability measure may be represented by a mixture of Value-at-risk measures, a very special case of distorted probability measures.

Theorem 6. Let $\Lambda=(0,1)$ be endowed with a probability space structure by σ-algebra of Borel subsets and a distribution function G. Let further $\mathbf{\Lambda}$ be a family of Value-at-risk measures (5). Then the mixture risk measure (6) is a distorted probability measure with distortion function $g(v)=1-G(1-v), v\in(0,1)$.

Clearly any distortion function g may be obtained by appropriate choice of mixing distribution function $G(v)=1-g(1-v), v\in(0,1)$. Note that a similar spectral representation of distorted probability measures via expected shortfall family was presented in [5].

REFERENCES

1. A. Novosyolov. *Mathematical Modeling of Financial Risks. Measuring Theory*. Novosibirsk: Nauka, 2001, 102p. (in Russian)
2. A. Novosyolov. Concept of Risk and Methods of its Measuring. Proceedings of the 1st International Scientific School "Modelling and Analysis of Safety, Risk and Quality", St. Petersburg, 2001, p. 77-80.
3. H. Markowitz. *Mean - Variance Analysis in Portfolio Choice and Capital Markets*. Cambridge, Massachusetts: Blackwell, 1990.
4. S. Uryasev. Conditional Value-at-Risk: Optimization Algorithms and Applications. *Financial Engineering News*, **2**, 3, 2000.
5. C.Acerbi. Risk Aversion and Coherent Risk Measures: A Spectral Representation Theorem. Milano (Italy), *working paper*, July 2001, 11p.
6. J. v. Neumann, O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton: Princeton Univ. Press, 1953.
7. S. Wang. Premium Calculation by Transforming the Layer Premium Density. *ASTIN Bulletin*, **26** (1996), p. 71-92.

8. N. Boezio. Risk in Investment Accumulation Products: Insights from Investor Behavior. Proceedings of the Symposium "Risks in Investment Accumulation Products of Financial Institutions", Schaumburg, IL, 2001, p.67-77.
9. A. Muller. Stop-loss Order for Portfolios of Dependent Risks. *Insurance: Mathematics and Economics*, **21** (1997), p.219-223.
10. V. Young. Discussion of Christofides' Conjecture Regarding Wang's Premium Principle. *ASTIN Bulletin*, **29**, 2 (1999), p.191-195.

RESEARCHES IN IDENTIFICATION OF LOGICAL AND PROBABILISTIC RISK MODELS WITH GROUPS OF INCOMPATIBLE EVENTS

Solojentsev E.D., Rybakov A.V.
●
Institute of Mechanical Engineering Problems of RAS,
sol@sapr.ipme.ru

Abstract: In this paper the results of the researches in identification of the logical and probabilistic (LP) risk models with groups of incompatible events are presented. The dependence of the criterion function on several parameters has been investigated. The parameters include: the total number of optimisations, the amplitude of parameters increments, the initial value of the criterion function (CF), the choice of identical or different amplitudes of increments for different parameters, objects risks distribution. An effective technology of defining the global extreme in the identification of LP-risk model for the calculation time, appreciable to practice has been suggested.

Key words: risk, logic, probability, model, identification, incompatible events

The logical and probabilistic risk models are almost twice as accurate and have seven times better robustness than other known classification methods [1,2]. However the task of multi-parameter and multi-criteria optimisation for training LP-models is characterised by exclusive difficulty [1-3]. In the process of identification of LP-risk models in business according to statistical data there arise a number of additional features and difficulties [1,2]:

- The criterion function F_{max} (CF) is a number of correctly recognised good and bad objects, i.e. it accepts the integer values and it is stepped;
- CF has some local extrema, and depends on the high number of real positive arguments;
- The derivatives of the criterion function with respect to probabilities $P1_{jr}$ cannot be computed.

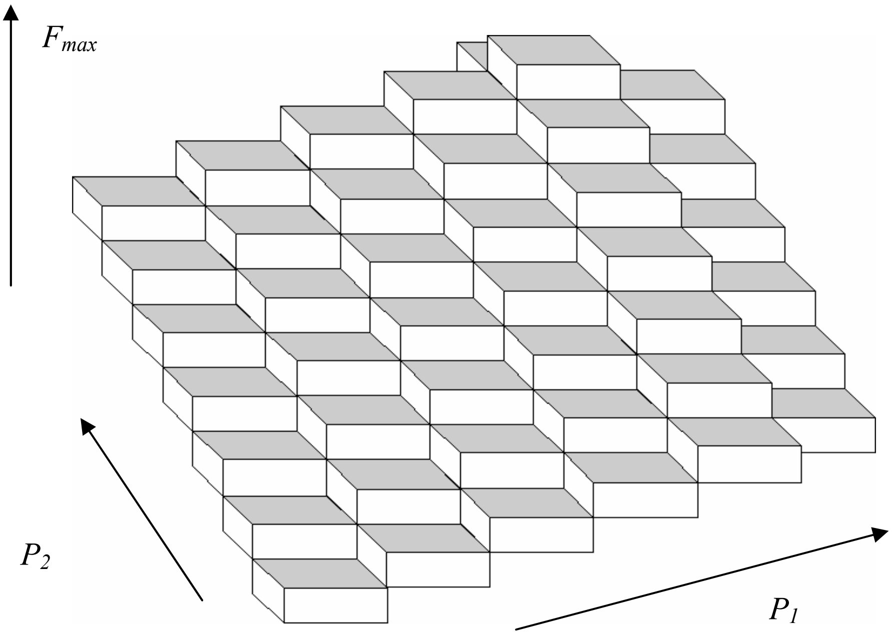


Fig.1. The stepped changing of the criterion function F_{max} from parameters P_1 and P_2

For each event-grade in GIE we consider three probabilities: W_{jr} is the relative frequency of the grade in the objects of the “object-signs” table, $P1_{jr}$ is the probability of the event-grade in GIE, P_{jr} is the probability of the event-grade to be substituted into the probability formula. The sums of the probabilities both W_{jr} and $P1_{jr}$ in GIE equal 1. Connection of these probabilities are considered in [1].

The criterion function F_{max}, presented in Fig.1, depends only on two arguments and changes with steps equal to 2. The platforms have different sizes. The arguments $P1_1$ and $P1_2$ belong to the interval [0,1], but their sizes can differ substantially. While approaching the extreme the platforms decrease in size.

The optimisation can get «stick» at any «platform», not reaching the maximum or crossing the maximum. The character of changing the criterion function in the multivariate space remains the same. Let us remind that the optimisation arguments space dimension for the credit risk LP-model equals 94 [1].

1. IDENTIFICATION OF LP-RISK MODELS

The risk object is described by a large number of signs, every sign has several grades. These signs and grades correspond to random events, which lead to a failure [1,2]. The events-signs (*j=1,n*) have logical connections and events-grades for each event-sign (*r=1,Nj)* form groups of incompatible events (GIE).

The identification of the P-risk model consists in the determination of optimal probabilities P_{jr} , $r = \overline{1, Nj}; j = \overline{1, n}$ *, corresponding to* events-grades. Let us formulate the identification (training) problem for a B- risk model [1,2].

Available data: the 'object-signs' table with N_g good and N_b bad objects and the risk B-model;

Expected results: to determine the probabilities of P_{jr} , $r = \overline{1, Nj}; j = \overline{1, n}$ for events-grades and the acceptable risk P_{ad}, dividing the objects into good and bad according the amount of risk.

***We need:* to maximise the criterion function, which is the number of correctly classified objects:**

(1) $$F = N_{bs} + N_{gs} \Rightarrow MAX,$$

where N_{gs} *and* N_{bs} *are the* numbers of objects classified as good and bad using both by statistics and the P- risk model (both estimates should coincide). From (1) it follows, that the errors or accuracy indicators of the P-risk model in the classification of good E_g and bad E_b objects and in the classification of the whole set E_m are equal:

(2) $$E_g = (N_g - N_{gs}) / N_g; E_b = (N_b - N_{bs}) / N_b; E_m = (N - F) / N.$$

Assumed restrictions:

1) probabilities P_{jr} *and* $P1_{jr}$ *must satisfy the stipulation*:

(3) $$0 < P_{jr} < 1, j = \overline{1, n}; r = \overline{1, Nj}.$$

2) the average risks of objects P_m *based* on the P- risk model and on the table P_{av} must be equal; while training the P- risk model we must correct the P_{jr} probabilities on every step of iterative training under the formula

(4) $$P_{jr} = P_{jr} * (P_{av} / P_m); j = \overline{1, n}; r = \overline{1, Nj}.$$

3) the acceptable risk P_{ad} must be determined with the given ratio of incorrectly classified good and bad objects, because of non-equivalence losses at their wrong classification:

(5) $$E_{gb} = (N_g - N_{gs}) / (N_b - N_{bs}).$$

2. OPTIMISATION IN THE IDENTIFICATION TASK

Identification of the LP- risk model by the random search method is based on the ideas used in the training of neural networks [4]. With reference to the identification task of the LP- risk model, the following formula for the calculation of the changes of events-grades probabilities may be put down:

(6) $$dP1_{jr} = K_1 * (1 / N_t) * tg(K_3); j = \overline{1, n}; r = \overline{1, Nj},$$

where: K_1 *is a* coefficient; N_t is the current number of optimisation; K_3 *is a* random number from [-π/2, + π/2], ***n*** is a number of events-signs, N_j is a number of events-grades in each GIE, i.a. for every event–sign.

In the formula (6) the CF is a current error in training. The number of optimisations N_t, before the end of the training process, can be very big. The «tangent» operation is the consequence of the training error distribution recording to Cauchy. Theoretically, this error is distributed according to the normal law, but not spend a lot of time on tabulated values calculation, we use the distribution of the training error under the Cauchy's law. It allows to reduce in 100 times the calculation time, which otherwise, for real problems, would continue for days and weeks.

For failure risk LP-model training the following modification of the formula (6) is suggested [1]:

$$dP1_{jr} = K_1 * (N_{opt} - N_t) * tg(K_3),\ j = \overline{1,n}; r = \overline{1,Nj}, \tag{7}$$

where: N_{opt} is the given number of optimisations. The new values of $P1_{jr}$ and P_{jr}, obtained at $F > F_{max}$ on every step N_t of optimisation are considered optimal and saved.

In the LP-risk model identification task, the criterion function cannot exceed the total number of objects in the statistical data. The formula (7) is quite applicable, but the time of calculation is too big (about 10 hours for a session of optimisation).

To reduce the time of calculation, in the formula (7) the "tangent" operation is eliminated. As a result the following expression is obtained [3]:

$$dP\,1_{jr} = K_1 * (N_{opt} - N_t) * K_3,\ j = \overline{1,n}; r = \overline{1,Nj}. \tag{8}$$

Using (7,8) the optimization happens so: if F>Fmax, then we remember the new $P1_{jr}$ and P_{jr}. If the criterion function does not strictly increase after the chosen number of trials N_{mc} in Monte-Karlo, then F_{max} is reduced by 2-4 units and optimisation continues.

In spite of the investigation in optimisation, carried out before, where the formulas (7) and (8) were used [1,2], the problem of optimisation in the identification task of LP-risk models is far from the final solution. The following fact proves it. In one of the research with the huge number of optimisations $N_{opt}=245\ 000$ and with the constant, almost optimal, value of the increment $dP1_{jr}$, we obtained $F_{max} = 824$ instead of $F_{max} = 810$ at the usual number of optimisations $N_{opt} \approx 245$. We had to carry out special investigations, the results of which are adduced below.

3. INVESTIGATIONS IN IDENTIFICATION / OPTIMISATION

If we generate a random number K_3 in the interval [-1, +1], then the absolute values of increments of probabilities $dP1_{jr}$, multiplied by 100, are transformed in percents (%). It is convenient, for practically it solves the problem of the evaluation of probabilities $P1_{jr}$ accuracy. For example, if the increment is $dP1_{jr}=0.0005$, it equals *0.05 %* . *W*e can say that the probability $P1_{jr}$ with the accuracy *0.05 %* is evaluated .

Using the formula (8), in the beginning of optimisation we have the following maximum amplitude of probabilities increments :

$$AP1_{max} = K_1 * N_{opt}\,. \tag{9}$$

In the end of optimisation the maximum amplitude of probabilities increments equals 0. Let us designate the current amplitude of probabilities increments as *AP1*. There is an optimal interval *OPT* of the amplitudes increments *AP1*, which position and width are unknown (Fig. 2). For the big values of *AP1* there is a small probability of increasing F_{max}, and for small values of *AP1* there is a high probability to stop at the local extreme of the reached value F_{max} (see Fig.1).

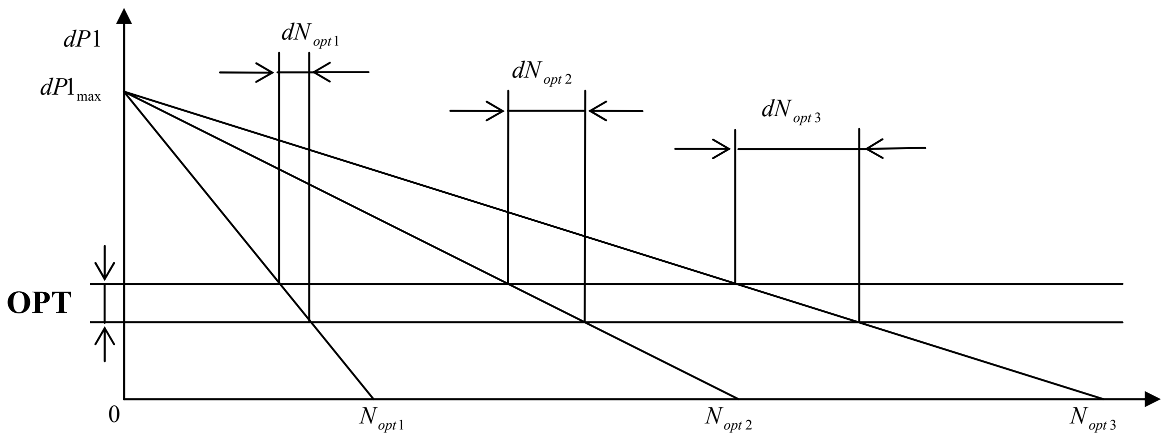


Fig.2. Graphs of relation between the number of optimisations N_{opt} and increments amplitudes AP1

The optimisation process (of training the LP-risk model) should be long enough in the optimal OPT interval . The value of dN_{opt} duration in the optimal *OPT* interval is equal

$$dN_{opt} = (OPT * N_{opt}) / AP1_{max} . \tag{10}$$

It also depends on the number of optimisations N_{opt} and the maximum amplitude of the increment $dP1_{max}$. The more N_{opt} is and the less $AP1_{max}$ is , the longer is the duration of dN_{opt}. The purpose of this work is the investigation of the dependence of the criterion function (accuracy of LP- risk model) on the following parameters in the training formula (8):

1. The number of optimisations N_{opt};
2. The increment minimum amplitude $AP1_{min}$, at which the optimisation is still possible;
3. The initial value of the criterion function F_{beg} ;
4. The choice of identical or different amplitudes *AP1* for different grades;
5. The increment maximum amplitude $AP1_{max}$;
6. Objects risk distribution in the statistical data.

Let us illustrate it. A question arises, whether to choose the identical or different values of increments amplitudes *AP1* for all events-grades ? In other words, whether the amplitudes $AP1_{jr}$ should depend on the values of probabilities $P1_{jr}$? In the training formulas of the LP-risk model (7) and (8) the increments amplitudes $AP1_{jr}$ are identical for all events-grades and do not depend on the values of their probabilities $P1_{jr}$. The increments $dP1_{jr}$ differ only because of the random simulation of the K_3 coefficient.

The model investigations for the LP-model of the credit risk were made on the PC. The credit risk structural LP-model has 20 events-signs (correspondingly GIE) and 94 events-grades. The credit risk L-function is [1,2] :

$$Y = X_1 \cup X_2 \cup \ldots \cup X_{20} \tag{11}$$

Verbally it can be formulated as follows: a failure occurs, if any one, or any two, … or all initiating events happen. After the orthogonalization of the L-function (11) the following P-risk model for the evaluation of the credit risk has been obtained:

$$P = P_1 + P_2 Q_1 + P_3 Q_1 Q_2 + \ldots . \tag{12}$$

The investigations were carried out in a set of 1000 credits, 700 of which were good and 300 - bad [5]. For calculation investigations we used the Software , designed in the object-oriented languages Visual C+++ and Java.

3.1 The choice of parameters N_{opt} , $AP1_{min}$, F_{beg}

In comparison with the optimal variant F_{max} = 824, the initial variant had the probabilities $P1_{jr}$ without the last four signs. So the optimisation starts at F_{beg} = 690-760. Such solution allowed to reduce calculation time.

The calculations were made for two values of increments maximum amplitudes: 1) $AP1_{max} =0.05$ *(5 %)* , 2) $AP1_{max} = 0.1$ *(10 %)* . We used the following numbers of optimisations N_{opt}*: 150, 300, 500, 750, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000.*

The results of investigations presented in Table 1 (Var.2-21) and Fig.3 , allow to make the following conclusions:

1) The criterion function F_{max} (column 6 in Table 1 and Fig.3) asymptotically increases with the growth of the number of N_{opt} optimisation ;
2) The minimum amplitude $AP1_{min}$ (column 9) equals approximately *0.0025 (0.25 %)*; at the smaller values of $AP1_{min}$ the optimisation does not happen and the number of the last optimisation N_{end} (column 10) is less, than the given number of N_{opt} optimisations. It is necessary to modify the law of the change of *AP1* during the training process *,* adding the constant line $AP1_{min}$ (Fig.4). It increases the chance to get the greater value of F_{max};
3) The big value of N_{opt} can lead to the disappearance of the B-C line (Fig. 4), which undoubtedly will deteriorate the process of optimisation.
4) The initial value of F_{beg} (column 5) should not be lowered, as it often leads to low final values of F_{max} (Fig. 5) because of the unsuccessful trajectory of optimisation process; in the considered case it is possible to accept F_{beg}*=750-760.*

Taking into consideration the just made conclusions , instead of the formula (8) the following formula for training the LP-risk model is suggested:

$$\text{(13)} \qquad \textit{If } AP1 < AP1_{min} \textit{, then } dP1_{jr} = AP1_{min} * K_3 \textit{,}$$
$$\textit{If } AP1 > AP1_{min} \textit{, then } dP1_{jr} = K_1 * (N_{opt} - N_t) * K_3.$$

The optimisation results using the formula (13) under $AP1_{min}$ *= 0.0025 (0.25 %),* different $AP1_{max}$ *= 0.098, 0.09, 0.03 (9.8 %, 9 %, 3 %), a* rather large number of optimisations N_{opt}*=5000-12000* and the high F_{beg} *=745* in Table 1 (var. 22-24) are shown. In all variants high values of F_{max} *=812-822* have been obtained.

Table 1. The investigations results in the choice of optimisation parameters

N	*Nopt*	*K_1*	*AP1ma*	*Fbeg*	*Fmax*	*dPc*	*AP1min*	*Nend*	*Notes*
1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
1	*2000*	*0.0001*	*0.2*	*776*	*786*	*0.204*	*0.1987*	*20*	
2	*300*	*0.000165*	*0.05*	*756*	*794*	*0.1969*	*0.00198*	*289*	*(3)*
3	*300*	*0.00033*	*0.1*	*712*	*790*	*0.221*	*0.00429*	*288*	*(3)*
4	*750*	*0.0000665*	*0.05*	*756*	*802*	*0.1641*	*0.00545*	*669*	*(3)*
5	*750*	*0.000133*	*0.1*	*692*	*790*	*0.2052*	*0.01316*	*652*	*(3)*
6	*1000*	*0.00005*	*0.05*	*750*	*802*	*0.1867*	*0.00350*	*931*	*(3)*
7	*1000*	*0.0001*	*0.1*	*708*	*792*	*0.2174*	*0.01580*	*843*	*(3)*
8	*2000*	*0.000025*	*0.05*	*776*	*808*	*0.1595*	*0.00747*	*1702*	*(3)*
9	*2000*	*0.00005*	*0.1*	*724*	*798*	*0.1802*	*0.01405*	*1720*	*(3)*
10	*3000*	*0.0000166*	*0.05*	*748*	*806*	*0.1867*	*0.00699*	*2581*	*(3)*
11	*3000*	*0.000033*	*0.1*	*708*	*806*	*0.1867*	*0.00501*	*2849*	*(3)*
12	*4000*	*0.0000125*	*0.05*	*744*	*812*	*0.1945*	*0.00791*	*3368*	*(3)*
13	*4000*	*0.000025*	*0.1*	*740*	*802*	*0.2121*	*0.00862*	*3656*	*(3)*
14	*5000*	*0.00001*	*0.05*	*754*	*806*	*0.1663*	*0.00556*	*4445*	*(3)*
15	*5000*	*0.00002*	*0.1*	*738*	*803*	*0.1586*	*0.00400*	*4801*	*(3)*
16	*6000*	*0.000016*	*0.1*	*710*	*810*	*0.1598*	*0.00625*	*5610*	*(3)*
17	*6000*	*0.0000183*	*0.109*	*736*	*810*	*0.1618*	*0.00495*	*5730*	*(3)*
18	*7000*	*0.0000071*	*0.05*	*764*	*810*	*0.2096*	*0.00407*	*6430*	*(3)*
19	*7000*	*0.0000142*	*0.1*	*734*	*810*	*0.1692*	*0.00745*	*6479*	*(3)*
20	*8000*	*0.0000062*	*0.05*	*764*	*810*	*0.1755*	*0.00985*	*6425*	*(3)*

N	*Nopt*	*K_1*	*AP1ma*	*Fbeg*	*Fmax*	*dPc*	*AP1min*	*Nend*	*Notes*
21	*8000*	*0.0000125*	*0.1*	*718*	*814*	*0.1802*	*0.00286*	*7772*	*(3)*
22	*12000*	*0.0000075*	*0.09*	*772*	*812*	*0.1737*	*0.0025*	*11754*	*(10)*
23	*8000*	*0.00000375*	*0.03*	*780*	*820*	*0.1526*	*0.0025*	*7662*	*(10)*
24	*8000*	*0.00000875*	*0.07*	*744*	*814*	*0.1733*	*0.0025*	*7801*	*(10)*
25	*5000*	*0.0000043*	*0.0215*	*812*	*820*	*0.1462*	*0.0025*	*23*	*(13)*
26	*5000*	*0.00000043*	*0.0025*	*810*	*824*	*0.1511*	*0.0025*	*34*	*(13)*
27	*8000*	*0.00000002*	*0.0025*	*810*	*826*	*0.1538*	*0.0025*	*678*	*(13)*
28	*8000*	*0.0000025*	*0.00458*	*806*	*822*	*0.1604*	*0.00609*	*507*	*(13)*
29	*8000*	*0.00000312*	*0.00572*	*806*	*822*	*0.1677*	*0.00452*	*1757*	*(13)*

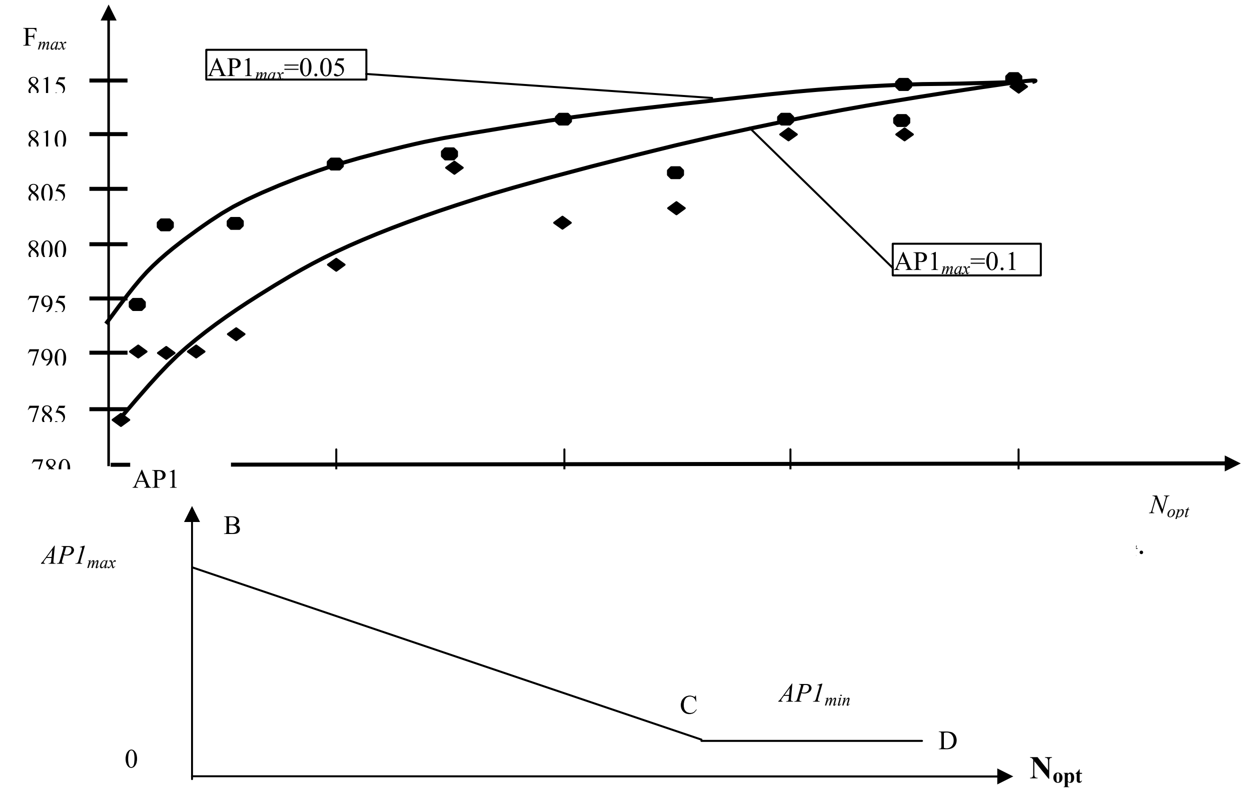


Fig.4. The graph of the current amplitude of increment AP1 modification

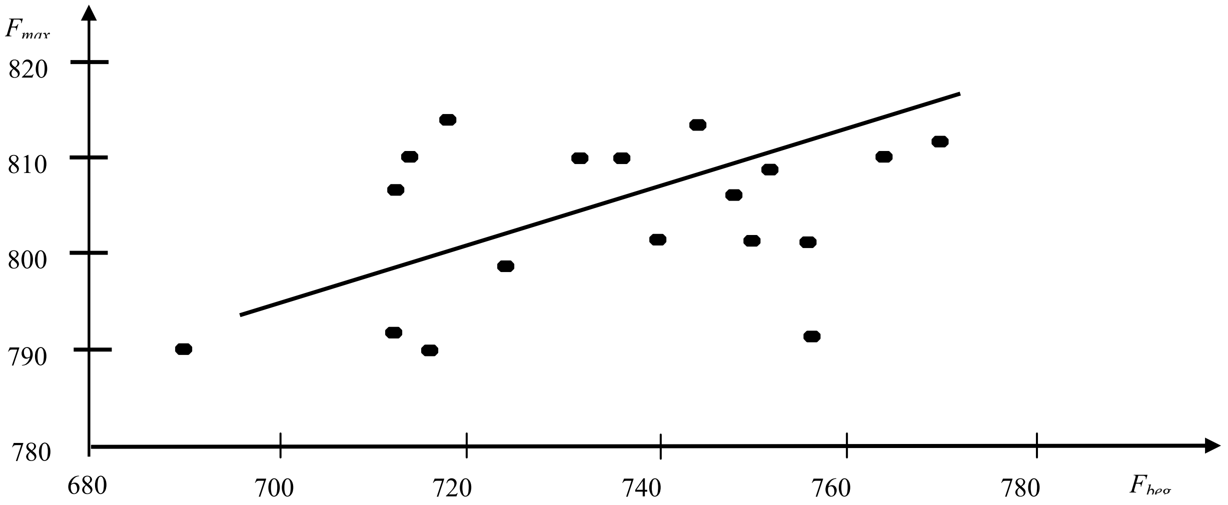


Fig. 5. Dependence of the criterion function F_{max} on its initial value

3.2 Different amplitudes $AP1_{jr}$ *of* increments for different grades

It should be noted, that the probabilities $P1_{jr}$ depend on: a number of grades in GIE, the frequencies of W_{jr} grades in objects and the grades contributions in the classification errors of objects. In the formula of training the LP-risk model (8) the increments amplitudes $AP1_{jr}$ are identical for all events-grades and do not depend on the magnitude of their probabilities $P1_{jr}$.

Let us change the formula of the training LP-risk model so that it takes into account the value of probability for each grade

(14) $$dP1_{jr} = K_1 * (N_{opt} - N_t) * K_3 * P1_{jr}..$$

Here the amplitudes for every event grade are equal

(15) $$AP1_{jr} = K_1 * (N_{opt} - N_t) * P1_{jr}$$

and the formula (14) can be the following:

(16) $$dP1_{jr} = AP1jr * K_3.$$

Let us also put down the formula (14) with the following modification:

(17) $$dP1_{jr} = K_1 * (N_{opt} - N_t) * ((1-a) + a * P1_{jr}) * K_3,$$

where *a* is a coefficient from the interval *[0 < a < 1]*. It determines the formula (8) at *a=0*, the formula (14) at *a=1* and all the modifications at other values of *a*.

In the formula (13) let us take into account the limitations, introduced earlier in the formula (8), and we shall get the following expression for training the LP-risk model:

(18) *If* $AP1_{jr} < AP1_{min}$, *then* $dP1_{jr} = AP1_{min}$,

If $AP1_{jr} > AP1_{min}$, *then* $dP1_{jr} = K_1 * (N_{opt} - N_t) * ((1-a) + a * P1_{jr}) * K_3$,

The investigations results in optimisation using the formula (18) at *a=1* ($AP1_{max}$ *= 2.15 % , 0.25 % , 0.45 % , 0.57%)* are represented in Table 1 (Var.25-29). They show that the high values of the F_{max} *=822-826* can be obtained at the limited number of optimisation attempts N_{end} (column 10). Actually the first optimisation already gives the high value of CF (F_{beg}*=806-810)*. The optimisation process ends at N_{end}= 23-1750 instead of the given numbers of optimisations N_{opt}*=5000-8000* (column 6). It seems, that the number of optimisations N_{opt} can be essentially reduced. To verify this hypothesis some extra investigations have been carried out.

The investigations were carried out at small numbers of optimisations N_{opt} *= 600, 450, 300, 150, 100, 50 and K1=0.00033, 0.00025, 0.00015, 0.0001*. The increments maximum amplitude $AP1_{max}$ varied in an interval *0.5% - 20%* from $P1_{jr}$. In Table 2 the CF values and the difference between maximum and minimum risks of objects in the statistics F_{max} */ APc* are shown. The results of the investigations should be considered as good *(*F_{max} *=810-822)* and completely confirming the effectiveness of the formulas (14), (17) and (18).

Also the investigations of the influence of *a* parameter on the optimisation results have been carried out. It was done at the small numbers of optimisations N_{opt}*=150* and K_1*=0.00015*. The maximum amplitude of an increment $AP1_{max}$ equals *0.0225**$P1_{jr}$.

Table 2. Values of F_{max} */ APc* at the small number of optimisations N_{opt} and *a=1*

Number of optimizations, N_{opt}	K_1=0.00033	K_1=0.00025	K_1=0.00015	K_1=0.0001
600	798 / 0.248	796 / 0.225	810 / 0.180	810 / 0.149
450	802 / 0.217	804 / 0.187	814 / 0.162	819 / 0.161
300	810 / 0.146	810 / 0.174	816 / 0.147	820 / 0.162
225	810 / 0.154	811 / 0.152	818 / 0.148	821 / 0.146
150	816 / 0.145	820 / 0.156	822 / 0.148	822 / 0.147
100	818 / 0.146	820 / 0.149	820 / 0.151	820 / 0.153
50	822 / 0.151	820 / 0.146	820 / 0.152	820 / 0.148

The investigations results, represented in Table 3, also confirm the effectiveness of the formulas (14),(17) and (18) at *a=1*. Really, at *a=1* F_{max} equals *820*, and at *a=0* F_{max} equals *802*.

Table 3. Values F_{max} at different values of a

Value a	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	1.0
Value F_{max}	802	800	798	804	808	810	808	810	818	820

3.3 Determination of the amplitude *AP1*$_{max}$ and the global extreme F_{max}

Let us consider again the choice of the increment maximum amplitude of probabilities $AP1_{max}$. The results of the change of F_{max} at the change of $AP1_{max}=K_1*N_{opt}$ in the interval *0.5-20 %* of $P1_{jr}$ are represented in Table 2. They demonstrate that the higher is $AP1_{max}$ the less is F_{max}. In Fig.6 the dynamics and the results of optimisation for five variants, having $N_{opt}=2000$, are shown:

- Variant 1: $AP1_{max}=0.05(5\%)$, $F_{max}=808$ (Var.8 in Table1), training under the formula (3);
- Variant 2: $AP1_{max}=0.1(10\%)$, $F_{max}=798$(Var.9 in Table1), training under the formula (3);
- Variant 3: $AP1_{max}=0.05\ (5\%)$, $F_{max}=820$, training under the formula (14) with a=1;
- Variant 4: $AP1_{max}=0.1\ (10\%)$, $F_{max}=804$, training under the formula (14) with a=1;
- Variant 5: $AP1_{max}=0.2\ (20\%)$,$F_{max}=786$(Var.1 in Table1),training under the formula (14).

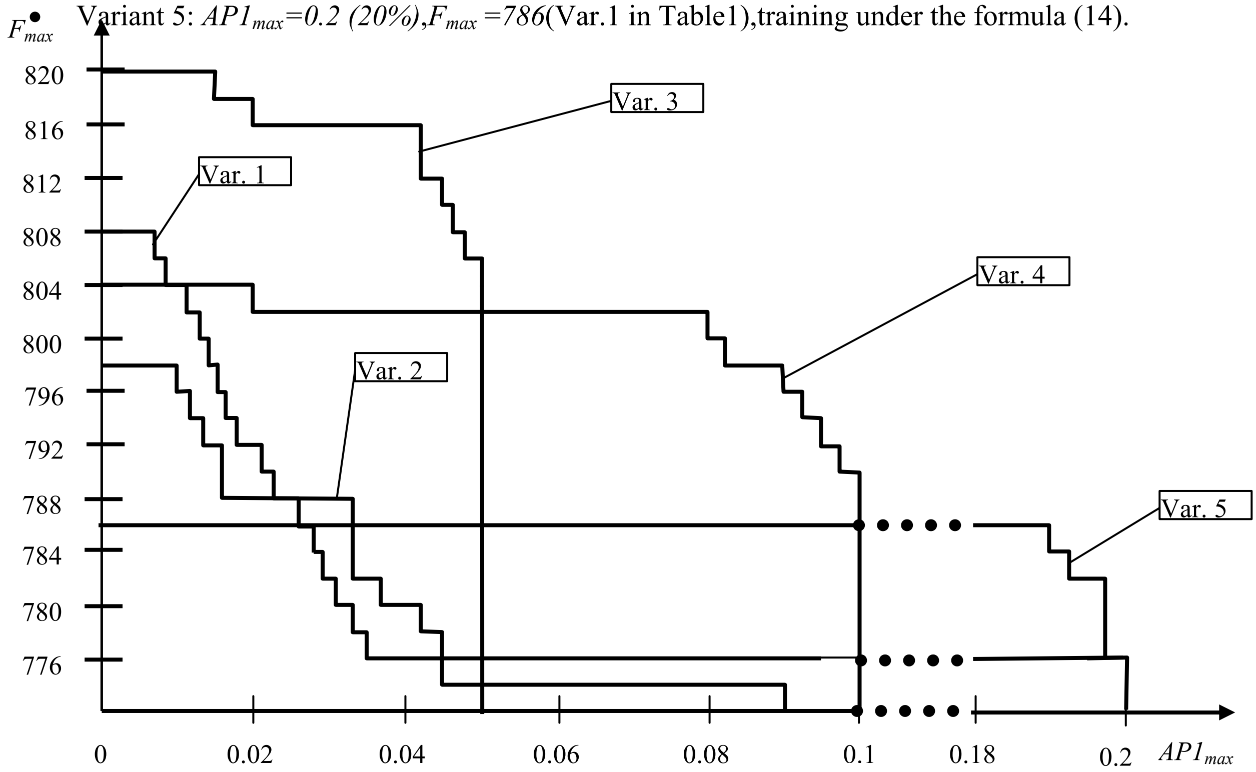


Fig.6. Dynamics and results of optimization depending on the AP1$_{max}$ *parameter*

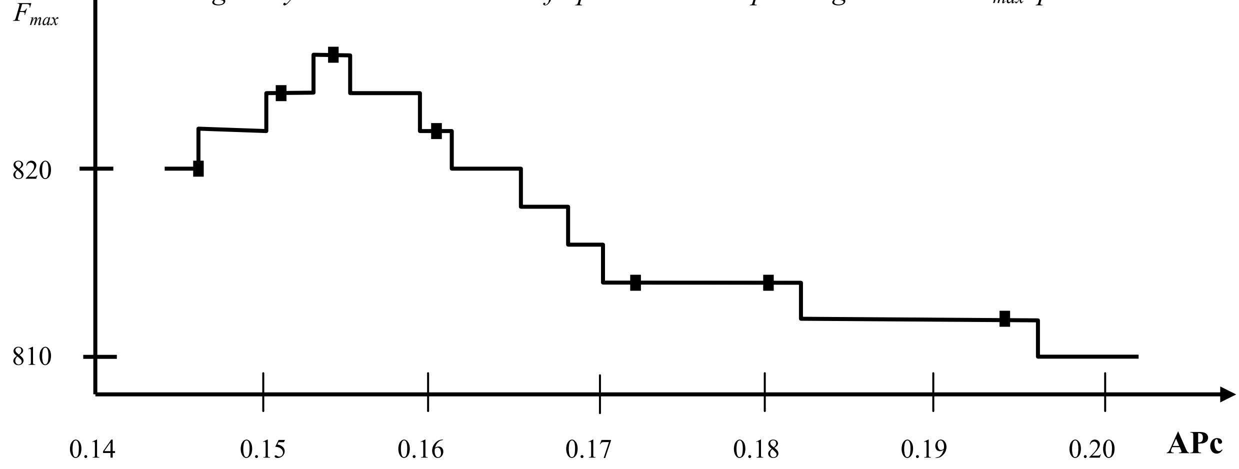


Fig.7.The connection of parameters F_{max} *and APc*

Variants 4 and 5 with high $AP1_{max}$, despite using the effective formula (18) and *a=1*, have bad training dynamics and results. In these variants CF are correspondingly *786* and *804*. The optimisation process finishes early, *(N_{end}=1608 and N_{end}=20)*. Additional optimisation attempts $N_{opt} - N_{end}$ have not increased CF. This example confirms that the increment amplitude $AP1_{max}$ should not be more than *0.02 - 0.05 (2-5 %)*.

We check the calculation of the global extreme of the CF by the graph (Fig.7). The function F_{max} has an extreme at some value of the difference *APc* between the maximum risk and the minimum risk of objects in statistics [2]. This difference, constructed for variants of computational investigations , presented in Table 1 and 2, demonstrates the robustness (stability) of solutions at a small dispersion of *APc* in the area of the global extreme of CF.

4. CONCLUSION

In the investigations the following main results have been obtained:

1. The effective technology of the criterion function global extreme search in the tasks of identification of LP-risk models under statistical data has been offered .It permits to solve the task of multi-parameter multi-criteria optimisation with integer CF for the time, applicable to practice (less than before).
2. We suggest to generate in the training formula a random number K_3 in the interval [-1, +1]. It permits to consider the absolute values of increments $dP1_{jr}$, multiplied by 100, in percents (%)) and to estimate the accuracy of probabilities $P1_{jr}$.
3. In the technology of the CF global extreme search, the following regularities of changing the CF should be used:
   - The CF asymptotically increases with the growth of N_{opt} optimisation number ;
   - The minimum amplitude $AP1_{min}$ of probabilities $P1_{jr}$ increments is established by 2-3 test calculations; at smaller values of $AP1_{min}$ the optimisation does not happen (less than 0.25 %);
   - The initial CF F_{beg} should not be lowered , as low values more often result in low final values of F_{max} because of the unsuccessful trajectory of the optimisation process;
   - Maximum amplitude of increments of $AP1_{max}$ must not exceed *0.02 - 0.05 (2-5%),* as the training speed lows down and the value of the CF F_{max} becomes less.
4. For the criterion function global extreme search new , more effective formulas of training (14), (17), (18) have been suggested ; they use different amplitudes of increments for probabilities of different events-grades.
5. It has been confirmed that we can test the determination of the global extreme of CF F_{max} by the graph of change of F_{max} in the function of difference *APc* between maximum and minimum risks of objects in statistics. The function F_{max} has an extreme at a certain value of *APc* .

REFERENCES

1. Solojentsev E.D., Karasev V.V.,Solojentsev V.E. Logic and probabilistic models of risk in banks, business and quality / edited by E.D.Solojentsev. SPb.: Nauka, 1999.-120 p.
2. Solojentsev E.D., Karassev V.V. (2001) Risk logic and probabilistic models in business and identification of risk models. - Informatica 25 (2001) 49-55.
3. Taha X. Introduce in research of operations . v.1,2. Moscow, Mir , 1985.
4. Wasserman Philips D. Neural Computing Theory and Practice. ANSA Research , Inc. VAN NOSTRARD REINOLD, New York, 1990.
5. Seitz J., Stickel E. Consumer Loan Analysis Using Neural Network. Adaptive Intelligent Systems. Proceed. of the Bankai Workshop, Brussels, 14-19 October, 1992. P.177-189

A NEW METHOD FOR THE APPLICATION OF RAMS TECHNIQUES TO QUALITY ASSURANCE (QA)

Renzo Righini
●
ENEA, Via Martiri di Monte Sole, 4 – Bologna – Italy
e-mail: renzo.righini@bologna.enea.it

Enrique Montiel
●
INESCOP, Poligono Industrial Campo Alto – Spain
e-mail:emontiel@inescop.es

Abstract

The application of RAMS techniques in all the phases of the lifecycle of each type of installation will surely guarantee its adequate exploitation in terms of production continuity and quality of the obtained products in the respect of prefixed constraints on the security of the working staff, safety and environment impact. In this frame, a particular importance must be attributed to the use of those techniques as support to quality assurance applied in the planning and building phases of the installation and of the products obtained by it. The present paper will include a short description of a method for the application of those techniques in this phase of the lifecycle and of the results that may be obtained by its application in shoes manufacturing, in particular those types where the technical requirements are higher, as it is the cases of certified products like “safety” footwear.

Key Words

Reliability, availability, quality assurance, risk assessment.

1. INTRODUCTION

Recently the discipline of Reliability, Availability, Maintainability and Safety (RAMS) techniques has changed from the traditional role of predicting reliability to that of involvement with the process of verification and validation of production lines and of their products.

By this way, the application of RAMS techniques in all the phases of the lifecycle of the assets and their products should contribute to obtain an amelioration of their quality (in terms of their safety, reliability, maintainability and availability) and a decrease of the production costs with a consequent improvement of their competitiveness.

In particular they should allow the achievement of the following main results:

- improvement of the quality, reliability and safety of the considered asset by the application of a suitable quality assurance program during all its design, planning and building phases;
- demonstration that the asset may be adequately exploited avoiding risks connected with working staff security, safety and environment impact;
- decrease of the production costs by an improvement of the quality of the production lines and of their connection and by a correct definition and improvement of the maintenance planning to be applied on them involving a more efficient use of existing capacity;
- make available advanced techniques allowing savings in the production by improving “lean manufacturing” and a “just-in-time adaptation” of the production lines to the quick changes of the market requests and trends.

2. AN INNOVATIVE SOFTWARE TOOL FOR THE APPLICATION OF RAMS TECHNIQUES ON INDUSTRY

In order to obtain the above results, ENEA (an important Italian Research Institute) is promoting, in collaboration with some industries, the development of a software tool for the application of RAMS techniques in normal industrial field besides the advanced ones (in fact, up to now, their use is limited to advanced fields such as nuclear, aerospace, etc. and for risk and safety analysis).

To this aim it will be constituted by the following main software methods and devices:

- a database for the recording of exploitation and test data;
- methods for the processing of reliability data on their basis;
- methods for the selection of the components on which maintenance must be applied and for its preliminary definition;
- methods for the system reliability analysis;
- methods for the application of RAMS techniques as support to quality assurance.

The database is a fundamental device in the application of all the software tool. In fact it must allow the achievement of important results such as a feedback of experience and the processing of meaningful and personalised reliability data.

Concerning the experience feedback that may be obtained by the enquiry of the data recorded in it, as better explained in the following, it has a fundamental importance in order to improve so the quality of the asset taken into consideration as its exploitation especially referring to the maintenance to be applied in such phase of the lifecycle. An adequate reliability and/or availability analysis cannot be left apart of the availability of meaningful reliability data on the components constituting the system taken into consideration; those data may be found in the literature, but this type of data doesn’t take into consideration the particular exploitation conditions to which the interested components are submitted with the consequence that they don’t allow the obtaining of meaningful results in particular analyses such as maintenance improvement or quality assurance. This type of data may be only obtained by the processing of suitable exploitation data previously collected on the components or component piece parts constituting the system taken into consideration.

To this aim a database component oriented will be developed. In order to obtain the results mentioned above it will be constituted by different files allowing the recording of the following information:

- univocal identification of the component and of its constituting parts not only in terms of their initials but also of their engineering characteristics, normative applied in their planning, building and testing, environment and operating conditions;
- recording of data on the results of possible tests carried-out on the component after its building;
- recording of information on the failures occurred on each component during its exploitation;
- recording of information on each maintenance (preventive programmed maintenance or failure repair) carried-out on the components so identified.

The database will be interactively connected with suitable computer programs for the data entry and for the enquiry of the recorded data. Concerning the data entry, the set-up program will include the possibility so of a transfer of data from other databases as the manual recording by the operators who collect the information. The enquiry program will be developed so that the end-user may obtain datasets to be used for the reliability data processing or for a direct exploitation feedback on the basis of his necessities: details on the subjects will be reported in the following.

In order to obtain meaningful reliability data starting from the rough data recorded in the database, suitable processing methods will be developed. In particular methods based on the classic so as on the bayesiam statistics will be set-up so that the processing will be possible also in the case that very few rough exploitation data are available. In this frame methods approximating the main statistics functions and allowing the processing of each type of reliability data (failure rates, repair times, etc.) will be developed. To this aim suitable homogeneous datasets obtained on a particular component type by the enquiry of the information recorded in the database (see above) will be used. At last further methods for the statistical analysis of imprecisely recorded data will be also set-up: they will meet the need arising from the practical observation that real field data very often cannot be precisely recorded; to this aim an original methodology linking mathematical statistics with the theory of fuzzy sets will be developed.

As better explained in the following, the reliability data so obtained will be used for different applications of RAMS techniques such as maintenance definition and improvement, support to quality assurance, risk assessment and system reliability analysis, etc.

The methods for the system reliability analysis included in the main software tool must allow the achievement of the following main results:

- demonstrate, in the design and planning phase, that the product will surely satisfy requested quality and reliability requirements;
- execution of a suitable safety analysis and risk assessment in order to demonstrate that the considered system (installation, equipment, etc.) will surely satisfy the requested requirements in terms of security of the operating staff, safety, environment impact, etc;
- improvement of the maintenance planning at system level by the research of the best maintenance configuration corresponding to the minimum management costs and maximum availability.

In order to obtain those results, two different methods, respectively based on Monte Carlo or analytical techniques have been developed. The first type of techniques allow a realistic simulation of the system also in presence of failure rates varying in the time and a consequent realistic evaluation of quantities such as availability or management costs, etc.: it follows that the code developed on their basis will be the best solution to face problems connected with the maintenance improvement. On the contrary the use of analytical techniques will allow a best evaluation of reliability: it follows that a code based on the use of such a type of techniques will constitute a valid support for the execution of safety and risk assessment or of other analyses in which the evaluation of reliability is a critical factor.

Maintenance is a critical factor in the management of each type of industrial system in consideration of its influence on availability, safety and management costs. The introduction of reliability criteria in the maintenance definition and improvement may surely contribute to the amelioration of those quantities. To this aim, the software tool in reference will include a suitable method allowing the application of innovative criteria such as the selection of the only components on which maintenance must be applied on the basis of their importance in the system in which they operate and of the time trend of their total failure rate or the preliminary definition of the maintenance strategies to be applied on the components so selected on the basis of such time trend of the failure rate: by this way many useless and repeated maintenance operations will be eliminated with positive effects on the system availability and management costs. Besides, the maintenance strategies so defined component by component will be improved at system level by the methods for the system reliability analysis in order to find the best maintenance configuration corresponding to the minimum of the management costs and maximum availability in the respect of prefixed constraints on safety, environment impact, etc

In fact maintenance must be carried-out on the components constituting a system in order to prevent the intervention of systematic failure causes which would surely generate the failure of the component (and the consequent increase of their failure rate) in a short time. On this basis it is a non-sense to apply preventive maintenance on components only having a decreasing and/or constant failure rate (their failure probability would not change after the maintenance execution). On the contrary maintenance must be surely applied on components having an increasing trend of the failure rate in their life in order to eliminate, before their intervention, the systematic failure causes generating such increase. Equally maintenance must not be carried-out on components not important in the system operation in order to avoid useless money consumption (in this case the waiting of the component failure and its repair is convenient). It follows that the code that will be developed will allow the automatic analysis of the time trend of the total failure rate of each component (previously processed by the methods described above on the basis of the exploitation data recorded in the database) and of its importance: on this basis the code will allow an automatic selection of the only components having an increasing total failure rate and important in the system operation as the only ones on which maintenance must be applied. In particular a suspension of the maintenance provided on components having only constant failure rate will be advised in order to wait for the time at which it becomes increasing because of the intervention of systematic failure causes. Once the components on which maintenance must be applied have been so selected, the developed method will allow the definition of the maintenance strategies to be applied on them on the basis of the operations necessary to eliminate, before their intervention, the systematic failure causes generating the increasing of the total failure rate and of the time trend of the correspondent failure rates: the time at which the increasing failure rate calculated for each systematic failure cause crosses the constant one corresponding to random failures is advised to apply the preventive maintenance operation necessary to eliminate such systematic failure cause. Besides, the code will carry-out a comparison between the costs of each maintenance and of the failures to be avoided by it: such a comparison will allow to verify if the execution of maintenance is convenient in comparison with the waiting of the failure.

In order to obtain the results described above, the data recorded in the database will be adequately enquired so obtaining a suitable experience feedback and the failure rates calculated for each failure cause (random causes included) will be exploited. By this way an innovative FMECA (Failure Mode Effect Cause Analysis) method will be developed.

At last the main software tool will include the method described in details in the following chapter allowing the application of RAMS techniques as support to the quality assurance applied on the assets and on their products during their planning, building and first exploitation phases.

3. THE METHOD FOR THE APPLICATION OF RAMS TECHNIQUES TO QUALITY ASSURANCE

In the following figure 1 the main phases of the planning, building and first exploitation of each type of item are reported together with their possible interactions with RAMS.

Usually ISO/CEN standards are applied to draw the specifications of the product to be obtained; besides, CAD/CAM tools are followed to plan and manufacture it. In those phases internationally tested norms and standards are strictly applied so that a prefixed quality of the product must be guaranteed.

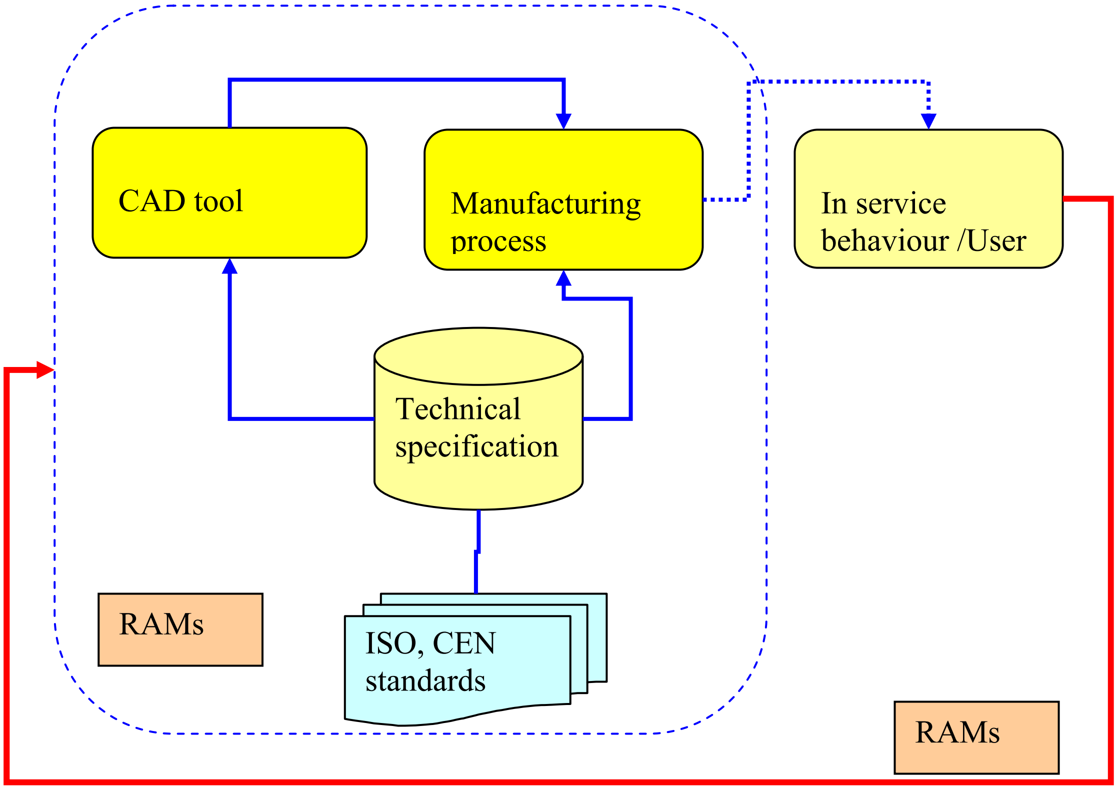


Figure 1. Areas of possible application of RAMs techniques

The method to be developed must allow the integration of RAMS techniques with those CAD/CAM systems: by its systematic application an amelioration of the design of the products, also including an adequate definition of trials, should be obtained. The final goal is the integration of RAMS techniques within design stages of products which are required to satisfy any technical requisites related to quality, safety, etc. Modules, to be considered in the proposed software are: materials requirements, performance requirements, design parameters, link with QA system.

In fact an adequate planning and building of each type of product cannot be left apart of statistics and reliability techniques.

First of all the design of a product must be "robust", where "robust design" means the obtaining of optimal performance of the product simultaneously with variation in manufacturing and field conditions. To this aim requested performance requirements of the product must be singled-out together with the factors which allow the achievement of those performances and, consequently, characterise them; the designers should specify target values of those factors and their possible tolerances so that the performance of the product in the field is not affected by variability in manufacturing or field conditions: suitable demonstrations must be carried-out to this aim. Besides, tests must be defined allowing the verification that the product has been realised respecting the nominal values of those factors with the consequence that the requested performances have been achieved. The execution of the above demonstrations, the interpretation of those tests and the verification that requested performances have been achieved necessarily involve the use of statistics techniques.

In any case, the verification that the requested performance requirements have been achieved does not guarantee that the product will have a long service life with few failures: traditional efforts of design, although necessary for the reasons expressed above, are often not sufficient to achieve both the functional performance requirements and a low rate of failures with time. To prevent these failures (i. e., to achieve a high reliability of the product), a specific "reliability program" must be applied during all the life-cycle of the product with particular reference to its planning and manufacturing. Such reliability program typically includes the following activities: setting overall reliability goals, apportionment of the reliability goals, stress analysis, identification of critical parts, failure mode and effect analysis, reliability prediction, design review, selection of suppliers, control of reliability during manufacturing, reliability testing, failure reporting and corrective action system.

On the basis of what reported above, referring to the design and manufacturing phases (see the part on the left of the fig. 1), RAMS techniques should allow the achievement of the following main results: first, an "a priori" verification that the target values of the factors characterising the performances of the product and the tolerance range around such a target value allow the achievement of those performances; besides, activities such as a setting of overall reliability goals, an identification of critical parts, a failure mode and effect analysis and a reliability prediction (consisting of an evaluation of the reliability of the product so demonstrating that prefixed requirements of the item to be produced will be guaranteed) should be carried-out in such a phase. To this aim the different methods and devices constituting the main software tool and described in the previous paragraph will be exploited: concerning the first results summarised above, the statistics methods will be applied; reliability data found in literature (or otherwise) on components similar to the ones constituting the item in reference, and the system reliability analysis codes will be used to "predict" the reliability and availability of the produced item. If safety requirements must be respected in its utilisation, a suitable safety analysis and risk assessment will be also carried-out by those codes during the planning phase.

Once the item has been manufactured, suitable tests will be carried-out on it to verify if the prefixed quality and reliability requirements have been achieved. Suitable records will be provided in the database to store the results of those tests. Besides, the statistics methods summarised in the previous paragraph will be applied to interpret the results of those tests. Tests are usually carried-out on a sample of products of the same type in order to achieve the following main types of results:

- control that prefixed quantities characterising the product are comprised within prefixed limits and that the results obtained may be extrapolated to the population that will be brought in the market;
- demonstrate that prefixed performance requirements may be achieved if the measured values of the factors characterising such performances are included within prefixed tolerance ranges;
- single-out possible failures due to "infant mortality";
- evaluate the reliability of the product on the basis of the failures occurred during the tests.

While good results may be obtained by the tests in the carrying-out of the above controls and in the singling-out of defects generating infant mortality, only a preliminary evaluation of the product reliability may be obtained by them.

On the contrary, more realistic results in the field of reliability evaluation may be obtained by a systematic recording of the results obtained during the exploitation or use of the product (see part on the right of fig.1) and by their interpretation on the basis of a systematic application of the above statistics methods (and system reliability analysis codes if the item is a complex system).

First of all, only the exploitation of the products allows a complete detection of failures usually known as infant mortality and due to defects generated by errors in the planning, inadequacy of the materials or of the techniques applied in the manufacturing: in fact, only partial information in the field may be obtained by the tests mentioned above. As better explained in the following, from a deep analysis of those failures, important suggestions may be obtained on modifications to be introduced into the various phases of the production in order to improve the quality and reliability of the obtained product. In such a frame, a fundamental importance is assumed by the evaluation of the time variation of the decreasing failure rate in the first phase of the exploitation (infant mortality) and by a qualified recording of the exploitation data into the database so that meaningful reliability data may be processed on their basis. The analysis of those failure rates and of the data recorded in the database (beginning from the information recorded on the techniques and materials used and applied in the manufacturing of the product up to the causes and modes of the failures) should allow the obtaining of suggestions on modifications to be introduced into the materials and into the design parameters in order to improve such a product. Obviously a first information will be obtained by its first exploitation but an iterative application of the methodology is advised: to this aim new data should be recorded (and successively processed and analysed) on the exploitation of the new product obtained after the introduction of those modifications so that the amelioration obtained by their introduction will be quantified; in this frame the different values of the failure rates obtained at each new application must be compared. The results so progressively obtained will also suggest important modifications to be introduced into the testing trials to be applied on the product so that it may be improved before its lunch into the market. In fact the presence of only a defect in few items of a stock involves the restitution of all the stock with consequent high loss of money.

Besides, the processing of the exploitation data progressively recorded in the database allows the obtaining of realistic reliability data (failure rates, middle time to failure, etc.) on the product obtained and a consequent realistic evaluation of its reliability: to this aim the methods for the reliability data processing will be exploited together with the system reliability analysis codes if the product is complex and constituted by more components.

A further method will be developed which will allow an automatic use of the methods and devices constituting the main software tool in order to obtain the results summarised above in the frame of the application of RAMS techniques as support to quality assurance.

In particular such development will involve:

- the verification that the available software methods and devices (database, statistics methods, methods for the system reliability analysis) constituting the main methodology and described in the previous chapter are enough for the achievement of the provided results. If they are not, singling-out of further methods that must be developed to this aim and their consequent development;
- an arrangement of the files constituting the database so that it will allow an adequate recording of the data necessary for the execution of the analyses described above such as tests interpretation, identification of the modifications to be introduced into planning and building of the products in order to limit or avoid their infant mortality, and so on;
- an adequate interface of the existing methods listed above so that their quick and end-user friendly application will be possible in order to allow the application of RAMS techniques to quality assurance in the terms summarised above;
- an exact definition of the analyses to be carried-out for such application of RAMS techniques to quality assurance and of the procedures to be followed in order to realise the results of those analyses (modifications to be introduced into the planning and manufacturing of the product and into the constituting materials, etc.). In this frame an integration of those techniques with the project systems adopted (CAD/CAM, etc.) will be also developed, providing designers of new products a valuable tool which could be considered as a virtual tester.

4. APPLICATION EXAMPLE ON SHOES MANUFACTURING INDUSTRY

Once the development of the above method will be completed in the terms described in the previous chapter, it will be tested by its application on existing industrial systems.

In this frame it will be also applied on safety shoes manufacturing. As said, the sw tool described in the present paper is yet in a development phase: it follows the presentation of results on the subject is not yet possible. In any case it is interesting to analyse in this paper the results that should be obtained by the application of the method subject of the previous chapter also because the manufactory in reference will directly take part in its development.

The final goal to be obtained should be the integration of RAMS techniques with CAD/CAM systems applied for shoes and components design in all the stages of such design with particular reference to safety shoes (i. e. a particular type of shoes used by workers and subject to European Directive of IPE - Individual Protective Equipment): by a systematic application of the method an amelioration of the design of the products also including an adequate definition of trials to be carried-out to this aim should be obtained. The importance of the use of such a technique relies in the fact that the legal responsibility of any accident/injury to workers is automatically affecting any providing company (footwear maker) in certain cases: so the sale of a pair of safety shoes is just the beginning of a process of accepting responsibilities for which no tools are available in the market.

The footwear design process, usually carried out by design specialists, is a complex process of several stages which requires detailed knowledge of footwear, both the shoe and the production, and the current fashion tendencies. Footwear design understands, apart from developing the design for a shoe, the development of the engineering drawings and the selection of the most appropriate last and the definition of technical specifications of the final product. The short lifecycle of a product such as a safety shoe (less than 1 year in-service) and its character of "passive" element of relative low cost and mass produced, makes the application of sophisticated techniques such as RAMS quite complex and expansive today. Even more, after sales are done, no relevant information of the product feeds-back the manufacturer, once the shoes are "in use". Anyway, in future markets where the "service and functionality" aspects of consumer goods (shoes) which have to fulfil very specific technical requirements during their use by consumers, RAMS or an evolution of them could support the product control along its lifecycle and a clear added value for the end users.

RAMS concepts are intended to be incorporated inside or close to CAD tools, putting intelligence into a product which many times is only checked throughout trial-error processes.

Referring to the diagram of figure 1, an area where actually lifecycle concepts are not really used is the use of the product by the consumer (in-service period, see the part on the right of the diagram), where no relevant information goes back to the manufacturer, except specific market complaints due to failure of the shoe. In any case, the presence of few defects rising in the first period of the life of a product involves the restitution of all the stock in which those few failed shoes were included. Independently of the money losses due to such restitution, the presence of those defects has a very negative effect on the image of the producer and on the competitiveness of his products. It follows that particular importance is assumed by a continuous contact between the producer and the customers so that the first one may take into account information on defects and failures found by the second ones, analyse them and introduce suitable modifications (in the planning, materials, etc.) and consequently improve his products. In this frame the RAMS techniques, applied as described in the previous chapter (see the database and the methods for the reliability data processing), could add value to shoes in generals and in particular to safety shoes, during the design-manufacturing stage (singling-out of the causes of the defects, consequent intervention into the materials, manufacturing processes, etc. in order to avoid those defects and to improve the quality of the product).

Problems caused by inadequate footwear are a significant cause of workplace absenteeism in companies. The cause-effect relationship is clear, but it is neither direct nor quantifiable. Although the prevention of foot problems and the use of the right footwear is fundamental, there is only one standard (UNE 23 087) that deals with safety footwear through impact studies. A few studies, like the "Safety footwear against mechanical risks: Guides for the election, use and maintenance" by the Spanish Workplace Health and Hygiene Institute (INSHT,) condense scattered information on this type of footwear.

An additional element which justifies the application of RAMS, is the legal responsibility of the safety shoe manufacturer in case of worker accident, as safety shoe is a protective element which has to fulfil the functionality for which it was designed and manufactured. In this frame RAMS techniques may be used to carry-out suitable safety studies: starting from the singling-out of the failures with consequences to the user security that may occur, they should allow the singling-out of the protections or planning devices necessary to avoid them. Besides, the failure rates that may be processed on the basis of those occurred failures will allow an "a priori" evaluation of the reliability of the product obtained so contributing to its qualification.

At last RAMS techniques will also be applied in the terms described in the previous paragraph in order to improve the design parameters and their tolerances so that the provided performance requirements are achieved.

5. CONCLUSIONS

The main features of a method for the systematic use of RAMS techniques as support to quality assurance have been described in the previous chapters: it may be applied to each type of product contributing to the improvement of its quality and reliability. In particular, the results that may be obtained by its application on shoes manufacturing (with particular reference to "safety shoes") demonstrate how the method may contribute not only to improve the quality of the product obtained but also to avoid its failures and the consequent risks; to this aim a fundamental importance is assumed by the application of RAMS techniques during all the planning and building phases so contributing to the obtaining of a robust design and, besides, in the execution of suitable safety and risk analysis also carried-out on the basis of the exploitation feedback supplied by the end-users. Obviously such logic may be applied on each type of asset, beginning from the products of a manufacturing industry up to each type of plant, installation, etc.

BIBLIOGRAPHY

1. European Directive 96/686 for Individual Protective Equipment.
2. Montiel E., (2004). Calzado personalizado: una oportunidad para competir. Congreso Calzatecnia. Mexico.
3. Montiel E. RTT-Footwear: en Experience in Technology Transfer. César Orgilés, Edited by INESCOP.
4. Righini, R., Fichera, C., (2000). A Monte Carlo Software Method for the Maintenance Improvement, Proceedings of ESREL 2000 Conference, Edinburgh, UK
5. Righini, R., Bottazzi, A., (1999). New Reliability Data Processing Methods for Maintenance Improvement, Proceedings of ESREL 1999 Conference, Munich, Germany
6. Righini, R., (1998). A Database for Reliability Centered Maintenance Management, Proceedings of PSAM4 Conference, New York City, USA

ASPECT-ORIENTED SOFTWARE RELIABILITY ENGINEERING

Igor Safonov
●
International Unity Science Institute
1011 Arlington Blvd., Suite 403
Arlington, VA 22209, USA
isafonov@aol.com

ABSTRACT
Aspect-Oriented Approach to Software Development allows us effectively to effectively extract, evaluate and solve the main problem of contemporary tendency in Information Technology (particularly, in an Application Software) – a unification is alternated by a personalization. Increasing customer concerns about Performance, Quality, Reliability and Security (PQRS concept) can be satisfied only by symbiosis synergy of adequate models, techniques and tools on all stages of the Software lifecycle. We propose original methodology, formal models and simple methods of Software Reliability Engineering based on our many years experience of concern separation and aspect orientation in Software Development for Specialized Computers, Business Application and Government Institutions.

Keywords: Aspect-Oriented Software Development, Software Reliability Engineering, Application, Optimization, Model Driven Architecture, Formal Foundations, Structure-Algorithmic System

1. INTRODUCTION

By the late 20th century, two main tendencies of Software Engineering (SE) – extensification and unification were alternated by intensification and personalization. The changing reflected on formulation and decision of optimization problems. Multi-years experience in SE make us strong supporters of the Aspect-Oriented Software Development [12, 13] and allowed us to create and use the Customer-Oriented Approach (COA) to SE for Reliability (Performance, Quality, Security, Unification, etc.) [19]. The COA combines ideas of several directions in SE, requiring global (External-Internal, Behavior-Structure) and local (Property-Oriented) separation of concerns (Architecture Alternatives [1], Goal Patterns [2], Process Algebra [3, 11], Formal Languages [6, 11], Workflow [4, 14], Exception Handling [17, 21]), Application-Oriented Operating Systems [10], and Evolutionary Games [8, 9]). In this article, we limit our Software Reliability Engineering by the example of Checkpointing-Recovery System [5, 15, 16, 18, 20] widely accepted by developers.

The Ontology of Software Reliability Engineering specifies External (Functions) and Internal (Aspects) any Software Application at different stages of Development by the quadruple of correspondent Structure, Behavior, Goal and Resource models and their decomposition (top-down approach). The Ontology uses unified Extended Algorithm Algebra Language (XAAL) for separation of Internal Behavior and Structure concerns in process of Analysis, Optimization and Synthesis with required or possible degree of detailed elaboration for every stage of Software lifecycle (starting from Customer Requirements and Formal Specification, and finishing with Deployment and Maintenance). The simple and fast optimization techniques make the adaptation of Software Internal Behavior possible in order to compensate restrictions of universal operating systems and personalize application properties.

Aspect-Oriented Approach to Software Reliability Engineering was implemented for dozens of Aircraft, Spacecraft and Submarines Navigational Computers, for the Tax Modernization System, for Year 2K Problem, for Search Engines, and for a Government (including the Personal DSS for a Country President). In this time, we are designing Safety Aspect-Oriented Software of the Emergency Loss Prevention System (ELPS) for local governments with severe Requirements and Restrictions to Performance, Reliability, and Security. The ELPS is based on the Safety Wireless Network for Incident Management.

The proposed Software Model can specify any complex or simple software system being viewed as an object of design, at different stages of design, with the needed level of details Mi (t), i = 1, 2, … , K, in which the structure and behavior of the system, the goals and resources of its design are represented, i. e.

$$Mi\ (t) = < MSi\ (t),\ MBi\ (t),\ MGi\ (t),\ MRi\ (t) >, \qquad i = 1, 2, \dots , K , \qquad (1.1)$$

where MSi (t) – Structure Model, MBi (t) – Behavior Model, MGi (t) – Goal Model, MRi (t) – Resource

Model, and t – time.

In process of design by stages, we accomplish that level of detailing, which allow us to develop the documentation needed for design and subsequent production of designed system. If MBi (t) , i = 1, 2, … , K , is a program, all functional and logical operators of which can be structurally interpreted in terms of models MSj (t) , i <= j <= K , then Mi (t) , i = 1, 2, … , K, will be called the Structure-Algorithmic System (STALS). The process of STALS development is the multistage sequential (in most cases, labor-intensive and time-consuming) approximation to the required (needed and sufficient) for implementation level of details. For example, in the design process of computer-aided control systems for technological processes the stages of feasibility study, technical design and functional coding are executed, in the special computer design – the stages of system, algorithmic, logic-functional and technical development, in the applied software package design – the stages of requirements, specification, design, programming, and testing.

2. SOFTWARE RELIABILITY SYNTHESIS

By the Software Reliability Synthesis Model of STALS we shall name the formal specification of the software for realization of the designed system's external (functions) and internal (aspects) behavior containing information sufficient for the functionally equivalent additions to and transformations of the specification with the purpose to control reliability parameters of the system, i. e. sufficient for Formal Software Reliability Synthesis. As Reliability Synthesis Model of Structure-Algorithmic System, we shall use the submodel < MRSS, MRSB > of the model < MS, MB >, where MRSS – the Model for Reliability Synthesis of Structure, MRSS is a part of MS, and MRSB – the Model for Reliability Synthesis of Behavior, MRSB is a part of MB.

Because, eventually the goal of Reliability Synthesis of control components and systems consists of giving the detecting and correcting features to the object functioning in real-time scale, we therefore can execute Reliability Synthesis to fit the MRSB, making the corresponding additions to and transformations of MRSS. In doing so, the working hypothesis is the statement that the level of abstraction can be chosen in such way that to every operator (logical condition) we can set in accordance with the structure component realizing this operator (logical condition).

We shall add to and (or) transform of MRSB by governing relationships of the form:

$$A1 = A2, \tag{2.1}$$

where A1 – an arbitrary Reliability Design Component (any operator or logical condition out of MRSB, any syntactically correct, i. e. canonically represented, part of MRSB, or the complete MRSB), and A2 – an algorithm functionally equivalent to A1, except that A1 and A2 vary in their detecting and correcting properties, or in probabilities of correct execution, or in levels of other particular parameter of the quality or cost.

Let P (A1) $\neq$ P (A2), where P (A1) – a reliability parameter (in this case – the probability of correct execution) of A1, and P (A2) – a similar reliability parameter of A2. In the majority of practically interesting cases of Reliability Synthesis P (A1) ‹ P (A2), and the alternatives when P (A1) › P (A2), can result from the Design for Reliability of objects with natural redundancy (for example, in process of Design for Reliability of the STALS, realizing on the basis of heterogeneous or homogeneous computer networks [5, 14, 16]). Other quality parameters (timeliness, accuracy, etc.) and a cost of A1 and A2 can be different, but sometime can be the same. By the Formal Reliability Synthesis of programs, we shall name the sequence (may be one-component) functional equivalent additions to and (or) transformations of MRSB by the governing relationships of the form (2.1).

For initial illustration of XAAL, a simple example may be useful. The exception handling pseudo code is from [17]:

A = a3(a2(A1*a1(A2 V E) V A2) V A3*a4(a2(A1*a1(A2 V E) V A2) V E)),

where A1 ::= call a local handler; A2 ::= go to a higher (action) level handling;
A3 ::= local error detection; a1 = 1, if "not handled", else a1 = 0
a2 = 1, if "local handler exist", else a2 = 0
a3 = 1, if "exception is propagated from the component", else a3 = 0

a4 = 1, if "exception is raised (error has been found)", else a4 = 0
or

$$A = a3(A4 \vee A3{*}a4(A4 \vee E)),$$

where A3 ::= local error detection;
a3 = 1, if "exception is propagated from the component", else a3 = 0
a4 = 1, if "exception is raised (error has been found)", else a4 = 0

$$A4 ::= a2(A1{*}a1(A2 \vee E) \vee A2),$$

where A1 ::= call a local handler; A2 ::= go to a higher (action) level handling;
a1 = 1, if "not handled", else a1 = 0
a2 = 1, if "local handler exists", else a2 = 0

The methodology of Formal Reliability Synthesis of Algorithms can be illustrated by an application example of the very common governing relationship:

$$A = \nu\{ K \}\cdot A \cdot\omega\{ \pi(E \vee K \cdot\nu\{ K \})\cdot A \} , \qquad (2.2)$$

where A – an arbitrary Component of the Design for Reliability, K – a restoration operator,
ν – a logical condition "The realization of A is capable to work", defined thus:
ν = 1 , if the realization of A is capable to work, ν = 0 otherwise;
ω – a logical condition "A executed correctly",
π – a logical condition "Malfunction", defined in the same way as ν.

It is known (for example, [7]) that nearly 90 % of computer errors are the input/output errors and nearly 90% of computer failures require halt/reboot activity. It corresponds universally recognized but not generally accepted concept of pre-operational testing and diagnosis and post-operational detecting and correcting. Let us notice that an operator K is, in general, complex operator representing, for example, the next sequence of operators K = K1 · K2 · K3 , where K1 – a procedure for search of an inoperative components in realization of A, K2 – a procedure for restoration of capability to work of the A realization, K3 – a procedure for testing of capability to work of the A realization.

In different computer systems, most often the procedure K3 is realized by a hardware or hardware-software composition, the procedure K1 – by a software, and the procedure K2 – by a peopleware. In essence, implementation of governing relationship (2.2) implies the next. Before an execution of A, the testing of capability of it is performed. If the A realization is capable, then operator A is executed, otherwise operator K has been executed until the capability of the A realization is completely restored, and then the A is executed.

After execution of the A, the checking of correctness of its execution is performed. If the A was executed correctly, then in accordance with MRSB, an execution of next operator or checking of next logical condition is initiated. If the A is executed improperly, then a cause of its improper execution is clarified – "malfunction" or "fault". In case of malfunction, the execution of operator A is repeated, and in case of fault, the operator K is executed until the capability of the A realization is completely restored, and then operator A is executed. After the repeated execution of A, again the checking of its correct execution is performed. If A is correctly executed, then in accordance with MRSB, an execution of next operator or checking of next logical condition is initiated, and so forth.

Because different realization variants of ν, ω, π and K are feasible, the selection problem of the best their combination (optimization problem) for certain set of components of the Design for Reliability is complex and labor consuming (especially, taking the fixing a set of detecting and correcting means for one component restricts the selection of corresponding means for other components into account). However, in design of a data processing and management information systems on the base of unified computers and peripherals, and also in design of software packages for universal computers, we are forced to use the finished means of realization of ν, ω and π under restricted and regulated options of realization of K more frequently. Because of this, in the majority of cases we don't need to use all options presented with realization of the governing relationship (2.2), and in process of Formal Reliability Synthesis of Algorithms we can use more simple governing relations, which are specific cases of relationship (2.2). One such formal technique to obtain these relationships is fixation of corresponding Boolean variables ν, ω and π identically equal to 1 (i. e., a fixation of the identical truth of corresponding logical conditions).

Under the condition that ν = 1, we obtain

$$A = A \cdot\omega\{ \pi(E \vee K \cdot)\cdot A \} . \qquad (2.3)$$

Under the condition that ω = 1, we obtain

$$A = \nu\{ K \}\cdot A . \qquad (2.4)$$

Under the condition that $\pi = 1$, we obtain

$$A = \nu\{ K \} \cdot A \cdot\omega\{ A \} . \tag{2.5}$$

Under the condition that $(\nu = 1)$ & $(\omega = 1)$, we obtain

$$A = A, \tag{2.6}$$

which is degenerate, i. e. its application means that no detecting and correcting features are conferred to the operator A .

Under the condition that $(\nu = 1)$ & $(\pi = 1)$, we obtain

$$A = A \cdot\omega\{ A \} , \tag{2.7}$$

Under the condition that $(\omega = 1)$ & $(\pi = 1)$, we obtain the governing relationship (2.4), and under the condition that $(\nu = 1)$ & $(\omega = 1)$ & $(\pi = 1)$ – the relationship (2.6).

It is obvious that the governing relationship (2.3) is applied in the case when kA >> P (A) (here, kA – availability coefficient of the A realization), the governing relationship (2.4) is applied in the case when kA << P (A), etc.

3. SOFTWARE RELIABILITY ANALYSIS

In XAAL, the sequential execution of operators (any units of action) A1, A2, … , AN corresponds to the multiplying of these operators:

$$A = A1 * A2 * \ldots * AN. \tag{3.1}$$

Let us assume that PA (TA), PA1 (TA1), PA2 (TA2), … , PAN (TAN) – the correct execution probabilities of the operators A, A1, A2, … , AN in the execution times T, T1, T2, … , TN respectively. It is obvious that

$$PA (TA) = [i = 1, N]\ \Pi\ PAi (Tai), \tag{3.2}$$

and

$$TA = [i = 1, N]\ \Sigma\ Tai. \tag{3.3}$$

For simplicity of the program execution reliability analysis, we use the generalized (canonical) a-disjunction operation instead of a-disjunction operation. The result A of its application to ordered operator sequence A1, A2, … , AN takes the form

$$A = a1(A1\ V\ a2(A2\ V \ldots V\ a[N-1](A[N-1]\ V\ A[N])) \ldots) , \tag{3.4}$$

where a1, a2, … , a[N-1] – logical conditions.

Let us assume that pai = Probability (ai = true). Then, supposing that logic conditions are checked momentarily and absolutely reliably, the correct execution probability of the disjunctive canonical program specified by the expression (3.4) is defined by formula

$$PA (TA) = [i = 1, N]\ \Sigma\ pai * PAi (TAi) + (1 – pa[N-1]) * PAN (TAN), \tag{3.5}$$

where

$$TA = [I = 1, N]\ \Sigma\ pai * TAi + (1 – pa[N-1]) *TAN. \tag{3.6}$$

The expression

$$B = a\{ A \} \tag{3.7}$$

means that the operator B is the result of the a-iteration operation to the operator A.

The correct execution probability of the operator B is defined by the formula

$$PB (TB) = pa / (1 – (1 –pa) * PA (TA), \tag{3.8}$$

where

$$TB = ((1 – pa) / pa) * TA. \tag{3.9}$$

If operators A1, A2, … , AN are executed in parallel the case may be noted

$$A = [A1, A2, \ldots , AN] , \tag{3.10}$$

and we can use the formulas

$$PA (TA) = [i = 1, N]\ \Pi\ PAi (TAi), \tag{3.11}$$

$$TA = \max (TA1, TA2, \ldots , TAN). \tag{3.12}$$

Because for modeling any program execution, it is enough to use operators (3.1), (3.4), (3.7) and (3.10), the probabilities of program correct execution and execution times can be evaluated by formulas (3.2), (3.3), (3.5), (3.6), (3.8), (3.9), (3.11) and (3.12).

We are reminded that the Aspect-Oriented Approach allows us separately develop External and Internal Behavior of Software. In context of Design for Reliability, the External Behavior executes functional data processing without taking into account the natural (unpremeditated) errors, malfunctions, faults, conflicts, etc., and the Internal Behavior realizes aspect data processing tolerated the program execution to conditions of above-mentioned interferences. In context of a Design for Security, the Internal Behavior realizes aspect data processing tolerated an access to and execution of programs to conditions of artificial (premeditated) interferences, without intervention in External Behavior.

The common property for Design for Reliability and Design for Security is redundancy, which can be shared between them only in exceptional situations. It is why resources for Reliability and for Security must be separated in majority cases.

The next technique of automated correction of Software execution uses the functional redundancy approach. Suppose some functional equivalent programs for the same A problem processing are ordered with accordance to their priority: A1, A2, … , AN. Suppose further that ν1, ν2, … , νN – the conditions of potential possibility for correct execution (availability) of A1, A2, … , AN , correspondently. Formally, we can write

$$A = \nu 1(A1 \vee \nu 2(A2 \vee \ldots \vee \nu N(AN \vee W)) \ldots), \qquad (3.13)$$

where W – the signal about the potential impossibility for correct execution of A.

A probability of correct execution of A is

$$PA\ (TA) = [I = 1, N]\ \Sigma\ kAi * Pai\ (TAi) * [l = 1, N; l \neq i\ ]\ \Pi\ (1 - kAl),$$

where

$$T = [I = 1, N]\ \Sigma\ kAi\ ([l = 1, N; l \neq i\ ]\ \Pi\ (1 - kAl))\ TAi\ ,$$

The availability levels kAi , i = 1, 2, … , N , are calculated by traditional structured reliability techniques.

Example. The program SK with structure (hardware) error detection and behavior (software) execution correction (aspect K) looks like

$$SK = S * \varphi(K * \varepsilon(S \vee W) \vee E), \qquad (3.14)$$

where S – the functional equivalent of SK, but without detection and correction; φ – the logical condition of error detection in process of S execution; ε -- the logical condition of data renewal for S repetition; K – the correction code (aspect): $K = A * \alpha(\ \beta\{BC\}\ \chi(\ \delta(F \vee D) \vee D) \vee F)$,

where A – addition of the number of malfunctions to the malfunction counter, B – address forming for the type malfunction counter, C – increment of the type malfunction counter, D – restoration of data for functional code (in our case, for S program), F – transfer of control to emergency diagnostics, α -- logical condition "common number of malfunctions less than N", β -- "the number of given type malfunctions has been counted", χ -- "the current and foregoing malfunctions occurred in the same point of functional code", and δ -- "the current and foregoing malfunctions have the same type".

Let us calculate change of the mean time and correct execution probability of S, realized by technique (3.14), if we know that TS = 100 sec., and PS (100) = 0.9. Also, PA (TA = 2 sec.) = 0.999, PB (2) = 0.999, PC (1) = 0.9995, PD (5) = 0.99, PF (2) = 0.999, pα = 0.8, pβ = 0.5, pχ = 0.1, and pδ = 0.1. First, we calculated, that TK = 8.776 sec., and PK (8.776) = 0.9896. Then, with help of (3.14): TSK = 110.8 sec., and P SK = 0.989. Thus, the mean time of functional program S execution, after addition the error detection and correction of aspect code, increased less than 11%, but the error probability decreased approximately 9 times.

Consideration is being given to the case, when the governing relationships being used in the Reliability Synthesis of the STALS are special cases of the relationship

$$A = \varphi\{\ K\ \}\ A\ \omega\{\ \pi(\ L \vee K\ )\ A\ \}\ , \qquad (3.15)$$

where A – a component of the Design for Reliability [1], with τA (the execution time of A) – a random value with the distribution function FA(t), and the A execution process is accompanied by the Poisson stream of malfunctions with the intensity Λm and Poisson stream of faults with the intensity Λf; L – the operator of restoration of conditions sufficient for repeating of A after a supposed malfunction; TL (the execution time of L) – the random value with the distribution function:

FL (t) = FLm\f (t) , if a malfunction happened, but a fault did not,
FL (t) = FLf (t), if a fault happened, and FL (t) = FLmf (t) in all other cases.
K – the operator of restoration of conditions sufficient for repeating of A after a supposed fault;
TK (the execution time of K) – the random value with the distribution function:

FK (t) = FKm\f (t) , if a fault happened,
FK (t) = FKf (t), if a malfunction happened, but a fault did not and

FK (t) = FKmf (t) in all other cases;
φ – the logical condition "realization of A is capable of working";
τφ (φ testing time) – a random value with the distribution function Fφ (t),
Pφ1 – the probability of a first type mistake (i. e., the probability of mistaken verification of a capacity for work of A realization),
Pφ2 – the probability of a second type mistake (i. e., the probability of mistaken verification of a unable to work of A realization),
ω – the logical condition "A execution is correct",
τω -- the random value with the distribution function Fω(t),
Pω1 – the probability of a first type mistake (i. e., the probability of mistaken verification of a correct execution of A realization),
Pω2 – the probability of a second type mistake (i. e., the probability of mistaken verification of a incorrect execution of A realization),
π – the logical condition "a malfunction is the cause of error",
Pπ1 and Pπ2 – the probabilities of corresponding mistakes of the first and second type (i. e., the probability of mistaken verification of a malfunction or fault).

The state of SAS at the instant t is a pair (X(t), Y(t)), where

$$X(t) \in \{-1, 0, 1\}, \quad Y(t) \in \{ A, L, K, \varphi, \omega, \pi, S, F \}.$$

X(t) – the parameter, characterized the Structure State of SAS and specified as follows:
X(t) = -1 , if in the instant t A realization is unable to work,
X(t) = 0 , if at the last A execution, preceded the instant t, a malfunction happened, but a fault did not, X(t) = 1 in all other cases;

Y (t) – the parameter, characterizing the Algorithmic State of SAS and specifying by a operator (logical condition) execution (testing) in the instant t, and S (F) – an arbitrary component of the Design for Reliability, an ending (beginning) of its execution is interpreted as the beginning (ending) of execution of the right side of the relationship (1).

The computational formulas are obtained for determination of the distribution function, mean time and dispersion of the execution time of the right side of relationship (1) and the probability of its correct execution in a given time. The application package is proposed for a statistical modeling of the STALS with purpose of a time-probability parameter estimation and optimization model stability testing.

4. SOFTWARE RELIABILITY OPTIMIZATION

Analysis of the accumulated experience in formulation and solution of optimization problems in design of algorithms and programs for computerized control and data processing systems can be useful for elaboration of profitable approach to formulation and solution of the appropriate optimization problems in conditions of a computer-aided design. Developers and customers of a developing object pursue the next prime (to a large degree, contradictory) goals.

1. A customer takes an interest to obtain the optimal in some a sense control system. This is most commonly done by specifying one or more quality parameters (efficiency, performance, etc.) of the design object for setting up an optimization problem. As this takes place, a certain amount of design resources identified with the design object can be expressed by the generalized cost criterion of manufacturing, operation and evolution of the object. In each specific case, this criterion can have one or another interpretation, for example it can characterize the realization complexity, volume of r/w memory, overall dimensions, energy consumed by the control system, number and skill of maintenance personnel. The identical with those problems, aimed at optimization of a design object, will be named the Design Object Optimization (Optimization Design).

2. Developers and customers of a control system take an interest to the Design Process Optimization. The interests of developers and customers coincide in an attempt to increase the productivity of a Computer-Aided Design System: a developer endeavors to timely meet contractual obligations to create a control system, and a customer – to timely obtain a capable to work system complete with operational documentation. However, a customer is less interested in the generalized cost of the design system (although the design cost may reflect on the project price) than a developer. At the same time, the quality parameters of a project documentation (especially, its validity) are the governing factors for a customer, because poor

project quality manifests itself in the time of operation in the most unexpected and undesirable form. Notice that the pace of a moral aging of methods and means for control of technological and business processes are ahead of the pace of control system development.

3. Neither the Design Object Optimization, nor the Design Process Optimization results the potentially possible effect of control automation without rational richness of control systems by models, algorithms, software and hardware for the Control Optimization. The realization volume of Control Optimization methods immediately not only causes on cost parameters of Control Systems, but it also influences on cost parameters of their Design Process. The interrelation between the realization volume of Control Optimization and many of quality parameters of Design and Control Systems is also obvious.

To partly illustrate the foregoing, let us cite one example of decision of Design Problem Optimization. The object of Design Problem Optimization is the algorithm of single execution of the control program

$$B = [A0, A4]\ \alpha5\{\ \alpha1(\ A1 \vee \alpha2(\ A2 \vee \alpha3(\ A3 \vee E\ )))\}\ A5\ .$$

(Interpretation of functional operators and logical conditions of the algorithm are contained in [18]. There is also more information about the XAAL.)

It needs to minimize the average time of the algorithm B execution under condition that a probability of its correct execution is no less than 0.995. The required accuracy of the problem solution is five decimal places. The initial data (the durations and correct execution probabilities of functional operators A0, A1, A2, A3, A4 and A5, true probabilities of logical conditions α1, α2, α3 and α5):

t0 = 0.01 min., t1 = 1 min., t2 = 1 min.,
t3 = 5.00 min., t4 = 0.001 min., t5 = 0.001 min.,
P0 = 0.999, P1 = 0.900, P2 = 0.950,
P3 = 0.950, P4 = 0.999, P5 = 0.999,
p1 = 0.800, p2 = 0.150, p3 = 0.999, p5 = 0.100.

For organization of error detection and correction of the system's internal behavior, in efforts to enhance its reliability, we apply the next governing relationship

$$Ai = \underbrace{Ai\ \omega i(E \vee Ai\ \omega i(E \vee \dots Ai\ \omega i(E \vee Ai)) \dots)}_{Xi}\ ,$$

where ωi – logical condition of correct execution of the functional operator Ai, E – the identical operator.

Let us presume

Pi (Xis) >= 0.99999 , i = 0, 1, 2, 3, 4, 5 , (s – sufficient)

and calculate the vector of sufficient conditions for accomplishment of required reliability level:

$$Xs = (1, 4, 3, 3, 1, 1).$$

Algorithm B is a linear regular algorithm of the form

$$B = L1\ L2\ L3\ ,$$

where

$$L1 = [\ A0, A4\ ]\ ;$$
$$L2 = \alpha5\{\ \alpha1(\ A1 \vee \alpha2(\ A2 \vee \alpha3(\ A3 \vee E\ )))\ \}\ ;$$
$$L3 = L5\ .$$

The sufficient conditions for algorithms L1, L2 and L3:

$$PL1\ (XL1) >= 0.9983\ ,\quad PL2\ (XL2) >= 0.9983\ ,\quad PL3\ (XL3) >= 0.9983\ .$$

Let us test the conditions:

$$PL1\ (0) = P0 \cdot P4 = 0.999 \cdot 0.999 = 0.998\ ;$$
$$PL2\ (0) = p5\ /\ (1 - (1 - p5)\ (\lambda1 \cdot P1 + \lambda2 \cdot P2 + \lambda3 \cdot P3 + \lambda4 \cdot 1)) =$$
$$= 0.1\ /\ (1 - 0.9(0.8 \cdot 0.9 + 0.03 \cdot 0.95 + 0.1698 \cdot 0.95 + 0.00017)) =$$
$$= 0.5525\ ;$$
$$PL3\ (0) = P5 = 0.999\ .$$

Thus, only the algorithm L3 = A5 meets the sufficient conditions, and because of this we can set X5 = 0 . Based on the known value of Xs, we select the initial values of the vector XL1, X0 = 1 and X4 = 1, i. e. XL1(0) = (1, 1). We solve the optimization problem for corresponding parallel algorithm by the coordinate hauling down technique starting from the sufficient value location

X0 = 0 and X4 = 1 , i. e. X L1 = (0, 1).

Now, let us define the accomplished probability of correct execution of algorithm L1:

$$PL1\ (XL1) = P0\ (\ 1 - (\ 1 - P4)**2 =$$
$$= 0.999\ (\ 1 - (\ 1 - 0.999)**2\ ) = 0.99899$$

and refine the sufficient conditions for algorithm L2:

$$PL2\ (XL2) >= P0\ /\ (\ PL3\ (XL3) \cdot PL1\ (XL1) =$$
$$= 0.995\ /\ (\ 0.999 \cdot 0.99899\ ) = 0.997.$$

Algorithm L2 is iterative regular algorithm

$$L2 = \alpha5\{\ C\ \}\ ,$$

where

$$C = \alpha1(\ A1\ V\ \alpha2(\ A2\ V\ \alpha3(\ A3\ V\ E\ )))\ ,$$

And based on this, we can define the restriction for algorithm C:

$$PC\ (XC) >= (\ PL2\ (XL2) + q5 - 1\ )\ /\ (\ PL2\ (XL2) \cdot q5\ ) =$$
$$= (\ 0.997 + 0.9 - 1\ )\ /\ (\ 0.997 + 0.9\ ) = 0.99967\ .$$

After solution of optimization problem for the corresponding disjunctive regular algorithm, we obtain:

$$x1 = 3,\ \ x2 = 1,\ \ x3 = 2, \quad \text{i. e.} \quad XL2 = (\ 3, 1, 2\ ).$$

Thus, the necessary and sufficient conditions for optimal ensuring of the required reliability level of the algorithm B are:

$$x0 = 0,\ \ x1 = 3,\ \ x2 = 1,\ \ x3 = 2,\ \ x4 = 1 \ \text{ and } \ x5 = 0,$$
$$\text{i. e .}\ \ X = (\ 0, 3, 1, 2, 1, 0\ ).$$

After Reliability Synthesis, the control program takes the form:

$$D = (\alpha0\ V\ \alpha4)\{\ [\ A0, A4 \cdot \omega4(\ E\ V\ A4\ )\ ] \cdot \alpha5\{\ \alpha1(\ A1 \cdot \omega1(\ E\ V$$
$$A1 \cdot \omega1(\ E\ V\ A1 \cdot \omega1(\ E\ V\ A1\ )))\ V\ \alpha2(A2 \cdot \omega2(\ E\ V\ A2\ )\ V$$
$$\alpha3(A3 \cdot \omega3(\ E\ V\ A3 \cdot \alpha3(\ E\ V\ A3\ )))\ V\ E\ )))\ \}\ A5\ \}\ .$$

Appraising the average time of single execution of the control program before and after an optimization, we obtain correspondently:

$$tB\ (0) = 15.123\ \text{min.} \quad \text{and} \quad tB\ (X) = 16.002\ \text{min.}$$

Thus, we met required reliability level by an increase the average time of the algorithm execution less than 6 %.

The problems of optimization for business automation and data processing specified by the XAAL are being solved for a variety of criteria. The most frequently used parameters are a time of algorithm execution, probability of correct execution, complexity, and price. Any from these parameters (and also the number of functional operator types) may be selected as the goal function F, and other parameters must fit the restrictions. We developed different techniques for the Software Aspect Optimization problem based on the Branch and Bound, Convex-Programming and Dynamic Programming Methods.

5. CONCLUTIONS

On base of the Aspect-Oriented Approach to Software Engineering, we developed and implemented the common methodology and set of techniques for Optimal Enhancement of Software Reliability. The techniques use unified for all stages of Software Development Extended Algorithm Algebra Language, simple probability models of Reliability Analysis, formal methods of Reliability Synthesis and effective methods of Reliability Optimization. The next step of this R&D is Aspect-Oriented Software Development for the Adaptive External (Market Intelligence) and Internal (Conflict Resolution) Corporate Governance, for the Decision Support Systems of City Governments in emergency situation, for Reliability and Security Optimization.

6. REFERENCES

1. AATT TO24. Communications System Architecture Development. – AATT TO24 SAIC. – http://www.grc.nasa.gov/WWW/avsp/wxap2000/SAIC/sld001.htm
2. I. Alexander. Goal Patterns Generate Scenarios. Paper at RESG Scenarios Day, 1999. http://easyweb.easynet.co.uk/~iany/consultancy/goalpatt/goalpatt.htm

3. J. H. Andrews. Process-Algebraic Foundations of Aspect-Oriented Programming. The Third International Conference on Metalevel Architectures and Separation of Crosscutting Concerns (REFLECTION 2001), Kyoto, Japan (September 25-28, 2001). A. Yonezawa and S. Matsuoka (Eds.). LNCS 2192, Springer-Verlag, pp. 187—209, 2001.
4. B. Bachmendo and R. Unland. Aspect-Based Workflow Evolution. In Aspect-Oriented Programming and Separation of Concepts. Proceedings of the International Workshop, A. Rashid and L. Blair (Eds.), pp. 13—19, Lancaster University, UK, 24 August, 2001.
5. B. Bieker and E. Maehle. User-Transparent Checkpointing and Restart for Parallel Computers. In Foult-Tolerant Parallel and Distributed Systems, pp. 385—399, Boston: Kluwer Academic Publishers, 1998.
6. S. Clarke and R. J. Walker. Towards a Standard Design Language for AOSD. Aspect-Oriented Software Development. Proceedings of the 1st International Conference on Aspect-Oriented Software Development, pp. 113-119, Enschede, The Netherlands, 2002.
7. E. Daniel, R. Lal, and G. Choi. Warnings and Errors: A Measurement Study of a UNIX Server. The 29th International Symposium on Fault-Tolerant Computing. Madison, Wisconsin, USA, June 15-18, 1999.
8. V. Degtiar and I. Safonov. Evolutionary Mechanism of Conflict Resolution. New York, NY: Plenum Publishing Corporation. No. 2401-0114, 1988.
9. V. Degtiar and I. Safonov. An Evolution-Stable Conflict-Reducing Mechanism with Side Payments. New York, NY: Plenum Publishing Corporation. No. 2602-0297, 1990.
10. A. Frohlich and W. Schroder-Preikschat. High Performance Application-Oriented Operating Systems – the EPOS Approach. In Proceedings of the 11th Symposium on Computer Architecture and High Performance Computing, pp. 3—9, Natal, Brazil, September 1999. http://citeseer.nj.nec.com/augusto99high.html
11. V. Glushkov, A. Barabanov, L. Kalinichenko, S. Michnovskiy, and Z. Rabinovich. Computers with Developed Interpretation Systems. Kiev: Naukova Dumka, 1970.
12. G. Kiczales, J. Lamping, A. Mendhekar, C. Maeda, C. Lopes, J.-M. Loingtier, and J. Irwin. Aspect-Oriented Programming. In ECOOP'97 – Object-Oriented Programming, 11th European Conference, LNCS 1241, pp. 220—242, 1997. http://citeseer.nj.nec.com/kiczales97aspectoriented.html
13. I. Kiselev. Aspect-Oriented Programming with AspectJ. Indianapolis, IN: SAMS, 2002.
14. S. W. Loke and A. Zaslavsky. Towards Distributed Workflow Enactment with Itineraries and Mobile Agent Management. E-Commerce Agents. J. Liu and Y. Ye (Eds.). LNAI 2033, pp. 283—294, Springer-Verlag, Berlin, Heidelberg, 2001.
15. J. S. Plank, Y. Chen, K. Li, M. Beck, and G. Kingsley. Memory Exclusion: Optimizing the Performance of Checkpointing Systems. Technical Report UT-CS-96-335, University of Tennessee, August 1996.
16. F. Quaglia, B. Ciciani, and R. Baldoni. A Checkpointing-Recovery Scheme for Domino-Free Distributed Systems. In Foult-Tolerant Parallel and Distributed Systems, pp. 93—107, Boston: Kluwer Academic Publishers, 1998.
17. A. Romanovsky. Exception Handling in Component-Based System Development. In the 25th International Computer Software and Application Conference (COMPSAC 2001), Illinois, USA, October, 2001.
18. I. Safonov. Design for Reliability of Control Algorithms. Vladivostok, USSR: VINITI, 1982.
19. I. Safonov. Trust Engineering and Risk Management of Complex Systems. Proceedings of the International Scientific School “Modelling and Analysis of Safety, Risk and Quality in Complex Systems”, pp. 62—65. June 18-22, 2001, Saint-Petersburg, Russia, 2001.
20. J. Whaley. System Checkpointing Using Reflection and Program Analysis. The Third International Conference on Metalevel Architectures and Separation of Crosscutting Concerns (REFLECTION 2001), Kyoto, Japan (September 25-28, 2001). A. Yonezawa and S. Matsuoka (Eds.). LNCS 2192, Springer-Verlag, pp. 44—51, 2001.
21. I. S. Welch, R. J. Stroud, and A. Romanovsky. Aspects of Exceptions at the Meta-Level. The Third International Conference on Metalevel Architectures and Separation of Crosscutting Concerns (REFLECTION 2001), Kyoto, Japan (September 25-28, 2001). A. Yonezawa and S. Matsuoka (Eds.). LNCS 2192. Springer-Verlag, pp. 280-281, 2001.

TRUST ENGINEERING AND RISK MANAGEMENT FOR SAFETY OF METROPOLIS AND MEGALOPOLIS CITIZENS

Brian Bailey
●
Digital Security International
2703 Arlington Blvd., Suite 101
Arlington, VA, USA
bbailey@dciencryption.com

Igor Safonov
●
International Unity Science Institute
1011 Arlington Blvd., Suite 403
Arlington, VA, USA
isafonov@aol.com

ABSTRACT. The article describes the problems and solutions in the field of safety enhancement in emergency situations of the complex urban agglomerations and analyses of the most actual problem for all metropolises and megalopolises – terrorism, proposing the rational models and techniques of counterterrorism strategy, based on knowledge and experience.

Keywords: metropolis, megalopolis, emergency, terrorism, trust engineering, risk management, safety, models, evolutionary games, tools.

INTRODUCTION

Terrorism threats and terrorist activities became the most important factors of people troubles and government care ahead of traditional risk factors of nature, economic and technological characters. We are in need of tools and resources for engineering and management of safety in emergency situations for big cities in the most economically attractive way. The conceptions of Trust and Risk are at the focus of our attention.

"Terrorism threatens us deeply because it puts into question our ordinary lives and the trust we need to conduct them. … Sociologists sometimes say that trust is the glue that holds society together. … You could say that terrorism poisons the social glue, inspiring fear that it just won't stick any longer. When you stop to think about it, terrorists could operate nearly anywhere. A taxi? Didn't one of those hijackers work for a while as a taxi driver? Your cup of tea? Who had access to the water used to make it? That polite young man at the Xerox machine? He might be making false documents to support somebody who wants to launch another attack." [14]

In the first part of the article we describe the problems and decisions in the field of safety enhancement in emergency situations for one of the complex urban agglomerations – Washington, the capital of the United States of America, the country located between two choppy oceans, achieved technological peaks, abused human possibilities, and attracted attention of terrorists. The second part of article analyzes the most actual for all metropolises [22] and megalopolises problem – terrorism, proposes the most rational models and techniques of counterterrorism strategy, based on experience and common sense.

There are known three kinds of Trust [13]: 1) Strategic Trust – trust that the organization is doing the right things (goal and strategies), 2) Organizational Trust – trust in the way things are being done (processes and decision making), 3) Personal Trust – trust in the people leading the organization (trust in you, trust in them). Analysis of these three kinds of trust shows us that we have a Goal, Behavior (strategies, processes, and decision making), and a Structure (you and them), but that we have not yet a Resource model. This means we cannot formulate and decide problems of Trust Engineering not only optimally, or even rationally. In a similar manner, we have the same incompleteness in Risk Management.

Big cities – big troubles. Not only for regular people, but also for businesses. On November 26, 2002, President Bush signed into law The Terrorism Risk Insurance Act of 2002 (TRIA). The primary objective of the TRIA is to mainly ensure the availability of commercial property and casualty insurance coverage for losses resulting from acts of terrorism. The TRIA will also allow for a transitional period for the private markets to regain stability, resume pricing and build capability to absorb damages in the future. But companies at risk of a terrorist attack are rejecting the expensive premiums sought by insurers for required coverage, sending a signal that TRIA could fail to meet its targets. Several high-risk groups have recently rejected the TRIA policies because the quotes were either too expensive or they felt they would be able to negotiate with rival insurers.

We propose the original decision of the problem using Anti-Terrorism Engineering and Management Approach (ATEMA), which is the part of Trust Engineering and Risk Management (TERM) approach and framework. The ATEMA problems, in contrast to more regular and traditional TERM problems, are characterized by the lack of understanding a terrorist processes, imperfect investigation of terrorist events, and not sufficiently developed models of decision making for loss prevention. Furthermore, the factor priorities are cardinally different. In this case, the immigration policies and technologies have become the subjects of care and study.

There are three major areas in which changes in immigration policies and technologies may be able to counter future terrorist threats: visa issuance and entry inspections, border controls, and interior enforcement. Of course, in considering the problems highlighted by the terrorist attacks, and the options to head off future attack, it is important to reach a trade off between enhanced security and reliability of safety procedures and privacy and liberty in Open Society.

The objects of our research and development are safety problems and procedures of megalopolis and metropolis citizens, partly based on one author experience in terrorism investigation and the Loss Prevention Program creation for Washington, DC Government, and on second author experience in cyber crime prevention technologies for governmental and corporate customers.

In the procedural direction, we can investigate the real world problems of criminal activities, which can help terrorists and threat citizens of megalopolises and metropolises, on base of mournful experience of New York City, District of Columbia, Tokyo, London, Moscow, and other big cities of the World. We must concentrate our attention on fact and document falsification as one of the major factor of terrorism oriented frauds, provide the most typical case studies (personal and corporate identity thefts, criminal placement in banks and businesses), and propose recommendations and procedures of fraud prevention and detection based on optimal or rational use the available resources and restricted time. In the technological direction, we can analyze, compare and recommend devices, systems and technologies of document control and fraud detection for government and businesses, propose methodology of multicriteria selection of appropriate equipment, services and its vendors, and direct the way to improvements in the techniques and technologies. In a general way, we must propose models and tools of resources trade off between procedures and technologies of Safety Enhancement and Loss Prevention for citizens, businesses and governments of metropolises and megalopolises.

A Loss Prevention Program (LPP) have been developed for prevention and elimination of human suffering, life and resource loss of the District of Columbia government, officers and employees in process of disaster events and emergency situations. Our problems were close, but no similar to problems of the District Response Plan (DRP) [10]: “The DRP provides a new framework for District Government entities to respond to public emergencies in the metropolitan Washington area. The DRP provides a unified structure for District emergency response operations to ensure a coordinated and effective operation. The plan describes how District agencies will work collaboratively within the District and with our regional and federal partners. The ultimate goal is to protect the public and respond efficiently and effectively to significant incidents that threaten life, property, public safety, and the environment in the District of Columbia.”

If the DPR has orientation toward the external goals and behavior of DC Government, the LPP was directed toward the internal goals of DC Government, in particular to protect own officers and employees, structure and behavior in emergency situations. Therefore, the first steps of the LPP development were concentrated in the next areas:

1. Reasons why the LPC is needed and what must be developed – employee motivation to participate in the LPC development and collaborate with the LPC developers.

2. Efficient goal decomposition on personal and team objectives for effective loss prevention and limitation directed to create a safe working environment ready for disaster events and emergency situations.

3. Complete (all functions and all employees) emergency responsibility distribution between all levels of job executors with rational redundancy for the LPP reliability.

4. Communication with and between top managers for information and knowledge feedback oriented on the LPP correction and adaptation with accordance with DC Government goals, it employee responsibilities, real circumstances and accessible resources.

MOTIVATION. DC Government managers, officers and other employees are motivated, like all others members of the human species, by species-wide needs for food, etc.; needs for safety, protection, and care; needs for gregariousness and for affection-and-love relations; needs for respect, standing, and status, with consequent self-respect; and by need for self-actualization or self-fulfillment of the idiosyncratic and species-wide potentialities of the individual person [18]. In emergency situations everything looks less important than safety and everybody may be characterized as living almost for safety alone.

SAFETY RULES & PROCEDURES. The common safety rules and procedures in emergency situations are inherited from normal situations, but must be reengineered for emergency because obvious restrictions in accessible resources and limited time for making decisions. These rules and procedures will be analyzed in context of forecasting (modeling) emergency events and situations and modified for real emergency conditions. Following DRP definition, during the normal situations (Normal Operations) DC agencies, divisions, managers, officers and employees “should be engaging in preparedness, training, and exercise activities to ensure continual readiness.” [10]

EMERGENCY RULES & PROCEDURES. The special safety rules and procedures in emergency situations must be created and developed on the base of world experience and DC peculiarity as a capitol of the USA. In this case, we need to take into account the next three operation levels of emergency proposed by the DPR: Operation Level 1 – a monitoring phase triggered by the potential threats for life, property, or the environment; Operation Level 2 – a partial activation of the CMT triggered by highly probable hazardous conditions and a strong potential for property damage and loss of life; Operation Level 3 – a full activation of the CMT triggered by extremely hazardous conditions that are imminent or occurring.

MONITORING & INSPECTIONS. Following to the Federal Response Plan (FRP) [12], the DC has identified 15 Emergency Support Functions (ESF) as the priorities of emergency preparedness and loss prevention. These functions are our guiding lines for internal monitoring and for external inspections. Of course, the priorities must be established with taking into attention the DC peculiarity as a capitol of US and operation level of emergency. Also, we reserve the right to extend or (and) modify the ESF list:

1. Transportation
2. Communications
3. Public Works and Engineering
4. Firefighting
5. Information and Planning
6. Mass Care
7. Resource Support
8. Health and Medical Services
9. Urban Search and Rescue
10. Hazardous Materials
11. Food
12. Energy
13. Law Enforcement
14. Media Relations and Community Outreach
15. Donations and Volunteer Management

According to a vulnerability assessment of the District, there are five major categories of hazards that may pose a threat to the District: 1) Natural Hazard – sever weather, hurricanes, tornadoes, flooding, or earthquakes; 2) Infrastructure Disruptions – utility and power failures, water supply failures, critical resource shortages, or exploding manhole covers; 3) Human-caused Events and Hazards – urban fires, special events, civil disorder, or transportation accidents; 4) Technological Hazards – hazardous materials, radiological, biological, or computer-related incidents; and 5) Terrorist Incidents – bomb threads, sabotage, hijacking, or armed insurrection, which threaten life or property. Terrorist attacks can also be conduits through which biological, chemical, and radiological agents can be employed.

REPORTING, INVESTIGATION, ANALYSIS & CORRECTION. All accidents and injuries must be reported immediately. Very important part of this activity is the performance management in conditions of emergency. The traditional performance management [1] is the systematic process of: planning work and setting expectations, continually monitoring performance, developing the capacity to perform, periodically rating performance, and rewarding good performance. This systematic process is never happened in real life and even in a no-emergency case is subject to mistakes, malfunctions, and failures. The performance management becomes especially difficult in emergency cases, but composition of Trust Engineering and Risk Management means does the problem solvable.

TRAINING & CONSULTING. Training and consulting are important components of the loss prevention and control process. Their importance cannot be overestimated. The training course must cover the basics of emergency management, the role of your department and your personal possibilities and responsibilities in case of emergency events and in process of emergency situation. Therefore, as you study our course and participate in training, think about adapting the information and knowledge to your particular job and to your personal safety. The format of our course is design to help DC managers, officers and employees learn and apply to team and person safety the principles, rules and procedures involve in emergency management and self-protection. First of all you will study the concept of Comprehensive Emergency Management (CEM) [34], which consist of three interrelated components: 1) All types of hazards, 2) An emergency management partnership, and 3) An emergency lifecycle. But before the detail explanation of these components, allow us to acquaint you with principal notions of the CEM.

Emergency is defined as any event or (and) situation, which threatens (threaten) to, or actually does, inflict damage to property or people. Large disasters can range from hurricanes and floods, to explosions and toxic chemical releases. Management has a traditional definition as the coordination of an organized effort to attain specific goals or objectives. In our case, emergency management means an organized effort to mitigate against, prepare for, respond to, and recover from an emergency. Comprehensive clarifies "emergency" by including all forms of natural, technological, human-caused and infrastructure hazards which threaten or adversely affect lives and properties; by bringing together the proper mix of resources from the federal, state, and local governments, from business and industry, and from the public; by adding phases of disaster lifecycles. The four phases of CEM are: 1) Mitigation, 2) Preparedness, 3) Response, and 4) Recovery.

Mitigation: Any activities that actually eliminate or reduce the occurrence of a disaster. It also includes long-term activities that reduce the effects of unavoidable disasters.

Preparedness: The activities are necessary to the extent that mitigation measures have not, or cannot, prevent disasters. In the phase, government, organizations, and individuals develop plans to save lives and minimize disaster damage. Preparedness measures also seek to enhance disaster response operations.

Response: The activities follow an emergency of disaster. Generally, they are designed to provide emergency assistance for casualties. They also seek to reduce the probability of secondary damage and to speed recovery operations.

Recovery: The activities continue until all systems return to normal or better state. Short-term recovery returns vital life support systems to minimum operating standards. Long-term recovery may continue for a numbers of years after a disaster. Their purpose is to return life to normal, or improved levels.

In our course, we recommended the DC government employees to recognize and share the basic philosophy of promoting safe and secure urban planning [20]:

1. To assume their respective roles, help each other, and liaise with each other in order to promote the realization of safe families, teams, and themselves.
2. To foster a wide range of department, team and community emergency activities and good relations with other employees for ensuring the safety and security.
3. The lessons, experience, and knowledge gained from living through disaster, crime, and accident will be put to good effect if everyday life and duty in order that we may be prepared for emergencies, and in order that we may hand our wisdom down to future generations.

HARDWARE AND SOFTWARE. The consulting services in case of emergency are oriented on future support by the Integrated Transportation and Public Safety Wireless Information Network (CapWIN) – the common project of DC and the States of Maryland and Virginia [4]. In the Washington Metropolitan Region – MD, VA, and DC, more than one hundred various fire, transportation, police and emergency medical services agencies are available to respond to emergency and life threatening incidents that impact public safety. These emergency services agencies utilize individual, proprietary communications systems that limit the user's ability to quickly share vital information with other responding agencies.

The CapWIN project will integrate transportation (ESF # 1) and public safety data and voice communication (ESF # 2) systems in two states and the DC and will be the first multi-state transportation and public safety integrated wireless network in the US.

Potential benefits of CapWIN for emergency events' and situations' consulting are: 1) "Real time" information to improve decision making and resource allocation; 2) Improve response to natural, technological and man-made disasters; 3) Increased accuracy and reliability of reports, investigations and analysis; 4) Direct communications between mobile units of departments and agencies; 5) Enhanced safety for government employees and their families.

PERSONNEL SELECTION & PLACEMENT FOR EMERGENCY ACTIVITY. Traditionally, this is to insure, that the best-qualified person is hired and placed based on job qualification standards, but in our case we need to orient all managers, officers and employees on very special conditions of emergency. Good personnel are the most valuable assets of an operation. Poorly performing workers can severely constrain and hamper a program. So it follows that personnel evaluation is a critical function of disaster managers.

The selection of the right person for a specific job is crucial in both normal and emergency situations. In pre-disaster situations, such as disaster mitigation and preparedness programs, the staff size is constant and usually small. A manager must be able to evaluate each person and assign him or her to the right task. In post-disaster environments, a program staff expands quickly for the emergency, and then contracts as rehabilitation and reconstruction phases occur. For this reason, the manager must constantly assess the staff to ensure that each project is being properly executed. When the size of the organization is reduced, the manager must carefully evaluate the staff to determine whom to let go. In disaster management, there are two purposes for personnel evaluation: to provide the basis for making staffing decisions during the transition between phases of a disaster and to help improve the performance of the operation by determining what aspects of an individual person's work need improvement. Thus, personnel evaluation is an important control technique.

The task of fairly, thoroughly and regularly evaluating the performance of others is a difficult one, but is indispensable to smooth operations. Subordinates need to know how they are doing; managers need to know how their subordinates are performing; and organizations need to know if personnel are being used effectively. Personnel evaluations must be approached carefully. If conducted poorly or with disregard of people's emotions, the evaluation will be disruptive, and it will serve little, if any, purpose. A manager's task is to develop a systematic evaluation process that is meaningful, fair and comprehensive. In modern management, the term "performance appraisal" is often used instead of "personnel evaluation," as it is considered to be less threatening.

WHAT & HOW. Information Technologies (IT) not only had absorbed a lot of scientific and empirical results from different fields of human activity, but also received a lot of own results, which can be feasibly implemented into other technologies. We hope to actively use the IT (primary, methodological) results for loss prevention (engineering) and control (management) programs. One of our approaches based on concern separation and aspect engineering. The approach is grounded on principal difference between External and Internal Behavior of any object what help us to separate goal functions and their aspects for more convenient implementation of system engineering and management.

The functions of external behavior are regular, but the functions of internal behavior are casual. Yet, if the functions of external behavior are goal (objective) determined, the functions of internal behavior are common for different goals (objectives). Absolute different departments or employees have a lot of common internal functions (for safety, performance, reliability, security, quality, etc.). Separation of concerns (particularly, in software engineering) has always been a very natural means to handle complexity of (software) development. However, modularizing concerns can be a very tricky task for the developer and raise some issues such as performance, crosscutting, or redesigning when the software is used in a context that is quite different from the overseen one. By handling crosscutting within the language or system, the recent approach of Aspect-Oriented Programming (AOP) seems to be a very promising way for helping developers to handle separation of concerns and to overcome the drawbacks of traditional design approaches.

However, if AOP introduces a new programming paradigm that complements existing ones, it is clear that it brings a new bunch of difficult but solvable problems, which can not be practically solved in Object-Oriented Programming. The main of them is an optimization. We developed a lot of different models and techniques algorithm and program optimization, which can be used in Trust Engineering for Emergency

Availability Support and in Risk Management for Emergency Loss Prevention. In order to define problems and generate novel courses of action, we need to draw on our experience to make judgment about [17]: reasonable goals and their attributes; the appearance of the anomaly; the urgency of solving a problem (whether to take anomalies seriously or treat them as transient that will go away); what constitutes an opportunity worth pursuing? Which analogues best fit the situation, and how to apply them? The solvability of a problem. It seems, as there are two primary sources of power for individuals in emergency situation problem solving:

Pattern matching (the power of intuition) provides us with a sense of reasonable goals and their attributes. It gives a basis for detecting anomalies and treating them with appropriate seriousness. It helps us to notice opportunities and leverage points, discover relevant analogies, and get a sense of how solvable a problem is. The judgment of solvability is also responsible for letting us recognize when we are unlikely to make more progress and that it is time to stop.

Mental simulation (the power of imagination) is the engine for diagnosing the causes of the problem, along with their trends. It plays a role in coalescing fragmentary actions to find a way to put them together. And it is the basis for evaluating courses of action. The themes covered thus far in reviewing problem solving and decision-making are the core components for a perspective on naturalistic decision making.

The next question is: Can we research the terrorist activity in the same way as we study the majority of surrounding us processes? In other words, can we exploit the scientific methods? The article with intriguing name “Modeling for Terrorism” [33] makes one of the first attempts to answer on this question in the affirmative agrees. Tom Stamer analyses three model approaches and corresponding techniques, proposed by Risk Management Solution, EQECAT, and Worldwide Corporation.

GAME THEORY APPROACH. The approach based on the Game Theory supposes that targets and techniques of possible terrorist attacks can be modeled by behavioral structures and parameters of terrorist organizations. The Risk Management Solution (RMS) developed an application called U.S. Terrorism Risk. The main goal of the U.S. Terrorism Risk is quantification of catastrophic terrorist attack risk. The model uses information from terrorism experts, estimates the probabilities and costs of property damages, business interruptions, casualties and injuries, taking into account 16 modes of attack. The modes are based on 4 types of terrorist weapons – biological, chemical, nuclear and radiological. High-resolution of simulation tool allows to model a lot of loss and damage agents, from blast pressure to airborne and ground-based contaminants. The simulated events cover close to 1,500 potential terrorist targets in the United States of America – business centers of megalopolises, government districts of metropolises, facilities, landmarks, etc. The model is focused on the most probable attacks and uses reflection approach to understand the corresponding models of an enemy.

PROBABILITY THEORY APPROACH. The approach based on the theory of probability, was developed by EQECAT Inc. It strikes by its dimensionality: the model takes into consideration hundreds of thousands terrorist targets and millions of events. President of EQECAT (Oakland, California) Richard Clinton says: “We believe our model is the only one currently available that is fully probabilistic and covers all relevant risk sources, including bomb blast, aircraft impact, and CBNR (chemical, biological, nuclear and radiological) weapons for all 50 states and the District of Columbia”. The National Council on Compensation Insurance (NCCI) selected the Terrorism Model of EQECAT for terrorism loss evaluation in every state of the USA. Here is the NCCI opinion about the model: “By definition, events that cause catastrophic losses occur infrequently but have the potential to create massive claims costs. … For example, predicting the annual number of hurricanes or major earthquakes with any precision is impossible, in spite of more than a century of experience and extensive meteorological and geological/seismic research. In the case of terrorism events, we are fortunate to have few historical data points for the U.S., but this means that forecasts for the likely number of future terrorist events can be little more than conjecture.

The catastrophe-modeling firm EQECAT, at NCCI’s request and with its support, developed a terrorism model that clearly details the devastating potential consequences of likely terrorist events. Using NCCI’s terrorism model to analyze a range of specific events (e.g., truck bombs, sarin gas, chlorine, anthrax) confirms that the workers compensation losses alone from a single event could have a devastating financial impact on a significant portion of the country’s property and casualty industry. This would create major hardships for the families of workers killed or injured in the attack, and extensive financial and administrative burdens for insurance regulators, policyholders, and the U.S. economy. NCCI’s analysis also confirms that this is a problem for all regions of the country—not just major metropolitan areas." The EQECAT model is supported by ABS Consulting's MIDAS software, which has been used for counterterrorism planning and response. It also helps insurers to optimize their risk portfolio, -- says Clinton. Because attacking the heartland of any country might be easy for terrorists, but could have a psychological impact on the country, Clinton recommends using the model and software for evaluation of probability and loss of midsize and small cities and towns.

DELPHI METHOD APPROACH. Developed by RAND Corporation the Delphi Method uses special procedure for processing of expert opinions and allows forecasting a place, time, means and impact of terrorist attacks. AIR Worldwide Corporation (Boston, Massachusetts) applies the method for estimation of numbers and sites of attacks. The AIR model is supported by the database of potential terrorist targets – buildings, bridges, tourist attractions and national infrastructure. The model was chosen for the terrorism preparedness exercise Silent Vector, where the roles of government leaders were played by former Virginia Governor James Gilmore, former Senator Sam Nunn, former FBI Director William Sessions, and former CIA Director James Woolsey. "The lessons learned from Silent Vector will help the government prepare for, and possibly deter, future attacks in the United States," says Jack Seaquist, product manager from the AIR.

GENERALIZED DYNAMIC SYSTEMS. All of these approaches are known for dozens of years and have been applied for forecasting and analysis of very complex processes with big degree of vagueness. The most powerful models and tools were created by Viktor Glushkov's team [2, 16, 23-28, 30,31] in the Kiev Institute of Cybernetics, named now the V.M.Glushkov Institute of Cybernetics. The experience of application Event- and Process-Forecasting systems, based on the Theory of Generalized Dynamic Systems, for Politic, Economic, Social, Science, and Engineering forecasting and analysis situations approved not only their wide possibility, but also demonstrated a lot of restrictions. Ignoring of these restrictions in the interest of special groups (lobbyists) is open to many hazards. For example, it may be in interests of the nuclear military industry. A shorthand text of the illustrated statement is cited below [29].

"Foresee and foreworn! Looking through an article "Don't be afraid of the nuclear winter" published in "Rossiyskaja Gazeta" May 16, 1992, we consider our professional duty to express discrepancy with stated in the article of the Associated Press information about necessity of reconsideration of climate consequence forecasts of the large-scale nuclear war – so called "nuclear winter". The article informs that scientists from a number of the US scientific centers consider that relatively weak and local changes of air temperature near earth in result of the Kuwait oilfield fires confirm that climate consequences of nuclear war may be small and the maiden earlier nuclear war forecasts must be revised. This assertion is mistaken. The analysis of climate consequences of Kuwait fires does not invalidate by any means the correctness of nuclear winter calculations. A nuclear winter is caused by nuclear bombing and followed by gigantic fires in big cities, when burning products elevate to a higher troposphere and stratosphere (to 10 km) and there they firstly extend toward the Northern Hemisphere followed by the Southern. Conflagrations in cities and oilfields differ essentially by composition of their combustibles, fire characters and consequences. The burning products from oilfield fires did not elevate to big height. That is why the temperature of air close to ground surface was comparatively less changed. By this means it is beyond reason to reconsider forecasts of nuclear winter because of information about conflagrations in Kuwait."

ASYMMETRIC INFORMATION AND EVOLUTIONARY GAMES. The Nobel prizes of last years (John C. Hasanyi, John F. Nash, and Reinhard Selten – 1994, James A. Mirrlees and William Vickrey – 1996, George A. Ackerlof, A. Michael Spence, and Joseph E. Stiglitz – 2001) and a movie "A Beautiful Mind" about a great mathematician John Forbes Nash attract public attention to the Game Theory and Asymmetric Information for modeling of economic conflicts, contemporary wars, emergency situations and counterterrorism activity. Yet, a "hungry" market also attracted a lot of popularizers and advertisers, which inadequately evaluate orientation and availability of these mathematic tools, causing the discredit of all the scientific movement, created by such Titans as John Von Neumann and John Maynard Smith. Using concepts taken from the theory of games formulated by John von Neumann in the 1940s, Maynard Smith introduced the idea of an Evolutionary Stable Strategy (ESS) in the 1970s. Assuming that two animals are in conflict, then an ESS is one that, if adopted by the majority of the population, prevents the invasion of a mutant strategy. Stable strategies by definition thus tend to be mixed strategies. Many aspects behavioral pathology of human relations from economic fraud to terrorist activities may be investigated and prevented with the Evolutionary Games models developed by Maynard Smith. Modeling of intra-corporation (collaboration) and inter-corporation (competition) relations demonstrated that infiltration of criminals can be detected and their influence can be restricted using local- or wide-area networks (Intranet, Internet, etc.) and corresponding software [6-9]. Our experience in development and application the Theory of Evolutionary Games, Asymmetric Information and Knowledge, Conflict Resolution and Disaster Prevention [6-9] did not destroyed our pragmatic optimism, but has taught us to be careful. Money and Knowledge are the main resources of Contemporaneity, Law and Ethics – Bottom line Frameworks of Progress. "As we struggle to come to terms with vulnerability and fear, pointing to a need for moral reflection and logical evaluation that tend to be rare in times of crisis." [14] Following [32], we try to orange institutionalization of individual and collecting knowledge about terrorism and counterterrorism and transfer the knowledge between individuals, groups and organizations. We do it using the common principles of Trust Engineering and Risk Management and separating functions/aspects in context of bipolar dimensions of Internal/External, Actual/Future, Explicit/Implicit and Experimental/Theoretical Knowledge. Steven R. Newcomb made a huge job by rapprochement and attachment of net models and technologies of data, information and expert knowledge processing, particularly for terrorism patterns recognition and terrorism activities forecasting [19]. Gordon Woo is making first steps from the mathematics of natural catastrophes to modeling of artificial ones in attempt to quantify terrorism risks and justify terrorism insurance [36]. By the time of this article completion, The Associated Press correspondent Gene Johnson reported [15] about five day large-scale counterterrorism exercises in Seattle (an imaginary "dirty bomb") and Chicago (fake threat of a biological agent). The exercises involve more than 8,500 people from 100 federal, state and local agencies, and it cost was estimated $16 million dollars. Hundreds of evaluators are watching the exercises.

ILLEGAL MIGRATION AND IDENTIFICATION FRAUD. "Those who enter Japan illegally cannot take up regular employment, and often get associated with Japan criminal organizations to become criminal elements." [11] USA Census Bureau estimates (March 2002) the population of foreign-born residents in country 32.5 million or 11.5 percent of the 282.1 million common population. Illegal alien population in the USA, by estimation of Census Bureau (2002), is shortly close to 9 million. The General Accounting Office (GAO) concluded that immigration fraud is rampant and the Immigration and Naturalization Service (INS) has no idea how to get it under control. The agency's lax bureaucratic practices have even helped open the door for terrorism. In a report released 15.02.02, GAO came to a conclusion that immigration benefit fraud is "pervasive and significant and will increase as smugglers and other criminal enterprises use fraud as another means of bringing illegal aliens, including criminal aliens, into the country." INS fraud falls into two categories: using fake document, and lying on an application for a green card or U.S. citizenship. When perpetrators of fraud are caught, little is done to them. The usual penalty for immigration fraud is a denial of benefits, not criminal prosecution." [5]

CONCLUSION.

"In the short term, a military approach to terrorism may protect us, but in the long term, we need to find solutions by pursuing education, development, dialog, negotiation, and law. In such contexts, we can only be assisted by an appreciation of values and value differences, and the limitation of violence as a means of conflict resolution." [14] We agree to this. Our future research and development oriented to combine mathematical and engineering tools for conflict resolution and disaster prevention, for safety of citizens.

REFERENCES

1. A Handbook for Measuring Employee Performance. – Workforce Compensation and Performance Service. Washington, DC: U.S. Government Printing Office, 2001. – http://www.opm.gov/perform/articles/1999/pdf10.asp
2. Balmin, Lev, and Igor Safonov. Optimal Scheduling of Complex R&D Programs. – Vladivostok: Far-East Polytechnic University, 1985.
3. Brinkhoff, Th. The Principal Agglomerations of the World. – http://www.citypopulation.de -- 12.11.2002

4. Capital Wireless Integrated Network. Strategic Plan 2001. May 9, 2001. – HTTP://www.capwinproject.com
5. D'Agostino, Joseph. GAO: INS Bungling Facilitates Fraud, Terrorism. February 22, 2002. – http://www.humaneventsonline.com/articles/02-25-02/dagostino.htm
6. Degtiar, Vladimir, and Igor Safonov. Evolutionary Mechanism of Conflict Resolution. – New York, NY: Plenum Publishing Corporation, No. 2401-0114, 1988.
7. Degtiar, Vladimir, and Igor Safonov. Behavioristic and Ethical Aspects for Computerized Collective Decision Support Systems. – Moscow: 1989.
8. Degtiar, Vladimir, and Igor Safonov. An Evolution-Stable Conflict-Reducing Mechanism with Side Payments. – New York, NY: Plenum Publishing Corporation, No. 2602-0297, 1990.
9. Degtiar, Vladimir, and Igor Safonov. Distributed System for Decision Ethics Support. – Proceedings of the Fifth All-Union Seminar "Synthesis Technique and Development Planning for Large-Scale System Structures". – Moscow: NKAU USSR, 1990.
10. District of Columbia. District Response Plan. April 4, 2002. – http://dcema.dc.gov/info/pdf/basic.pdf
11. Ensuring urban security. – http://www.chijihonbu.metro.tokyo.jp
12. FEMA. The Federal Response Plan. April 1999. – http://www.app1.fema.gov/fema/fed1.htm
13. Galford, Robert, and Anne Seibold Drapeau. The Trusted Leader. – New York, NY: The Free Press, 2002.
14. Govier, Trudy. A Delicate Balance. – Boulder, CO: Westview Press, 2002.
15. Johnson, Gene. Terror Drills in Seattle, Chicago. – The Associated Press. – 12.05.2003.
16. Karas, Viacheslav, and Igor Safonov. Organization of Research and Development for Concurrent Projects. - Proceedings of the Ninth All-Union Symposium "Logical Control in the Industry". – Tashkent: MISIS, 1986.
17. Klein, Gary. Sources of Power. How People Make Decisions. – Cambridge, MA: The MIT Press, 1998.
18. Maslow, A. Motivation and Personality. – New York, NY: Harper & Brothers, 1954.
19. Newcomb, Steven. Forecasting Terrorism: Meeting the Scaling Requirements. – Extreme Markup Languages 2002, Montreal, August 2002. – http://www.coolheads.com/SRNPUBS/extreme2002/forecasting-terrorism.hml
20. Outline of the Ordinance. – http://www.gity.kobe.jp/cityoffice/15/092/Jorei/Outline.htm
21. Principles of Management. Lesson 11: Personnel Evaluation. – University of Wisconsin Disaster Management Center, 2002. – http://dmc.engr.wisc.edu/courses/principles/AA04-11.html
22. Ruble, Blair. Second Metropolis. Pragmatic Pluralism in Gilded Age Chicago, Silver Age Moscow, and Meiji Osaka. – Cambridge, UK: Woodrow Wilson Center and Cambridge University Press, 2001.
23. Safonov, Igor. Methods and Systems of Forecasting Using Expert Appraisals. – Kiev: Znanie, 1973.
24. Safonov, Igor. Optimization of the Structured-Algorithmic Systems. – Kiev: Znanie, 1978.
25. Safonov, Igor. SELENA - the Expert-Modeling System Presenting Symbiosis of DSS and CAD. - Proceedings of the Regional Conference "Mathematical and Programming Techniques for MIS and CAM Design". – Penza: Polytechnic University, 1986.
26. Safonov, Igor. Information Security and Information Terrorism. - Capital and Vicinity, (6), 1996.
27. Safonov, Igor. Trust Engineering and Risk Management of Complex Systems. – Proceedings of the International Scientific School "Modeling and Analysis of Safety, Risk and Quality in Complex Systems." – Saint-Petersburg, Russia: Russian Academy of Sciences, 2001.
28. Safonov, Igor, Sergey Poroshin, and Nikolay Yukhin. Expert-Modeling System for Economic Analysis and Development Optimization of Production. – Proceedings of the All-Union Conference "Problems and Techniques of Science-Technical Progress Acceleration Using MIS", Part 1. – Moscow: VNIIPOU, 1985.
29. Safonov, Igor, and Alexander Tarko. Foresee and foreworn! – Business World, May 29, 1992.
30. Safonov, Igor, and Anna Tolmacheva. Optimization and Interaction in CAD of MIS. - IFAC Workshop "Computer-Aided Control Systems Design". – Moscow: Institute of Control Science, 1980.
31. Safonov, Igor, and Igor Tsikunov. Automation of the Situation Analysis. – Proceedings of the Conference "Analysis and Synthesis of Finite Automata." – Saratov: State University, 1973.

32. Schuppel, Jurgen, Gunter Muller, and Peter Gomes. The Knowledge Spiral. – Knowing in Firms. – London: SAGE Publications, 1998.
33. Starner, Tom. Modeling for Terrorism. – Risk & Insurance, April 1, 2003.
34. The Emergency Program Manager. – Emergency Management Institute. Federal Emergency Management Agency. – 2001.
35. What NCCI's Terrorism Modeling Demonstrates. December 10, 2002. – http://www.ncci.com/nccisearch/news/ceocorr/terrorism_model.htm
36. Woo, Gordon. Quantitative Terrorism Risk Assessment. – Risk Management Solutions Ltd. – www.rms.com/NewsPress/ Quantitative_Terrorism_Risk_Assessment.pdf

RISK ANALYSIS ON THE BASIS OF PARTIAL INFORMATION ABOUT QUANTILES

Lev V. Utkin
●
Department of Computer Science, St. Petersburg Forest Technical Academy,
St. Petersburg, Russia
e-mail: utkin@stat.uni-muenchen.de

Thomas Augustin
●
Department of Statistics, University of Munich, Germany
e-mail: thomas@stat.uni-muenchen.de

Abstract. Risk analysis under partial information about probability distributions of states of nature is studied. An efficient method is proposed for a case when initial information is elicited from experts in the form of interval quantiles of an unknown probability distribution. This method reduces a difficult to handle non-linear optimisation problem for computing the optimal action to a simple linear one. A numerical example illustrates the proposed approach.

Keywords: risk analysis, imprecise probability theory, optimisation problem, quantile, decision making, randomized strategy, expert opinions

INTRODUCTION

One of the main objectives of performing risk analyses is to support decision-making processes. Risk analysis provides a basis for comparing alternative concepts, actions or system configurations under uncertainty. A variety of methods has been developed for estimating losses and risks. When events occur frequently and when they are not very severe, it is relatively simple to estimate the risk exposure of an organization, as well as a reasonable premium when, for instance, an insurance transaction is made. Commonly used methods rely on variations of the principle of maximizing expected utility, tacitly assuming that all underlying uncertainty can adequately be described by a precise and completely known probability measure. However, when the uncertainty is complex and the quality of the estimates is poor, e.g., when evaluating low-probability, catastrophic events, the customary use of such rules together with overprecise data could be harmful as well as misleading. Therefore, it is necessary to extend the principle of maximizing expected utility to deal with complex uncertainty. This allows powerful evaluation under vague and numerically imprecise information. An efficient way for realizing such methods is the framework provided by imprecise probability theory [3,5,6].

Very often the initial data about unwanted events are elicited from experts, who are typically asked about quantiles of a random quantity (states of nature). Based on this information, and on the choice of a parameterized family of distribution functions, a fitted distribution function is chosen that represents the available information in some best way to some extent. However, as pointed out, for instance, in [2], experts better supply intervals of quantiles rather than point-values because their knowledge is not only of limited reliability, but also imprecise. Moreover, as discussed above, the choice of one particular distribution function fitted to the quantiles would lead to substantial errors in risk analysis. Therefore, new procedures for computing optimal actions under conditions of partial information about states of nature in the form of imprecise quantiles are proposed in the paper. Efficient methods for computing optimal unrandomized and randomized actions based on solving the linear optimisation problems are investigated. A numerical example illustrates the methods.

A GENERAL APPROACH TO RISK ANALYSIS AND IMPRECISE QUANTILES

Consider the basic model of decision theory: One has to choose an action from a non-empty, finite set $A=\{a_1,...,a_n\}$ of possible actions. The consequences of every action depend on the true, but unknown state of nature $t\in\Omega=\{t_1,...,t_m\}$. The corresponding outcome is evaluated by the utility function

$$u:\quad (A\times\Omega)\to\mathbf{R}$$
$$(a,t)\mapsto u(a,t)$$

and by the associated random variable u(a) on $(\Omega,Po(\Omega))$ taking the values $u(a,t)$. Alternatively a loss function $l(a,t)$ is assigned, which can be embedded into the framework proposed by setting $u(a,t)=-l(a,t)$. Often it makes sense to study randomized actions, which can be understood as a probability measure $\lambda=(\lambda_1,...,\lambda_n)$ on $(A,Po(A))$. Then $u(\cdot)$ and $\mathrm{u}(\cdot)$ are extended to randomized actions by defining $u(\lambda,t):=\sum_{s=1}^{n}u(a_s,t)\lambda_s$.

If the states of nature are produced by a perfect random mechanism (e.g. an ideal lottery), and the corresponding probability measure p on $(\Omega,Po(\Omega))$ is completely known, the Bernoulli principle is nearly unanimously favored. One chooses that action λ^* which maximizes the expected utility $\mathrm{E}_p\,\mathrm{u}(\lambda):=\sum_{j=1}^{m}u(\lambda,t_j)p(t_j)$ among all λ. Here E_p is the expectation operator with respect to the distribution p.

Suppose that information about states of nature is represented as a set of r judgements $\underline{b}_i\le\mathrm{E}_p f_i\le\overline{b}_i$, i=1,...,r, on the expectations of some random quantities $f_1,...,f_r$. This set restricts all distributions p on $(\Omega,Po(\Omega))$ by a set M such that every distribution p from M satisfies all the inequalities. An action λ^* is optimal iff for all λ, $\underline{\mathrm{E}}_M\,\mathrm{u}(\lambda^*)\ge\underline{\mathrm{E}}_M(\mathrm{u}(\lambda))$. Here $\underline{\mathrm{E}}_M$ is the lower prevision (expectation) taken over all probability distributions p from M. Then the optimal action λ^* can be obtained by maximizing $\underline{\mathrm{E}}_M(\mathrm{u}(\lambda))$ subject to $\lambda_1+...+\lambda_n=1$. This leads to the non-linear optimisation problem:

$$\min_{p\in M}\sum_{j=1}^{m}\sum_{s=1}^{n}u(a_s,t_j)\cdot\lambda_s\cdot p(t_j)\to\max_{\lambda_s\ge 0}\qquad(1)$$

subject to $\underline{b}_i\le\mathrm{E}_p f_i\le\overline{b}_i$, i=1,…,r, $\lambda_1+...+\lambda_n=1$.

Similar expressions can be written in a case of the continuous set of states of nature $\Omega=[A,B]$. In this case, the expected utility is $\mathrm{E}_p\,\mathrm{u}(\lambda):=\int_A^B u(\lambda,t)p(t)dt$. Here $p(t)$ is a density function which is consistent with the set of initial judgements about states of nature. In the paper, we will consider the continuous set of states of nature.

In the probabilistic approach, experts are typically asked about quantiles of a random variable X defined on a continuous sample space Ω. The smallest number $t\in\Omega$, such that $\Pr\{X\le t\}=k/100$, is called the k% quantile and denoted qk%. In this approach, the experts are often asked to supply the 5%, 50% and 95% quantiles. In other words, an expert supplies t_1,t_2,t_3 such that $\Pr\{X\le t_1\}=0.05$, $\Pr\{X\le t_2\}=0.5$, $\Pr\{X\le t_3\}=0.95$, respectively. Generally, if r experts provide their judgements about q_i quantiles, i=1,...,r, of an unknown cumulative probability distribution of the continuous random variable X, this information can be represented as $\Pr\{X\le t_i\}=q_i$, i=1,...,r. In terms of the imprecise probability theory, q_i can be viewed as identical lower and upper previsions (expectations) of the gamble $I_{[0,t_i]}(X)$, i.e., $\underline{\mathrm{E}}I_{[0,t_i]}(X)=\overline{\mathrm{E}}I_{[0,t_i]}(X)$. Here $I_{[0,t_i]}(X)$ is the indicator function taking the value 1 if $X\in[0,t_i]$ and 0 if $X\notin[0,t_i]$. However, judgements elicited from experts are usually imprecise and unreliable due to the limited precision of human assessments. In other words, experts provide some intervals of quantiles in the form $X_i = [\underline{t}_i, \overline{t}_i]$. This can be formally written as

$$\Pr\{X \le [\underline{t}_i, \overline{t}_i]\} = q_i,\ i = 1,...,r \qquad (2)$$

Every interval X_i produces a set of probability distributions such that the lower distribution contains the point $q_i(\overline{t}_i)$ and the upper one contains the point $q_i(\underline{t}_i)$.

Decision making with imprecise quantiles

Let us define what $\underline{\mathrm{E}}_M\,\mathrm{u}(\lambda)$ means in the case when initial information about *p* is given in the form of quantile intervals. Suppose that we knew precise values of q_i quantiles t_i, *i*=1,...,*r*. Denote $T = (t_1,...,t_r)$ and the set of possible vectors *T* by {*T*}. Let $\underline{\mathrm{E}}_M(\mathrm{u}(\lambda) \mid T)$ be the lower expectation of the function $u(\lambda)$ under condition of precise values *T* of quantiles. Since at least one of the points t_k belonging to the interval $X_i = [\underline{t}_i, \overline{t}_i]$ is a true value of the corresponding quantile, then there holds

$$\underline{\mathrm{E}}_M\,\mathrm{u}(\lambda) = \min_{\forall t \in X_i, i=1,...,r} \underline{\mathrm{E}}_M(\mathrm{u}(\lambda) \mid T).$$

By using the natural extension [3,4,5] for computing the lower prevision $\underline{\mathrm{E}}_M(\mathrm{u}(\lambda) \mid T)$, we get the following linear programming problem:

$$\underline{\mathrm{E}}_M(\mathrm{u}(\lambda) \mid T) = \max_{c, w_i}\left(c + \sum\nolimits_{i=1}^{r} w_i q_i\right) \qquad (3)$$

subject to $w_i, c \in \mathbf{R}, i = 1,...,r,$ and $c + \sum\nolimits_{i=1}^{r} w_i I_{[0,t_i]}(t) \le u(\lambda, t), \forall t \in \Omega$.

UNRANDOMIZED STRATEGY

The unrandomized strategy supposes that $\lambda = (0,...,0,\lambda_s,0,...,0)$, $\lambda_s = 1$. Let us consider how to find the value *s* corresponding to the optimal action.

Proposition 1. Suppose that $q_1 \le q_2 \le ... \le q_r$ and $\lambda = (0,...,0,\lambda_s,0,...,0)$, $\lambda_s = 1$. Denote q_0=0, q_{r+1}=1, t_0=*A*, t_{r+1}=*B*. Then the solution to problem (3) exists if (i) $t_1 \le t_2 \le ... \le t_r$, (ii) $t_i < t_{i+1}$ for $q_i < q_{i+1}$, *i*=1,...,*r*, and this solution is

$$\underline{\mathrm{E}}_M(\mathrm{u}(\lambda) \mid T) = \sum\nolimits_{i=0}^{r} (q_{i+1} - q_i) \min_{t \in [t_i, t_{i+1}]} u(a_s, t).$$

Let us consider an approximate solution of the decision making problem in the case of interval quantiles. Let us divide the sample space Ω into *N* intervals by points $A = \tau_0, \tau_1, ..., \tau_{N-1}, \tau_N = B$. Then the set {*T*} becomes finite and contains vectors of the form $(\tau_{l(0)}, \tau_{l(1)}, ..., \tau_{l(r)})$ such that $\tau_{l(i)} \in [\underline{t}_i, \overline{t}_i]$, i.e., *l*(*i*) is an index of a point belonging to X_i.

Proposition 2. Suppose $t_i \in [\underline{t}_i, \overline{t}_i]$, *i*=1,...,*r*. If there exist such *i* and *j* that $\underline{t}_i > \overline{t}_j$ and $q_i < q_j$, then judgements are conflicting, otherwise the optimal action is

$$a_s = \arg\max_s \min_{s=1,...,n} \min_{T \in \{T\}} \underline{\mathrm{E}}_M\,\mathrm{u} \cong \arg\max_s \min_{s=1,...,n\ (\tau_{l(0)}, \tau_{l(1)}, ..., \tau_{l(r)})} \sum\nolimits_{i=0}^{r} (q_{i+1} - q_i) \min_{t \in [\tau_{l(i)}, \tau_{l(i+1)}]} u(a_s, t).$$

In particular, if all utility functions $u(a_s,t)$, s=1,…,n, are decreasing as t is increasing, then $a_{\text{opt}} = \arg\max\limits_{s=1,\dots,n}\left(\sum_{i=0}^{r}(q_{i+1}-q_i)u(a_s,\overline{t}_{i+1})\right)$, if $u(a_s,t)$, s=1,…,n, are increasing, then $a_{\text{opt}} = \arg\max\limits_{s=1,\dots,n}\left(\sum_{i=0}^{r}(q_{i+1}-q_i)u(a_s,\underline{t}_i)\right)$.

RANDOMIZED STRATEGY

The technique proposed in the previous sections leads to a series of non-linear optimisation problems in the case of the randomized strategy. Therefore, it is necessary to consider a different method for computing λ. Here the modification of an approach proposed by Augustin [1] based on using sets of extreme points is applied. The optimisation problem for computing the optimal randomised action is

$$\min_{p\in M}\int_A^B\left(\sum_{s=1}^{n}u(a_s,t)\lambda_s\right)p(t)dt \to \max_{\lambda_s\geq 0}, \qquad (4)$$

subject to $\lambda_1+\dots+\lambda_n=1$ and $\Pr\{X\leq[\underline{t}_i,\overline{t}_i]\}=q_i$, $i=1,\dots,r$.

Let us introduce the variable

$$G=\min_{p\in M}\int_A^B\left(\sum_{s=1}^{n}u(a_s,t)\lambda_s\right)p(t)dt$$

and consider the sense of (2). If to call the expectation $\underline{\mathrm{E}}_M(\mathrm{u}(\lambda)\,|\,T)$ and the set of constraints $\Pr\{X\leq t_i\}=q_i$, i=1,...,r, for every fixed T by an imprecise model, then (2) corresponds to the union of a set of imprecise models taken over all possible vectors T, i.e., the set M of distributions p restricted by constraints (1) is the union of sets M_T. According to [3], a set of extreme points of the united model is the union of extreme points of the imprecise models corresponding to vectors T, i.e., $extr(M)=\cup_{T\in\{T\}}extr(M_T)$. This implies that a set of problems (4) can be reduced to the problem:

$$\max_{\lambda_s\in\mathbf{R}_+,G\in\mathbf{R}} G$$

subject to

$$G\leq\int_A^B\left(\sum_{s=1}^{n}u(a_s,t)\lambda_s\right)p(t)dt,\ p\in\bigcup_{T\in\{T\}}extr(M_T),\ \sum_{s=1}^{n}\lambda_s=1. \qquad (5)$$

Now we have to find the extreme points for each $T\in\{T\}$. Let us rewrite the available information about quantiles corresponding to T in the following form:

$$\int_A^{t_1}p(t)dt=q_1,\ \int_{t_1}^{t_2}p(t)dt=q_2-q_1,\dots,\int_{t_r}^{B}p(t)dt=1-q_r\ .$$

All equalities can be considered independently in the sense that they do not have common variables. If we approximately represent the integrals as sums, then the i-th hyperplane produced by the i-th equality has the following extreme points:

$$(q_i-q_{i-1},0,\dots,0),\ (0,q_i-q_{i-1},\dots,0),\dots,(0,0,\dots,q_i-q_{i-1}).$$

Hence the set M_T has the extreme points of the form:

$$p(t)=\sum_{i=0}^{r}(q_{i+1}-q_i)\delta(t-\tau_i),\ \tau_i\in[t_i,t_{i+1}],$$

where $\delta(t-\tau_i)$ is the Dirac function which has unit area concentrated in the immediate vicinity of the point τ_i; $t_0=A$, $t_{r+1}=B$, $q_0=0$, $q_{r+1}=1$.

After substituting these extreme points into constraints (5), we get

$$G \le \sum_{s=1}^{n} \lambda_s \sum_{i=0}^{r} \int_{t_i}^{t_{i+1}} u(a_s,t)p(t)dt, \forall p \in extr(M_T).$$

If we take one set of extreme points by fixed T, then there holds

$$G \le \sum_{s=1}^{n} \lambda_s \sum_{i=0}^{r} u(a_s,\tau_i)(q_{i+1}-q_i), \forall \tau_i \in [t_i,t_{i+1}]. \qquad (6)$$

Let us consider an approximate solution of the decision making problem in the case of interval quantiles. By dividing the sample space Ω into N intervals by points $A=\tau_0,\tau_1,...,\tau_{N-1},\tau_N=B$ (see the section “Unrandomized strategy”), we get a finite set of constraints

$$G \le \sum_{s=1}^{n} \lambda_s \sum_{i=0}^{r} u(a_s,\tau_i)(q_{i+1}-q_i), \forall \tau_i \in [\tau_{l(i)},\tau_{l(i+1)}],\ \forall l(i) \qquad (7)$$

Proposition 3. Suppose $t_i \in [\underline{t}_i,\overline{t}_i]$, i=1,...,r. If there exist such i and j that $\underline{t}_i > \overline{t}_j$ and $q_i < q_j$, then judgements are conflicting, otherwise the optimal randomized action is approximately defined by solving the following linear programming problem:

$$\max_{\lambda_s \in \mathbf{R}_+, G\in \mathbf{R}} G,$$

subject to (7) and $\sum_{s=1}^{n}\lambda_s = 1$.

Since the right side of (7) has to be as small as possible and $q_{i+1}-q_i \ge 0$, then the set of constraints is reduced to one constraint in the case of increasing or decreasing utility functions. By considering the set {T}, we can say that constraints (7) have to be satisfied for arbitrary values t_i and t_{i+1} such that $t_i \in X_i$ and $t_{i+1} \in X_{i+1}$, i=0,...,r. It is obvious that $\min_{\tau_i \in [t_i,t_{i+1}]} u(a_s,\tau_i)$ by $t_i = \underline{t}_i$ for increasing utility functions (by $t_i = \overline{t}_{i+1}$ for decreasing utility functions) is less than by any $t_i \ge \underline{t}_i$ ($t_i \le \overline{t}_i$). This implies that we remain one constraint

$$G \le \sum_{s=1}^{n} \lambda_s \sum_{i=0}^{r} (q_{i+1}-q_i)u(a_s,\underline{t}_i) \qquad (8)$$

in the case of increasing utility functions or one constraint

$$G \le \sum_{s=1}^{n} \lambda_s \sum_{i=0}^{r} (q_{i+1}-q_i)u(a_s,\overline{t}_{i+1}) \qquad (9)$$

in the case of decreasing utility functions.

Proposition 4. The linear optimisation problems with constraints (8) or (9) have the following solution:

$$G = \max_{s=1,...,n} \sum_{i=0}^{r} (q_{i+1}-q_i)u(a_s,\tau_i),$$

$$\lambda_k = \begin{cases} 1, & k = \arg_s \max_{s=1,...,n} \sum_{i=0}^{r} (q_{i+1}-q_i)u(a_s,\tau_i) \\ 0, & \quad otherwise \end{cases},$$

where $\tau_i = \underline{t}_i$ for increasing utility functions, $\tau_i = \overline{t}_{i+1}$ for decreasing utility functions.

Proposition 4 implies that the randomised optimal action for the considered decision problem is equivalent to the unrandomized one (see Proposition 2).

NUMERICAL EXAMPLE

Suppose experts provide 5%, 50%, 95% quantiles of the probability distribution of a random variable defined on the sample space $\Omega = [0,120]$. This implies r=3 and q_1=0.05, q_2=0.5, q_3=0.95. Expert judgements are given in Table 1. Suppose we have to choose one of two actions $\{a_1, a_2\}$ in accordance with utility functions

$$u(a_1,t)=\exp(-0.1t)\,,\ u(a_2,t)=0.5-0.012x\,.$$

Table 1. Interval quantiles provided by experts

5%		50%		95%	
Lower	Upper	Lower	Upper	Lower	Upper
2	4	12	15	19	19

Since the utility functions are decreasing, it follows from Proposition 4 or Proposition 2 that

$$k=\arg_s \max_{s=1,\dots,n} \sum\nolimits_{i=0}^{r} (q_{i+1}-q_i)u(a_s,\bar{t}_{i+1})\,.$$

If s=1, then we get the lower expected utility

$$\begin{aligned}&(q_1-q_0)u(a_1,\bar{t}_1)+(q_2-q_1)u(a_1,\bar{t}_2)+(q_3-q_2)u(a_1,\bar{t}_3)+(q_4-q_3)u(a_1,\bar{t}_4)\\&=(0.05-0)\exp(-0.1\times 4)+(0.5-0.05)\exp(-0.1\times 15)\\&+(0.95-0.5)\exp(-0.1\times 19)+(1-0.95)\exp(-0.1\times 20)=0.208.\end{aligned}$$

If s=2, then the lower expected utility is

$$\begin{aligned}&(q_1-q_0)u(a_2,\bar{t}_1)+(q_2-q_1)u(a_2,\bar{t}_2)+(q_3-q_2)u(a_2,\bar{t}_3)+(q_4-q_3)u(a_2,\bar{t}_4)\\&=(0.05-0)(0.5-0.012\times 4)+(0.5-0.05)(0.5-0.012\times 15)\\&+(0.95-0.5)(0.5-0.012\times 19)+(1-0.95)(0.5-0.012\times 20)=0.302.\end{aligned}$$

The above numerical results imply that the optimal action is a_2.

SOME REMARKS ABOUT DISCRETE STATES OF NATURE

If the set of states of nature is discrete, $\Omega=\{t_1,\dots,t_m\}$, then information about interval quantiles can be represented as

$$q_i \le \Pr\{X \le [\underline{t}_i,\bar{t}_i]\} \le q_{i+1},\ 1-q_i \le \Pr\{X \ge [\underline{t}_i,\bar{t}_i]\} \le 1-q_{i-1}, i=1,\dots,r\,.$$

In this case, the linear programming problem for computing $\underline{\mathrm{E}}_M(\mathrm{u}(\lambda)\,|\,T)$ is of the form:

$$\underline{\mathrm{E}}_M(\mathrm{u}(\lambda)\,|\,T)=\max_{c,c_i,d_iw_i,v_i}\left(c+\sum\nolimits_{i=1}^{r}(c_iq_i-d_iq_{i+1}+w_i(1-q_i)-v_i(1-q_{i-1})\right)$$

subject to $c_i,d_iw_i,v_i \in \mathbf{R}_+, c\in\mathbf{R},\ i=1,\dots,r,$ and

$$c+\sum\nolimits_{i=1}^{r}\left((c_i-d_i)I_{[0,t_i]}(t)+(w_i-v_i)I_{[t_i,m]}(t)\right)\le u(\lambda,t), \forall t\in\Omega\,.$$

Generally, it is difficult to find any solution to the above problem in the explicit form. However, this problem can be numerically solved for every $T\in\{T\}$, and the optimal action is computed by maximizing $\min_{T\in\{T\}}\underline{\mathrm{E}}_M(\mathrm{u}(\lambda)\,|\,T)$ over all possible actions.

CONCLUSION

Computationally simple algorithms have been obtained for calculating optimal actions under partial information about quantiles. It is worth noticing that we have focused in this paper on the basic decision problem. However, the ideas of this paper should be also applicable to more complex decision problems, for example, multi-criteria decision making, or the case where additional sample information is available.

REFERENCE

1. Augustin, Th. Expected utility within a generalized concept of probability – a comprehensive framework for decision making under ambiguity. *Statistical Papers*, 43:5-22, 2002.
2. Dubois, D. and Kalfsbeek, H. Elicitation, assessment and pooling of expert judgement using possibility theory. In C.N. Manikopoulos, editor, *Proc. of the 8th Inter. Congress of Cybernetics and Systems*, pages 360-367, Newark, NJ, 1990. New Jersey Institute of Technology Press.
3. Kuznetsov, V.P. *Interval Statistical Models,* Radio and Communication, Moscow (1991). (in Russian)
4. Utkin, L.V. and Gurov, S.V. New reliability models based on imprecise probabilities. Chapter 6. Edited book: *Advanced Signal Processing Technology by Soft Computing*. World Scientific, 110-139, 2001.
5. Walley, P. *Statistical Reasoning with Imprecise Probabilities*, Chapman and Hall, London (1991).
6. Weichselberger, K. *Elementare Grundbegriffe einer allgemeineren Wahrscheinlichkeitsrechnung*, volume I, Intervallwahrscheinlichkeit als umfassendes Konzept. Physika, Heidelberg, 2001.

RISK ANALYSIS ON THE BASIS OF JUDGMENTS SUPPLIED BY UNKNOWN EXPERTS

Lev V. Utkin, Sergey P. Shaburov
●
"Tehtrans" JSC
Fontanka River Quay, 117, St.Petersburg, 190031, Russia
e-mail: lvu@techrw.spb.su, spsh@techrw.spb.su

The development of a system requires fulfilling the available standards of reliability and safety. Due to possible complexity of the system, its parameters often are determined by experts whose judgements are usually imprecise and unreliable due to the limited precision of human assessments. Therefore, an approach for computing probabilities of expert judgments and for analysing the risk of decision about satisfying the parameters to standards of reliability and safety is proposed in the paper. A numerical example considering a microprocessor system of central train control illustrates the proposed approach.

Keywords: expert judgments, imprecise probabilities, multinomial model, Dirichlet distribution, risk analysis, reliability and safety.

INTRODUCTION

The development of a system requires fulfilling the available standards of reliability and safety. Due to possible complexity of a system, it is difficult to precisely assess the system parameters characterizing its reliability and safety. Therefore, very often these parameters are determined by experts. Judgments elicited from human experts may be a very important part of information about systems on which limited experimental observations are possible. At the same time, they are usually imprecise and unreliable due to the limited precision of human assessments. When several experts supply judgments or assessments about a system, their responses are pooled so as to derive a single measure of the system behaviour. Judgments of reliable experts should be more important than those of unreliable ones. Various methods of the pooling of assessments, taking into account the quality of experts, are available in the literature [1]. These methods use the concept of precise probabilities for modelling the uncertainty and the quality of experts is modelled by means of *weights* assigned to every expert in accordance to some rules. It should be noted that most of these rules use some available information about correctness of previous expert opinions. This way might meet several difficulties. First, the behaviour of experts is unstable, i.e., exact judgments related to a system elicited from an expert do not mean that this expert will provide results of the same quality for new systems. Second, when experts provide imprecise values of an evaluated quantity, the weighted rules can lead to controversial results. For instance, if an expert with a small weight, say 0.1, provides a very large interval, say [0,10], for a quantity (covering its sample space), it is obvious that this expert is too cautious and the interval he supplies is non-informative, although this interval covers a true value of the quantity. On the other hand, if an expert with a large weight, say 0.9, supplies a very narrow interval, say [5,5.01], the probability that true value of the quantity lies in this interval is rather small. We can see that the values of weights contradict with the probabilities of provided intervals. It should be noted that sometimes we do not know anything about quality of experts or assignment of weights meets some ethical difficulties. This implies that weights of experts as measures of their quality can not be measures of the quality of provided opinions.

How in this case to compare the assessed system parameter with the available standards of reliability and safety? How to compute the risk of decision making after this comparison?

The main aim of the paper is to develop an approach for computing probabilities of expert judgments and to provide a tool for risk analysis taking into account these probabilities. At that these probabilities are not regarded as a result of the previous expert experience, but as a result of recent judgments provided by unknown experts. The experts are unknown in the sense that we have no prior information about their quality.

What are conditions for probabilities of judgements? First, they have to take into account the incompleteness of the available information and even total ignorance. Second, the probabilities have to take into account the overcautiousness of experts when they supply too large and non-informative intervals. Third, the probabilities have to take into account the overconfidence of experts when they supply intervals that are too narrow (or point-values) [3]. Fourth, the probabilities have to be simply updated after obtaining new judgments. Fifth, the probabilities are assigned not to experts, but to intervals provided by the experts. The first, second, and third conditions can be satisfied if to use *imprecise* or *interval-valued probabilities* [4,5,7]. The fourth condition is fulfilled if to assume that probabilities of intervals are governed by the Dirichlet distribution.

STATEMENT OF THE PROBLEM AND THE BASIC IDEA FOR ITS SOLUTION

Suppose that *N* experts assess a parameter *u* of a system defined on $U=\{u_1,...,u_L\}$. They supply a set of intervals $\{A_1,A_2,...,A_N\}$ of *u* such that every interval $A_i \subseteq U$ contains elements from *U* with indices J_i, i.e., $A_i=\{u_j \ : \ j \in J_i\}$. At that, the number of elements in J_i is l_i. Let u_0 be a value of the standard safety. Our aim is to find probability that *u* is smaller than u_0, i.e., $\Pr\{u \leq u_0\}$.

Suppose that the set $\{A_1,...,A_N\}$ has identical elements such that there are $c_1, c_2,...,c_n$ identical intervals. Here *n* is a number of different intervals. Then $N=\sum_{i=1}^{n} c_i$. Let us calculate possible numbers of occurrences of every element of *U*. Associate the set A_i with an oblong box of size l_i with one open side and the set *U* with *L* small empty boxes of size 1. The *i*-th oblong box contains c_i balls which can move inside the box and we do not know location of balls in the *i*-th box because its open side is behind. Then we cover small boxes by the *i*-th oblong box and c_i balls enter in l_i small boxes with numbers from a set J_i. We do not know exact location of balls, but we know that they are in boxes with numbers from J_i. The same procedure is repeated *n* times. What can we say about possible numbers of balls in the small boxes now? It is obvious that there exist different combinations of numbers of balls except the case when $l_i=1$ for $i=1,...,n$, i.e., all sets A_i consist of one element. Suppose that the number of the possible combinations is *M*. Denote the *k*-th possible vector of balls by $\mathbf{n}^{(k)}=(n_1^{(k)},...,n_L^{(k)})$, $k=1,...,M$. If to assume that the sets A_i occurred independently and a ball in the *i*-th small box has some unknown probability π_i, then every combination of balls in small boxes produces the *standard multinomial model*. *M* possible combinations of balls produce *M* equivalent standard multinomial models. The models are equivalent in the sense that we can not choose one of them as a more preferable case.

For every model, the probability of an arbitrary event $A \subseteq U$ depends on $\mathbf{n}^{(k)}$, that is, we can find $P(A \mid \mathbf{n}^{(k)})$. So far as all the models are equivalent, even by precise probabilities of all categories only lower and upper probabilities of *A* can be computed

$$\underline{P}(A)=\min_{k=1,...,M} P(A \mid \mathbf{n}^{(k)}),\ \overline{P}(A)=\max_{k=1,...,M} P(A \mid \mathbf{n}^{(k)}).$$

In particular, if all sets A_i consist of single elements, that is, all oblong boxes are of size 1, then *M*=1 and

$$\underline{P}(A)=P(A \mid \mathbf{n}^{(k)}),\ \overline{P}(A)=P(A \mid \mathbf{n}^{(k)}).$$

The following problem is to define $\mathbf{n}^{(k)}$ and $P(A \mid \mathbf{n}^{(k)})$. In the case of multinomial samples, the Dirichlet distribution is the traditional choice.

Remark 1. It is worth noticing that the Dirichlet distribution should be regarded as one of the possible multinomial models that can be applied here.

Remark 2. Even if experts provide only characteristics of separate components of the system, their use leads to calculation of system parameters which also can be regarded as expert judgements (functions of expert judgements).

Remark 3. If *U* is some interval of real numbers, then we can always transform this universal set to a set with finite numbers of elements. Suppose that we have to find probabilities of an event *A*. Denote $A_{n+1} = A$. Let $\{\mathbf{i}\} = \{(i_1,...,i_n,i_{n+1})\}$ be a set of all binary vectors consisting of *n*+1 components such that $i_j \in \{0,1\}$. For every vector $\mathbf{i}$, we determine the interval B_k ($k = 1,...,2^{n+1}$) as follows:

$$B_k = \left(\bigcap_{j\ :\ i_j = 1} A_j \right) \bigcap \left(\bigcap_{j\ :\ i_j = 0} A_j^c \right),\ i_j \in \mathbf{i}.$$

As a result, we obtain a set of non-intersecting intervals B_k such that $B_1 \cup ... \cup B_{2^{n+1}} = U$. Moreover, all intervals A_i can be represented as the union of a finite number of intervals B_k. This implies that every interval B_k corresponds to an element u_k of the transformed universal set U^* with the finite number (2^{n+1}) of elements.

IMPRECISE DIRICHLET MODEL

The *Dirichlet* (*s*,α) *prior distribution* for π, where $\alpha = (\alpha_1,...,\alpha_L)$, has probability density function

$$p(\pi) = C(s,\alpha) \cdot \prod_{j=1}^{L} \pi_j^{s\alpha_j - 1},\ s > 0,\ \alpha \in S(1,L),$$

where *S*(1,*L*) denotes the interior of the unit simplex, the proportionality constant *C* is determined by the fact that the integral of $p(\pi)$ over the simplex of possible values of π is 1, α_i is the mean of π_i under the Dirichlet prior and *s* determines the influence of the prior distribution on posterior probabilities.

Walley [6] pointed out several reasons for using a set of Dirichlet distributions to model prior ignorance about probabilities π:

1) Dirichlet prior distributions are mathematically tractable because they generate Dirichlet posterior distributions;

2) sets of Dirichlet distributions are very rich, because they produce the same inferences as their convex hulls and any prior distribution can be approximated by a finite mixture of Dirichlet distributions;

3) the most common Bayesian models for prior ignorance about probabilities π are Dirichlet distributions.

The *imprecise Dirichlet model* is defined by Walley [6] as the set of all Dirichlet (*s*,α) distributions such that $\alpha \in S(1,L)$. For this model, the *hyperparameter s* determines how quickly upper and lower probabilities of events converge as statistical data accumulate. Walley [6] defined *s* as a number of observations needed to reduce the imprecision (difference between upper and lower probabilities) to half its initial value. Smaller values of *s* produce faster convergence and stronger conclusions, whereas large values of *s* produce more cautious inferences. At the same time, the value of *s* must not depend on *L* or a number of observations. The detailed discussion concerning the parameter *s* and the imprecise Dirichlet model can be found in [2,6].

By returning to the multinomial models considered in the example with boxes and balls and assuming that probabilities of balls are governed by the Dirichlet distribution, we can write the lower $\underline{P}(A,s)$ and upper $\overline{P}(A,s)$ probabilities of an event *A*, whose elements have indices from a set *J*, as follows:

$$\underline{P}(A,s) = \min_{k=1,...,M} \inf_{\alpha \in S(1,L)} \frac{n^{(k)}(A) + s\alpha(A)}{N+s},\ \ \overline{P}(A,s) = \max_{k=1,...,M} \sup_{\alpha \in S(1,L)} \frac{n^{(k)}(A) + s\alpha(A)}{N+s},$$

where $\alpha(A) = \sum_{j \in J} \alpha_j,\ n^{(k)}(A) = \sum_{j \in J} n_j^{(k)}.$

ANALYSIS OF EXPERT JUDGMENTS

Now we have to find $n^{(k)}(A)$ and $\alpha(A)$. The lower and upper probabilities $\underline{P}(A,s)$ and $\overline{P}(A,s)$ can be rewritten as

$$\underline{P}(A,s)=\frac{\min_{k=1,\dots,M} n^{(k)}(A)+s\cdot\inf{}_{\alpha\in S(1,L)}\alpha(A)}{N+s},\qquad \overline{P}(A,s)=\frac{\max_{k=1,\dots,M} n^{(k)}(A)+s\cdot\sup_{\alpha\in S(1,L)}\alpha(A)}{N+s}.$$

Note that $\inf_{\alpha\in S(1,L)}\alpha(A)$ is achieved at $\alpha(A)=0$ and $\sup_{\alpha\in S(1,L)}\alpha(A)$ is achieved at $\alpha(A)=1$ except a case when *A*=*U*. If *A*=*U*, then $\alpha(A)=1$ for the minimum and maximum.

In order to find the minimum and maximum of $n^{(k)}(A)$ we consider three intervals A_1, A_2, A_3 such that $A_1\subseteq A$, $A_2\cap A=\varnothing$, $A_3\cap A\neq\varnothing$ and $A_3\not\subset A$. Numbers of their occurrences are c_1, c_2, c_3, respectively, and $c_1+c_2+c_3=N$. It is obvious that all balls (c_1) corresponding to the set A_1 belong to the set *A* and $n^{(k)}(A)$ can not be less than c_1. On the other hand, all balls (c_2) corresponding to the set A_2 do not belong to *A*. This implies that $n^{(k)}(A)$ can not be greater than $N-c_2$. A part of balls corresponding to A_3 may belong to *A*, but it is not necessary. Therefore, $\min n^{(k)}(A)=c_1$ and $\max n^{(k)}(A)=N-c_2$. By extending this reasoning on an arbitrary set of A_i, we get the minimal, denoted $L_1(A)$, and maximal, denoted $L_2(A)$, values of $n^{(k)}(A)$:

$$L_1(A)=\min_{k=1,\dots,M} n^{(k)}(A)=\sum_{i\ :\ A_i\subseteq A} c_i,\qquad L_2(A)=\max_{k=1,\dots,M} n^{(k)}(A)=N-\sum_{i\ :\ A_i\cap A=\varnothing} c_i=\sum_{i\ :\ A_i\cap A\neq\varnothing} c_i.$$

Then there hold

$$\underline{P}(A,s)=\frac{L_1(A)}{N+s},\ \overline{P}(A,s)=\frac{L_2(A)+s}{N+s}.$$

SPECIAL CASES

Let us consider some important special cases.

1) *Coinciding judgments.* Suppose that there are *N* coinciding judgments, i.e., $A_1=\dots=A_N=[\underline{a},\overline{a}]$. This means that all experts have the same opinions about unknown value of *u*. Then we can write $\underline{P}(A_i)=N/(N+s)$, $\overline{P}(A_i)=1$. If $N\to\infty$, then there holds $\underline{P}(A_i)=\overline{P}(A_i)=1$. This property supports the intuitive sense. Indeed, if we have many identical judgments, we begin to think that these judgments are true even if we do not know anything about each expert. If $N=1$, then there hold $\underline{P}(A_i)=1/(1+s)$, $\overline{P}(A_i)=1$. The result corresponding to the case $N=1$ shows that the precise Dirichlet model ($s=0$) gives lower and upper probabilities of events 1. This is the incorrect conclusion. How can we totally believe one judgment? This contradiction can be avoided by using the imprecise Dirichlet model ($s>0$).

2) *Conflicting judgments.* Suppose that there are two conflicting judgments $A_1=[\underline{a}_1,\overline{a}_1]$, $A_2=[\underline{a}_2,\overline{a}_2]$, $\overline{a}_1<\underline{a}_2$, i.e., $A_1\cap A_2=\varnothing$. Then there hold $\underline{P}(A_i)=1/(2+s)$, $\overline{P}(A_i)=(1+s)/(2+s)$. If there are *N* conflicting judgments, then $\underline{P}(A_i)=1/(N+s)$, $\overline{P}(A_i)=(1+s)/(N+s)$. If $N\to\infty$, then there holds $\underline{P}(A_i)=\overline{P}(A_i)=0$.

3) *Noninformative judgments (overcautiousness of experts).* Suppose that there is one judgment $A_1=[\inf U,\sup U]$. Then $\underline{P}(A_1)=\overline{P}(A_1)=1$. This implies that the overcautiousness of experts can be properly modelled. Indeed, even if we do not believe to an expert, but he (she) provides very overcautious judgments, then probabilities of these judgments have to be large though the expert is unreliable.

RISK ANALYSIS

If there is the standard value u_0 of reliability or safety, then the parameter u can be compared with u_0 by means of computing the probability distribution function at point u_0. Suppose that U is the real line restricted by some values inf$U$ and sup$U$. Then we can define lower and upper cumulative probability distribution functions of a random parameter u, about which we have data in the form of intervals A_i, $i = 1,...,n$. By using the above results and taking into account the fact that $\alpha(A) = 1$ by A=U, we get

$$\underline{F}(u_0,s) = \underline{P}(\{u \in U : u \le u_0\},s) = \begin{cases} \dfrac{1}{N+s} \displaystyle\sum_{A_i : \sup A_i \le u_0} c_i, & u_0 < \sup U \\ 1, & u_0 = \sup U \end{cases},$$

$$\overline{F}(u_0,s) = \overline{P}(\{u \in U : u \le u_0\},s) = \begin{cases} \dfrac{s}{N+s} + \dfrac{1}{N+s} \displaystyle\sum_{A_i : \inf A_i \le u_0} c_i, & u_0 > \inf U \\ 0, & u_0 = \inf U \end{cases}.$$

The above expressions are obtained by considering lower and upper probabilities of the interval $A = [\inf U, u_0]$. If it is necessary to find probabilities of the complementary event $\{u \ge u_0\}$, then the following equalities can be used:

$$\overline{P}(\{u \in U : u \ge u_0\},s) = 1 - \underline{P}(\{u \in U : u \le u_0\},s) ,$$

$$\underline{P}(\{u \in U : u \ge u_0\},s) = 1 - \overline{P}(\{u \in U : u \le u_0\},s) .$$

Example. We consider a microprocessor system of central train control “Tract” developed by “Tehtrans” JSC. According to specialized standard related to safety of such systems, the probability of a hazardous failure during 10 years must be less than 2.6×10^{-6}, i.e., u_0=2.6×10^{-6}. In order to prove the safety of the developed microprocessor system, 10 experts supply indirectly intervals for probabilities of hazardous failures of the system (see Table 1). The intervals A_i are given in the second column. Numbers of identical intervals c_i are given in the third column. Values of $L_1(i)$ and $L_2(i)$ are given in the fourth and fifth columns. If we take $s = 1$, then $\underline{P}(A_i,1)$ and $\overline{P}(A_i,1)$ are given in the sixth and seventh columns. By using the expressions for lower and upper distribution function, we find lower $\underline{P}(\{u \le 2.6 \times 10^{-6}\},1)$ and upper $\overline{P}(\{u \le 2.6 \times 10^{-6}\},1)$ probabilities that the system parameter (probability of a hazardous failure) less than 2.6×10^{-6}:

$$\underline{P}(\{u \le 2.6 \times 10^{-6}\},1) = 9/11 = 0.82 , \quad \overline{P}(\{u \le 2.6 \times 10^{-6}\},1) = 1 .$$

It is obvious that risk of decision should be calculated on the basis of the lower probability. Therefore, we get the value of risk 1-0.82=0.18. This is rather large value and, therefore, it is necessary to carry out some additional expert elicitation or to improve the system, for instance, by using additional redundancy of main components that has been done.

Table 1. Expert intervals and their probabilities

i	$A_i \times 10^{-6}$	c_i	$L_1(i)$	$L_2(i)$	$\underline{P}(A_i,1)$	$\overline{P}(A_i,1)$
1	[0.6, 1.6]	3	5	8	0.45	0.82
2	[0.1, 1.6]	1	8	10	0.73	1
3	[0.0, 3.1]	1	10	10	0.91	1
4	[0.6, 2.6]	1	6	8	0.55	0.82
5	[0.6, 1.1]	2	2	8	0.18	0.82
6	[0.1, 0.6]	2	2	4	0.18	0.45

CONCLUSION

The method for analyzing the expert judgments has been considered in the paper. On one hand, it is very simple from computational point of view. On the other hand, it does not use information about experts and takes into account imprecision of expert information and deficiency in statistical data. The application of imprecise Dirichlet model allows us to avoid possible errors of traditional statistical analysis under condition when the number of available judgments is rather small. The resulting assessments can be simply updated after obtaining new expert judgments. It is worth noticing that the method can be simply extended on a case of heterogeneous judgments when experts provide information different in kind. Moreover, the method also can be extended on a case of experts with known parameters of their quality.

REFERENCES

1. Cook, R.M. *Experts in Uncertainty. Opinion and Subjective Probability in Science.* Oxford University Press, New York, 1991.
2. Coolen F.P.A. An imprecise Dirichlet model for Bayesian analysis of failure data including right-censored observations. *Reliability Engineering and System Safety*, 56, 61-68, 1997.
3. D. Dubois and H. Kalfsbeek. Elicitation, assessment and pooling of expert judgement using possibility theory. In C.N. Manikopoulos, editor, *Proc. of the 8th Inter. Congress of Cybernetics and Systems*, pages 360-367, Newark, NJ, 1990. New Jersey Institute of Technology Press.
4. V.P. Kuznetsov. *Interval Statistical Models*. Radio and Communication, Moscow, 1991. in Russian.
5. P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.
6. P. Walley. Inferences from multinomial data: learning about a bag of marbles. *Journal of the Royal Statistical Society, Series B*, 58:3-57, 1996.
7. K. Weichselberger. *Elementare Grundbegriffe einer allgemeineren Wahrscheinlichkeitsrechnung*, vol. I Intervallwahrscheinlichkeit als umfassendes Konzept. Physika, Heidelberg, 2001.

CONTRIBUTION TO CONSEQUENCES ANALYSIS USING FUZZY PROBABILITY

David VALIS,
●
Ph.D.; University of Defence, Czech Republic;
david.valis@unob.cz

Abstract: This article deals both with dependability and risk analysis from a complex point of view. Both these fields seem to be similar in many aspects, but unfortunately no congruence in sources of basic characteristics has been reached, yet. Statistical files are often very vague in terms of monitoring dependability measures or risk factors. There is a great need to use another point of view to describe these factors. One of those measures and fragments of risk or dependability are consequences both in terms of an event occurrence and failure occurrence. By using a new approach, better interconnection between these both fields and deeper applicability would be provided. A theory of fuzzy probability could be one of these new methods that could facilitate modelling of quantitative factors.

Key words: Risk, Dependability, Consequences, Management, Complex systems

1. INTRODUCTION

Dependability and risk are concepts that similarly as many other underwent a complex historical development and even today they are interpreted in many different ways and are used in various contexts. Analyses of these characteristics have also undergone a long development.

In connection with this paper, the dependability will be understood as a certain characteristics of studied objects (products, systems) that we endeavour to affect using the analysis, prognosis, calculations, modelling, testing and other tools. In a similar way we shall interpret a risk that is also a certain feature or characteristics of object that expresses its certain potential. We also strive to influence a risk using analysis, prognosis, calculations, and modelling, testing and other tools. First, we shall focus on dependability and its brief description and after that we shall aim at risk. The main attention will be paid to consequences which are related both with risk (in terms of an event occurrence) and dependability (in case of failure which is only more precise expression of an event). The consequences recognition, estimation, assessment, analysis and evaluation are more or less dependent onto a man decision. From historical point of view we do not have very often the possibility to verify and validate the human decisions very precisely. Unfortunately we need to have a tool for both better decision making and for respective validation of our decisions. The fuzzy theory seems to be very suitable for this purpose. It works both with uncertainty and vagueness. As the consequences determination is influenced both with vagueness and uncertainty mostly using language values this method suits to this very well. We would like to speak about that more detailed.

2. DEPENDABILITY AND RISK RELATIONSHIP

To describe dependability, we shall use a valid definition as given in the Terminology Standard (ČSN IEC 50 (191)). Here, dependability is defined as follows: *Dependability is collective term used to describe availability performance and its influencing* factors*: reliability performance, maintainability performance and maintenance support*. This definition implies a fact that a capability of the object of fulfilling required functions is not usually determined by characteristics of the object itself, but also by external factors, e.g. by a level of maintenance support needed.

The term dependability is used only for general description and characteristics defined in this way cannot be generally expressed by any numerical measure. Dependability of every product is understood as an integral part of the total sum of features that influence an ability to meet established and assumed needs of the user. In general, this ability is called a quality. Its individual sub-factors, e.g. availability, reliability and maintainability can be expressed numerically with the help of concrete factors.

The term risk has also undergone a long development. Meeting risk in accomplishing certain activities and studying objects and various effects, almost always it has negative emotions on us.

For the needs of various analyses, several terms have been adopted mainly from the English speaking countries. The best known is ***Risk and Hazard***. These terms have their own content, which is not usually used in purely autonomous form, but in connection with other circumstances (areas), to which it refers. For the needs of our study and this article it can be expressed as follows:

Let it exists a certain source of risk, either tangible (environment, object, human being) or intangible (activity). This source can have both positive, but as in our case also negative impact to its surroundings (other tangible or intangible elements). The existence of this impact is not always so important. The existence of such risk (i.e. negative impact) becomes important only when its impact or importance results from an interaction, which exists between an element (individual, group, technical object or activity) and a source (environment or activity).

In this moment it is necessary to realize that risk as such *does not exist*, *if there is no interaction between the source of risk and object (element) that has a certain relationship to this source*. It is necessary to take into account that interaction can also have various forms. It may be, for example, a ***voluntary***, ***involuntary***, ***random***, ***intentional,*** *etc.* interaction. The effect of these impacts can be attributed especially to an environment, in which the object occurs during its existence. Any such impacts shall be generally called ***area of risk***.

The important and integral part of all analyses will be precise, quality and sufficient identification of just this source of risk. Without this source we can hardly deal with a risk in a qualified way.

Contrary to dependability, which nearly in all cases applies to technical objects, a risk has such characteristics that it refers to a wide area of human activities and known disciplines.

This is only a brief introduction into a dependability and risk relationship. In further text we shall focus on analysis of dependability and risk resulting from the events – accidents and failures of technology. Contrary to an analysis of dependability, we can often meet only with a purely qualitative expression at the technical objects risk analysis. This is not because we would not want more, but simply because we are able to achieve nothing more concrete.

3. POSSIBLE APPROACH IN ANALYSING

Both risk or dependability analysis can be carried out in various stages of technical life of the object studied. Some partial steps of analysis can be very similar in certain aspects. However, to achieve the basic data for quality analysis need not always be so easy.

For dependability analysis we use various partial indicators that ultimately interpret a required characteristic in a numerical form. To express these indicators, it is possible to use a relatively wide scope of mathematical procedures or expressions. It is an expression of events resulting in various effects such as failure, recovery, achieving of limit state, performance degradation, etc. To be able to describe or model these events, various tools of dependability can be used. They are mostly based on a theory of probability**.** We shall not deal with their description. Individual indicators can be achieved using tools applied in dependability in all stages of technical life of the object.

The same approach applies for a risk analysis, since occurrence of some event is always that what links risk and dependability. However, for a risk it refers to events that in accordance with an interaction have an unwanted and dangerous development for us. To better understand when individual events can occur, it is possible to analyze risks in a similar way as in analysis of dependability. In single groups (stages) it is possible to apply a great amount of deterministic or stochastic methods that are the most suitable for analysis in a given period. It refers to:

- *Inherent dependability (risk)* – is dependability (risk) „embedded" into an object during its design and production. It does not involve worsening effects of operational conditions, environment, maintenance techniques, human factor, etc.;

- *Operational dependability (risk)* - is a dependability (risk) with considering the effects of operational and other conditions;
- *Estimated (predicted) dependability (risk)* - is dependability (risk), which is a result of calculations, analysis and prognosis of dependability (risk) of the projected object. Thus, it is the result of used estimation methods, input data on dependability (risk) of elements, used calculation model of the system dependability (risk), knowledge and abilities of an analyst who carries out an estimation, etc.

Up to now we talked about the events connected with the function and operation of technical object, it is advisable to rather narrow our view. In our study, first, we shall focus on an analysis of failures and their effects. In order to explain our procedure and to unambiguously outline the way of perceiving events, we shall apply several types of scales that seem to be suitable for our goal. These scales enable attaching a word meaning to events that occur. By the events we shall understand occurrence of the failure. But for many procedures, both immediate and consequent, attachment of these verbal meanings is not always sufficient. Sometimes, we would like to work with the numerical values that we would obtain from qualified data. In accordance with the above-mentioned scales, the seriousness of event (failure) effects can be divided into several groups. These are the groups that deal with a severity of event (failure) effects and its intensity (frequency, rate) of occurrence. Table 1 shows one of possible classification of severity of effects (e.g. **Mil-Std 1309, SAE ARP 5580-FMEA, IEC 608 12**).

Tab. 1: Failure severity levels

Category	*Consequence to Personnel*
Catastrophic	***Single or multiple death(s)***
Critical	***Multiple serious injuries or severe occupational illness***
Marginal	***A single severe injury or occupational illness; and or multiple minor injuries or minor occupational illnesses***
Minor	***At most a single minor injury or minor occupational illness***

According to intensity (frequency or probability) of event (failure) occurrence that has some meaning for our analysis, several groups can be distinguished. An example of their possible classification is shown in Table 2.

Tab. 2: Evaluation of probability

Probability	*Definition*
Frequent	***Likely to be continually experienced***
Probable	***Likely to occur often***
Occasional	***Likely to occur several times***
Remote	***Likely to occur some time***
Improbable	***Unlikely, but may exceptionally occur***
Incredible	***Extremely unlikely given the assumptions recorded about the system. It can be assumed that the event will not occur.***

With respect to the fact that individual scales are described in words only, it is very important who makes the analysis of a given event. As there exist a number of classes in a group (meaning „distance" between individual expressions), it is very important how an evaluator solves the problem in question. Let us assume that an experienced specialist who will not influence it by his subjective feelings will do this analysis. However, it is obvious that a subjective influence of a human being cannot be fully excluded. Then it is possible to say that it is a certain form of an expert (subjective) analysis, which is in a great extent dependent on knowledge, maturity and sometimes perhaps on the mood of an expert. Let us suppose that a level of subjectivity is on acceptable level.

From the results achieved by adding the word meaning to the effect it is possible to consequently create a certain form of matrix figure of acceptability or dismissal of failure occurrence characterized by a given effect and frequency of occurrence. On the basis of ratio of individual elements related to the event studied, we can decide if a given solution is acceptable for this event (failure). Whether the event is acceptable for us or not can be simply identified from a graphical-matrix representation as shown in Table 3.

Tab. 3: Hazard evaluation

	Severity			
Frequency	Minor	Marginal	Critical	Catastrophic
Frequent	▨	▨	▨	▨
Probable		▨	▨	▨
Occasional		▨	▨	▨
Remote			▨	▨
Improbable				▨
Incredible				

▨	Zone of non acceptable hazard.
	Zone of acceptable hazard.

From the above-mentioned steps is evident how to follow individual events-failures on the technical object. However, it is very important for us and it is again common both for the risk analysis and for dependability analysis to adequately determine severity of effects of individual failures. The above-mentioned scale shown in Table 1 it is evident that severity can be classified in a certain manner. Using known mathematical tools, it is not such a problem for the occurrence rate (shown in Table 2). Question at issue ensues when in a case of a severity of effects we want to be sure by our decision and when we expect that our surety will have a sound basis.

Another thing, which is to be considered, is a need to take certain steps of intended analyses, during which it is necessary to work with numeric value. But how to quantify such a vague characteristics? Some sectors will express this characteristic simply by money. In technical practice it is not so simple. One of the possibilities how to address this issue is to use a theory of fuzzy probability which the authors seem very suitable for a given purpose.

4. ANALYSIS OF SEVERITY OF EFFECTS WITH USING A THEORY OF FUZZY PROBABILITY

Let us assume that any technical object in any instant of time can occur in any operational state (operational condition – state "0", failure state – state "2" or partially failure state – functionality is limited, but not lost – state "1"). A transfer between these states is subject to stochastic laws. As suitable means to depict transfers between individual operational states is use a theory of Markov processes. However, we shall not deal with a description of transfers between individual operational states. A greater attention will be paid to mathematical modelling of effects related to a transfer between individual states.

As transfers between states are connected with a number of effects, it is very important to deal with them in more detail. The most important and from the respect of the function of the object also the most critical is a transfer from an operational state into a failure state. This transfer can result in the worst effects. However, it will depend what is the mechanism of a transfer. If a transfer is caused by a scheduled downtime of the object because of the preventive maintenance, it is unpleasant matter, but better than if, for example, a transfer caused by an unexpected failure with devastating results.

To evaluate severity of effects of failures of technical objects, we decided to use fuzzy set theory**.** Since this theory uses vague terms that already appear in classification of severity of failure effects, then a decision on acceptability of failure and determination on the importance of the object on which the failure appeared. Simultaneously, it is possible using this theory to assign numerical value to the studied circumstance and thus we consider it suitable. Through this theory it is also possible to include **severities of failure effects *D*** of single objects into a fuzzy set**.** Here, we shall assume that single fuzzy sub-sets consist of **coefficients of failure effect severity.** Based on the seriousness of these effects it will be later determined to what level are the given groups indispensable. To classify the **failure effect criticality i**n relation to the inherent availability of technical object we have selected the following three criteria of influence on:

- Function - $\mathbf{D_1}$,
- Safety - $\mathbf{D_2}$,
- Recovery-related costs - $\mathbf{D_3}$.

For every of these criteria we created an ascending scale of coefficients to enable to assess a seriousness of possible effects of failure related to the individual criteria. The severity scale is determined by a set **I** with four elements $I \in \{1;2;3;4\}$, while a value of coefficient of individual effect of failure in relation to selected criteria is denoted D_i, where $i \in <1,2,3>$. The principle is that with an increasing value of coefficient increases also a severity of effect. These values serve as the basis to express a severity of failure effect ***D***. The resulting coefficient ***D*** is at the same time a coefficient of seriousness of a given object and a relation expresses it:

$$\mathbf{D = D_1 . D_2 . D_3;\ D_{min} = 1, D_{max} = 64} \qquad (1)$$

To construct a fuzzy sub-set, a "fuzzification of values" is used. Actual observed values of physical values are bounded and are expressed by means of real numbers. Therefore as a universum of fuzzy numbers that represent vague concepts related with a classification of failure effects, a suitable closed interval for every of them will be sufficient. We will reach single classes of failure effects (seriousness) by dividing the resulting coefficient ***D*** into suitable sub-intervals (see above). For practical use and graphical representation a trapezoidal fuzzy number is suitable, see Fig. 1, where **μ** expresses a function of applicability and **x** obtained fuzzy number.

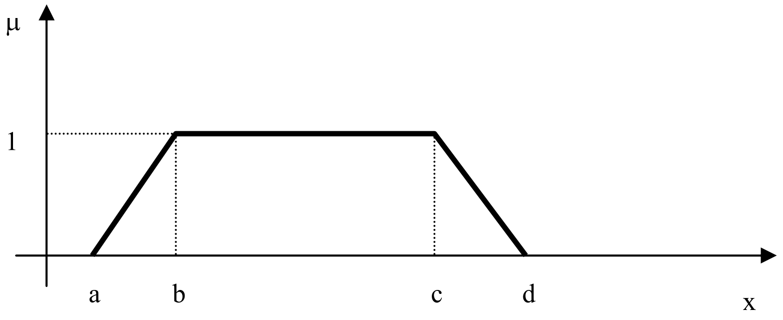


Fig. 1: *Example of trapezoidal fuzzy number*

To establish a concrete value of function of applicability for fuzzified value of selected quantity, it is sufficient to identify in which interval this value usually occurs. This interval is then a core of fuzzy number sought and we denote it as < b,c >. For our case, this core is always expressed by marginal values of single severity of failure effects (see above). Further, we determine what values the quantity does not acquire for certain. We assume a set of these values in the form $(-\infty;a) \cup (d;\infty)$, while $a<b\leq c<d$. Then an interval <a;d> is a support-set „**A**" of a fuzzy number sought.

We shall express a function of applicability of fuzzy number sought into a set „**A**" as follows:

$$\mu_A(x) = max\left(min\left(\frac{x-a}{b-a}, \frac{x-d}{c-d}, 1 \right), 0 \right). \qquad (2)$$

This expression, incl. trapezoidal shape of fuzzy number can be used for all classified severities of failure effects (see below). For the needs of technical application and based on the above mentioned scale relating to individual criteria, we divide severity of failure effects in accordance with the above-mentioned Table 1 into the following groups:

Minor	state **„0“**	fuzzy set < 1;4 >;
Marginal	state **„1“**	fuzzy set < 6;16 >;
Critical	state **„1“**	fuzzy set < 18;36 >;
Catastrophic	state **„2“**	fuzzy set < 48;64 >.

An example of fuzzy set with division of seriousness of effects is shown in Fig 2.

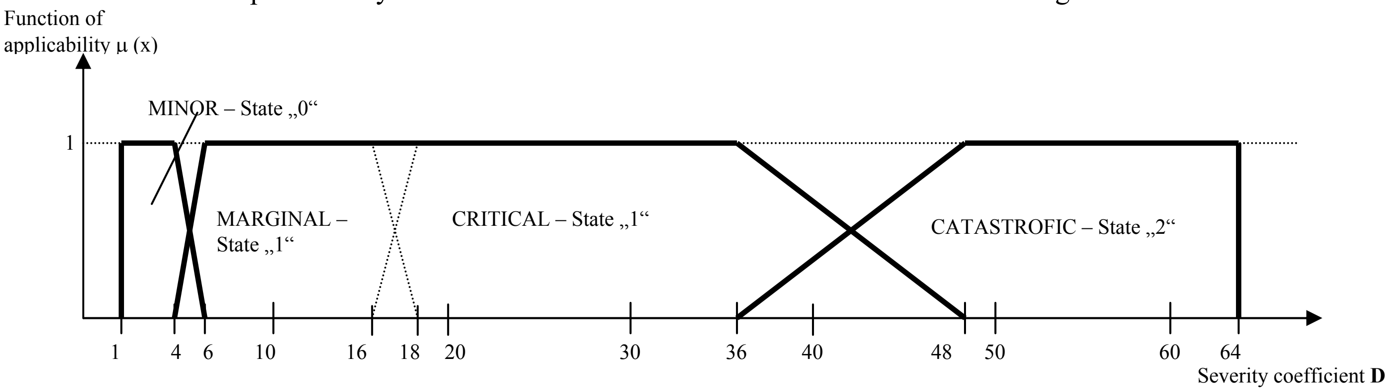


Fig. 2: *Mathematical graphical model of applicability of individual severities of failure effects into the fuzzy sets*

5. CONCLUSION

With the use of the above-mentioned procedure enabling analysing severity of failure effects of technical objects, it is possible to review any technical applications. Its use is especially suitable for those objects, the failure of which has a serious impact on the whole, society, environment, etc. These are primarily strategically and technically important objects such as, for example, power facilities, chemical and petrochemical installations, etc. This method can be also applied for military applications.

The results show not only how serious the effects of failures of given object (elements) are, but also a rate of importance of a given object (element). In addition, it can result in a mathematical model, which allows show how and in what way individual objects can transfer between its functional states. For users, these results are important primarily because they indicate beforehand weaknesses that must be paid attention to already in the design phase or that should be more focused on in the operation proper (preventive review) or preventive maintenance measures.

ACKNOWLEDGEMENT

This research was supported by the Ministry of Education, Youth and Sports of the Czech Republic, project No. 1M06059 - Advanced Technologies and Systems for Power Engineering

BIBLIOGRAPHY

[1]. Mezinárodní elektrotechnický slovník (International Electro-technical Vocabulary, Praha: Český normalizační institut 1993, ČSN IEC 50 (191).

[2]. NOVÁK, V.: Fuzzy množiny a jejich aplikace (Fuzzy Sets and Their Application), Praha: Matematický seminář SNTL 1986.

[3]. NOVÁK, V. : Základy fuzzy modelování (Fuzzy Modelling Foundations). Praha: BEN 2000.

[4]. VALIŠ, D.: Analysis of vetronics application consequences onto military battle vehicles dependability,

Brno: VA v Brně 2003, Dissertation thesis.

[5]. VALIŠ, D., FUCHS, P.: Metody analýzy a řízení risk (Risk Management Analysis Methods), Liberec: TUL 2004.

[6]. VALIŠ, D.: Fundamentals of description, perception and value determination of Risk. Liberec: Technical University 2005.

[7]. VALIŠ, D.: Fuzzy probability used for description of risk and dependability factors, In: "The 4th International Conference on Safety and Reliability", Kraków 30th May – 2nd June 2006, 9 pages (Proceedings – ISSN 1895-8281, p. 323-331).

[8]. VINTR, Z.: Specifikace požadavků na bezporuchovost technických objectů (Technical Object Reliability Requirement Specification). Brno: VA v Brně 1998, inaugural dissertation.

[9]. VINTR, Z.: Reliability and safety providing for railway applications, In: "Trans & Motauto 2005", Bulgaria, Veliko Turnovo.

CONTRIBUTION TO STOCHASTIC METHODS OF COMPLEX SYSTEMS RISK ANALYSIS

David VALIS,
●
Ph.D.; University of Defence, Czech Republic;
david.valis@unob.cz

Abstract: The paper deals with risk assessment of complex systems. As we investigate situations regarding military applications the fragments of risk management are very important for us. Risk and dependability characteristics of military battle equipment have the same importance for us as those measures which have to serve to perform battle missions itself. There is no time on the battle field to solve unpredicted and unexpected situations caused by high risk level or unreliability which might lead to loss of both equipment and crew. Due to high level of risk we face on the battlefield many systems have to be robust enough or have to be redundant to succeed.

Key words: Risk, Dependability, Management, Complex systems

1. INTRODUCTION

As we know there is number of characteristics which might be investigated and solved regarding military applications. Some of them are typically related to performance of the object although others are related to supporting characteristics. The supporting characteristics do not mean that they play second class role but usually are not preferred as much as those related to performance. In branch of our interest we talk about risk, dependability and its attributes. The common and well known dependability characteristics are often announced and used for various calculations as well as describe the item itself. We typically know these characteristics from different types of tests performed during development and testing phase. Such characteristics are related to so called inherent dependability – inherent availability. Apart of these specifications we need to know also the real behaviour in the battle field – in real deployment while completing mission. In the real deployment we talk about characteristics related to “so called” operational dependability – operational availability. These characteristics are not calculated theoretically but their calculation is based on practical and real possible situation. Such as real picture about technical item behaviour namely military battle vehicles is the most important for us. Several measures join the set of “dynamic dependability” characteristics. To be able to carry out the dynamic dependability analysis we have to know the edge conditions and our limitations for that. Dynamic in this terms means to have the information we need just in time. We may choose several possibilities for getting the time related characteristics regarding the military battle vehicle for instance. As dependability analysis serve for failures investigation we use them for getting more information about an event which in terms of risk understanding means the initial source. If we know the source of potential harm we consequently may work with the basic and well known tools for risk identification, assessment, analysis and finally evaluation. As battle vehicles are supposed to complete missions in very adverse and hostile conditions with very high level of success required and many times also in very diversified areas we have to look after the quality characteristics very well. We count among them both risk and dependability analysis which are very closely connected and their characteristics and measures serve for determination of proper picture for battle vehicle behaviour. With running time we are not happy enough with the measures and characteristics got from tests. We would like to get more precise and so called absolute (dynamic) characteristics regarding risk and dependability. That is why we have been looking for new approaches and methods suitable for this purpose. One of the most appropriate seems to be the Markov analysis. Beyond of dynamic characteristics we also need to know the potential risk level in case of unexpected event occurrence both during training phase and during real deployment while completing a mission.

If we talk about dynamic dependability and risk characteristics we take into account those events which have the major impact onto vehicle’s function – a failure. The only failures we assess are the failures from internal reasons. We do not count the possible failures caused by external reasons – in case of battle vehicles caused by hit or attack while performing a mission. In following parts we deal with all above mentioned issues.

2. RISK ON BATTLE FIELD AND ITS ASSESSMENT

In our lives we can recognise and we know plenty of circumstances which may generate existence of a risk. As we talk about a risk we subconsciously feel something wrong, negative, and unpleasant. We feel endanger or possible a hazard, endanger, jeopardy, imperilment, etc. The more we know about risk and its fractions the harder we cope with it/them. In some situations we can not do anything else than get used it. In another cases we may avoid it, reduce it or ignore it. There are many ways how to observe a risk and how to handle with it. The whole discipline dealing with risk has the name "Risk management" and its fragments have the crucial importance for us. Due to dealing with military battle vehicles we have to recognise a bit more than standard risk spectrum – risk profile we usually see regarding civilian vehicles. As the battle vehicles have to perform their mission in very difficult environment under very adverse conditions the spectrum of possible impacts is very high. We talk about sources of risk. A battle vehicle has the potential to be in collaboration with more than one source of risk both internal and external. It does not really matter if the vehicle carries out training or if it is in real deployment. Of course the real deployment may bring more consequences in case of an event occurrence. A failure in training does not need to be necessarily as crucial as in case of real mission. A failure occurrence either in training or in real mission puts the vehicle into involuntary situation which is raised due to military tasks it has to fulfil. Due to very high possibility to be immediately attacked in the battle the risk arisen is also very high. Regarding the above mentioned we use following description of risk for further work.

Let it exists a certain source of risk, either tangible (environment, object, human being) or intangible (activity). This source can have both positive, but as in our case also negative impact to its surroundings (other tangible or intangible elements). The existence of this impact is not always so important. The existence of such risk (i.e. negative impact) becomes important only when its impact or importance results from an interaction, which exists between an element (individual, group, technical object or activity) and a source (environment or activity).

In this moment it is necessary to realize that risk as such *does not exist*, *if there is no interaction between the source of risk and object (element) that has a certain relationship to this source*. It is necessary to take into account that interaction can also have various forms. It may be, for example, a ***voluntary***, ***involuntary***, ***random***, ***intentional,*** *etc.* interaction. The effect of these impacts can be attributed especially to an environment, in which the object occurs during its existence. Any such impacts shall be generally called ***area of risk***.

The important and integral part of all analyses will be precise, quality and sufficient identification of just this source of risk. Without this source we can hardly deal with a risk in a qualified way. Regarding to these facts we may understand that risk can be assessed both qualitatively and quantitatively (of course in both cases as well). Basic expressions which put risk into commonly understandable form and which enables us further dealing with risk are as follows. First and very well known (nowadays classical) description in form of an equation which may serve both for qualitative and quantitative assessment is as follows:

$$R = P \times C \tag{1}$$

Where: R – Risk;
P – Probability;
C – Consequences.

This expression allows us to carry out both qualitative and quantitative assessments. Problem is that we do not have any numerical expressions with physical unit.

Second very well known form for risk expression is following formula:

$$R = \frac{P \times C}{M} \times E \quad [unit] \tag{2}$$

Where: R – Risk;
P – Probability;
M – Measures;
E – Exposition.

This expression allows us also to carry out both qualitative and quantitative assessments. Very big advantage is that we may have physical units related to risk for further analysis.

For every element of the above mentioned equations are more or less clear procedures for their determination. We have to understand that the risk assessment as part of risk management is subdivided into two possible ways. In terms of finding solution we either talk about "Logic (sometimes determination) Access" or "Probabilistic Access". In case of probability is the situation more than clear. Although in the English speaking countries we have to distinguish between the terms "Probability" and "Likelihood" the determination is clear enough. In case of exposition we do not have to discus very much the possibility for unit and function determination. We may expect problems in terms of measures or consequences determination. Such decisions are more or less based onto expert expressions. The way is not necessarily bad but it does not give us the possibility to validate or verify a statement made.

From this point of view we recommend using new progressive forms and procedures for measures and consequences determination as well as from our historical experience. As we very often work with language and qualitative measures which are consequently somehow connected to scales (numerical expressions of qualitative expressions) we would like to be sure enough that our decision was not bad and in same circumstances under same conditions one day latter will be made in the same way. Theory of fuzzy probability and fuzzy logic seems to suit to this purpose very well. For more details how to solve such an issue see [6] or [].

From the risk assessment point of view regarding military battle equipment we may be confronted both with two known ways of stochastic distributions. We use for the random variable description distributions known as the counting and the continuous. Both of them have their importance and place both in terms of observed item and consequently risk/dependability analysis. As we want to know the so called absolute/dynamic characteristics regarding observed item we have to distinguish between both of them in the Markov´s analysis as well. The detailed description of both of them follows.

3. COUNTING DISTRIBUTION OF OBSERVED VARIABLE AND DEPENDABILITY

Based onto part describing risk assessment above we now have been looking for expression of object behaviour. Such behaviour will give us appropriate picture about real conditions of the object and will allow us to prepare possible mission scenario with such object. From mathematical point of view we may distinguish between two ways of observing object behaviour. Such as behaviour is based onto measures and characteristics used. In this part we would like to describe a possible way for dependability assessment of complex technical system which is represented by counting value in case of observed variable related to a failure. We know the basic characteristics and measures related to object. Also in this case – solving the issue related to counting variable – we use the Markov analysis for getting several characteristics of dynamic dependability. From the "good example" reasons we have chosen automatic cannon which shoots using rounds. Is a failure on a round occurs the part restoration system allows to re-charge faulty round with a new one. We talk about partial repair. The system may basically stay in two states as described bellow using scenarios for their description.

The mission is completed. In the first case there can be a situation when all the ammunition of a certain amount which is placed in an ammunition belt is used up and a round failure occurs or it is used up and a round failure does not occur. In this case a backup system of pyrotechnic cartridges is able to reverse a system into an operational state. Using up can be single, successive in small bursts with breaks between different bursts, or it might be mass using one burst. Shooting is failure free or there is a round failure occurrence n. In case a round failure occurs, a system which restores a function of pyrotechnic cartridges is initiated.

There are two scenarios too – a system restoring a pyrotechnic cartridges function is failure free, or a pyrotechnic cartridge fails.

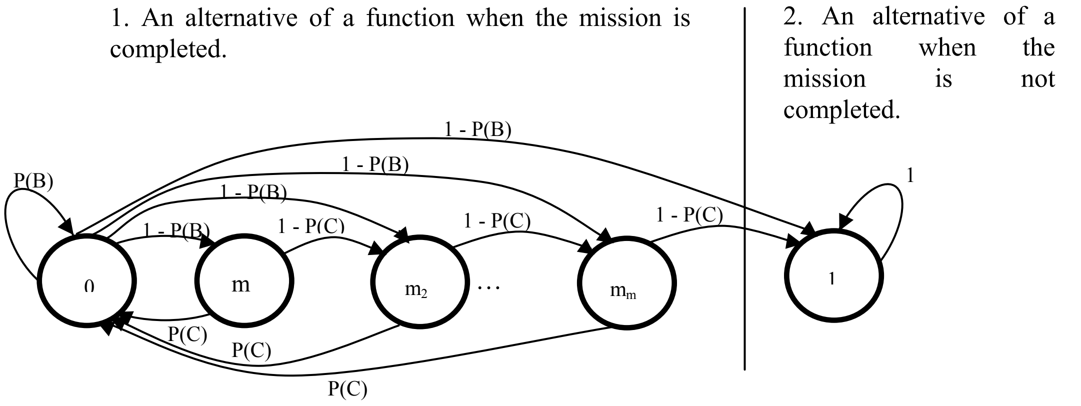
If a function of pyrotechnic cartridges is applied, it can remove failure m-times. So a number of restorations of the function is the same as the number of available pyrotechnic cartridges. In order to complete the mission successfully we need a higher amount of pyrotechnic cartridges *m*, or in the worst case the number of pyrotechnic cartridges should be equal to a number of failures.

Another alternative is the situation that a round fails and in this case a pyrotechnic cartridge fails too. A different pyrotechnic cartridge is initiated and it restores the function. This must satisfy the requirements that an amount of all round failures *n* is lower or at least equal to a number of operational (undamaged) pyrotechnic cartridges *m*.

The mission is completed in all the cases mentioned above and when following a required level of readiness of a block A.

The mission is not completed. In the second case the shooting is carried out one at a time, in small bursts or in one burst, and during the shooting there will be *n* round failures. At the time the failure occurs a backup system for restoring the function will be initiated. Unlike the previous situation there will be *m* pyrotechnic cartridges´ failures and a total number of pyrotechnic cartridges´ failures equals at least a number of round failures, and is equal to a number of implemented pyrotechnic cartridges M at the most. It might happen in this case that restoring of the function does not take place and the mission is not completed at the same time because there are not enough implemented pyrotechnic cartridges.

The relation of transition among the states can be expressed by the theory of Markov chains.


1. An alternative of a function when the mission is completed.
2. An alternative of a function when the mission is not completed.
P(B)
1 - P(B)
1 - P(B)
1 - P(B)
1 - P(B)
1 - P(C)
1 - P(C)
1 - P(C)
1
0
m
m2
mm
1
P(C)
P(C)
P(C)

Picture 1: Description of transitions among the states

Characteristics of the states:

0 state: An initial state of an object until a round failure occurs with a probability function of a round P(B). It is also a state an object can get with a pyrotechnic cartridge probability P(C) in case a round failure occurs $P(\overline{B}) = 1 - P(B)$, or $P(C|\overline{B}) = \dfrac{P(C \cap \overline{B})}{P(\overline{B})}$.

$m_1 \ldots m_m$ state: A state an object can get while completing the mission. Either a round failure occurs in probability $P(\overline{B}) = 1 - P(B)$, or there is a pyrotechnic cartridge failure in probability $P(\overline{C}) = 1 - P(C)$.

1 state: A state an object can get while completing the mission. It is so called an absorption state. Transition to the state is described as probability $P(\overline{C}) = 1 - P(C)$ of a failure of last pyrotechnic cartridge as long as an object was in a state „k_n“ before this state, or it can be described as probability of a round failure occurrence $P(\overline{B}) = 1 - P(B)$ as long as an object was in a state 0 before this state and all pyrotechnic cartridges are eliminated from the possibility to be used.

Transitions among different states as well as absolute probability might be put in the following formulae:

$$P(0) = P(B) + P(C_{k_1 0}) + P(C_{k_2 0}) + P(C_{k_3 0}) + \ldots + P(C_{k_{1n} 0}) \tag{3}$$

$$P(m_1) = 1 - P(B) \tag{4}$$

$$P(m_m) = (1 - P(B)) + (1 - P(C)) \tag{5}$$

$$P(1) = 1 \tag{6}$$

Transition probabilities are described using matrix of transition probabilities $\overline{P}$

$$\overline{P} = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix} \tag{7}$$

The arrows in picture 1 describe that the transition probability may occur with positive value. If we know the form of transition probability matrix $\overline{P}$ and original initial distribution of variable $p_i(0)$ than we can express the absolute probability of random variable p_i(n) as follows:

$$p_i(n) = \sum_{k \in I} p_k(0) p_{ki}(n), \qquad i \in I \tag{8}$$

This formula is possible to be expressed also in matrix form as follows:

$$\overline{P}(n) = \overline{P}(0)\overline{P}^n \tag{9}$$

We might describe the behaviour of the item in stationary state in terms of probability using limit probabilities p_j defined as follows:

$$p_j = \lim_{n \to \infty} p_{ij}(n), \qquad j \in I \tag{10}$$

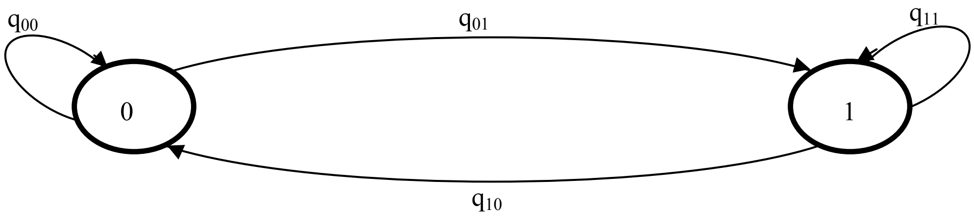
The importance of limit probabilities lies in expressing of weakening of initial conditions. With help of this statement we can get quiet exact picture about behaviour of our item observed. We are either happy enough to know that after going off the initial condition the item will with stay in one state with certain probability. Or we may use the help of absolute probabilities and to determine in which state the item will be after going off specific number of some measured units. This ways allows us to get the dynamic (in time) picture about the object observed.

4. CONTINUOUS DISTRIBUTION OF OBSERVED VARIABLE AND DEPENDABILITY

As we described the counting variable regarding the observed item above we also may use the continuous variable for getting a picture about the object behaviour. We are looking for random function NF X(t), where X(t) gets values from set I={0,1,2}. We call the items from set I as the states of observed process. If the parameter involved (time for instance) **t** = <0,∞), than we call the random function NF X(t) as Markov´s chain with continuous parameter. We also call such a chain homogenous if following formula is valid:

$$p_{ij}(s,t) = p_{ij}(0,s-t) = p_{ij}(t-s);\ s < t. \tag{11}$$

It is clear from above mentioned formula that the transition probabilities among each states are dependent on difference of arguments **t-s** and are independent on arguments **t** and **s** selves. Such a model is valid for those items and systems which are not capable to perform any operation even in reduced mode when a failure occurs. From the states point of view they immediately transfer from state "0" – operating state to state "1" – disabled state. This form is the most frequently used and for those items or systems with partial performance capabilities is extended of at least one mean state. Items or systems behaving in this way are not very suitable for us due to potential danger of complex inability to perform any function in case of failure. The transitions among states might be described either using probabilities or rates (as displayed bellow). The transitions among states might be any and the model has following form:



As well as in the previous part with counting parameter we use the same description for states. The assignment "0" means that the item/system is in operating state and the assignment "1" means that the item/system is in disabled state. Such a description may be applied on different completes (e.g. vehicles) systems (e.g. weapon system) or subsystems (e.g. engine) in frame of military equipment. We are also able to create plenty of different scenarios for each state description.

For transition rate is valid this form:

For i = 0 and j = 1 than it will be:

$$q_{ij} = \frac{1}{MTBF} = \frac{1}{E_P(X)}, \qquad (12)$$

where E_P (MTBF – Mean Time Between Failures) – is the mean value of time to failure and i∈{0;1}, j∈{0;1}, whereas j ≠ i.

For i = 1 and j = 0 than it will be:

$$q_{ji} = \frac{1}{MTTR} = \frac{1}{E_O(x)}, \qquad (13)$$

where E_O (MTTR – Mean Time To Reparation) – is the mean value of time to repair and i = 0, j = 1.

We presume following apart of above mentioned mathematical notations. The following formula is valid for the Markov´s chain with continuous parameter. We define the transition rate as follows. Lets have ***h*** which denotes an increment of the argument ***t***, than value $\boldsymbol{q_{ij}}$ where

$$q_{ij} = \lim_{h\to\infty} \frac{p_{ij}(h)}{h}, \text{ for } \boldsymbol{i \neq j} \qquad (14)$$

whereas p_{ij} denotes transition probability from state *i*into state *j* during an interval with length *h*, than we call the value $\boldsymbol{q_{ij}}$ as transition probability from state *i* into state *j*. Using formula (14) the following is also valid:

$$p_{ij}(\mathrm{h}) \approx \mathrm{q_{ij}.h}. \qquad (15)$$

If the $p_{ii}(h)$ denotes transition probability from state *i* into state *j* during a time interval *t*, than we call the value q_i, where

$$q_i = \lim_{h\to\infty} \frac{1-p_{ii}(t)}{h}, \text{ zde pokládáme } \mathrm{q_i = -q_{ii}}, \qquad (16)$$

as transition rate from state *i*. Using formula (15) the following form is also valid:

$$\mathrm{p_{ii}(h) \approx 1 - q_i.h}. \qquad (17)$$

Values q_i and q_{ij} also fulfil condition:

$$\mathrm{q_i} = \sum_{j\in I, j\neq i} q_{ij}, \text{ for all } i \in \mathrm{I}, \qquad (18)$$

where I is a set of states considered I∈{0;1;2;…}

We also would like to introduce the equations for transition probabilities calculation. The forms are as follows:

$$p_{ij}'(t) = \sum_{k \in I} p_{ik}(t).q_{kj}, \qquad \text{for } i,j \in \text{I}. \tag{19}$$

We also would like to introduce the equation system for absolute probabilities calculation. The forms are as follows:

$$p_i'(t) = \sum_{k \in I} p_k(t).q_{ki}, \qquad \text{kde i} \in \text{I}. \tag{20}$$

It is necessary to know the particular transition rates among states for exact calculation above mentioned differential equations. These equations are to give us exact information about the system and especially in what time the system will be in a particular state.

We see as suitable using the theory of "**Inherent availability of complex system composed from many mutually independent components**" for each measures (like the transition rate for instance) calculation. The results of these differential equations will give us the transition probabilities as well as the absolute probabilities for expressing what time the system will be in what state. Such a piece of information is exactly well related with the dynamic dependability measures. Our decision making would be much harder without this kind of information. That is why we do appreciate such as procedures for dynamic dependability indication especially regarding military vehicles.

5. CONCLUSION

This paper describes the procedures which are suitable for dynamic risk/dependability characteristics assessment. We have been desperately looking for new and progressive methods which allow us to get more precise view on military (battle) equipment. The more information about such as equipment we have the more successful the possible deployment might be.

One of things we have to take into account and not appear like it does not interest us is risk. The risk is very high both in training time and in real deployment as well as the risk profile. The first part of the paper deals with the basic understanding of risk and elementary formulas for its expression. The following parts show the dynamic dependability assessment and investigation both for counting and for continuous situations. We need to be aware using each procedure and respect each conditions in particular procedure.

Both of procedures shown above have been proved in frame of the Czech Armed Forces on respective equipment. In these examples has been confirmed the ability of mathematical procedures to express the system behaviour in terms of the dynamic dependability. The results were corresponding with reality as well as with our expectations.

ACKNOWLEDGEMENT

This contribution has been made with support of "Research Purpose Fund" of Faculty of Military Technology Nr. 0000 401, University of Defence in Brno.

BIBLIOGRAPHY

[1] KROPÁČ, J.: Vybrané partie z náhodných procesů a matematické statistiky, Brno: VA v Brně 2002, Skripta S-1971.

[2] KROPÁČ, J.: Základy náhodných funkcí a teorie hromadné obsluhy, Brno: VAAZ 1987, Skripta S-1751/A.

[3] VALIŠ, D.: Analysis of vetronics application consequences onto military battle vehicles dependability, Brno: VA v Brně 2003, Dissertation thesis.

[4] VALIŠ, D.: Fundamentals of description, perception and value determination of Risk. Liberec: Technical University 2005.

[5] VALIŠ, D., Contribution to Reliability and Safety Assessment of Systems. In: Sborník příspěvků konference – Opotřebení Spolehlivost Diagnostika 2006, Brno: Universita Obrany, 31. říjen – 1. listopad 2006, str. 329 - 337, ISBN 80-7231-165-4.

[6] VALIŠ, D., Assessment of Dependability of Mechatronics in Military Vehicles. In: Sborník příspěvků konference – Opotřebení Spolehlivost Diagnostika 2006, Brno: Universita Obrany, 31. říjen – 1. listopad 2006, s. 309 - 319, ISBN 80-7231-165-4.

[7] VALIŠ, D., VINTR, Z., Dependability of Mechatronics Systems in Military Vehicle Design. In: Proceedings of the European Safety and Reliability Conference “ESREL 2006” (September 18 – 22, 2006, Estoril, Portugal), London/Leiden/New York/Philadelphia/Singapore: Taylor&Francis Group 2006, p. 1703 - 1707, ISBN 10 0 415 41620 5.

ISSN 1932-2321