PRIMARY FACTORS AFFECTING SAFETY OF CONTROL SYSTEM

K. Rástočný, M. Foltán

Department of Control and Information Systems, Faculty of Electrical Engineering,
University of Žilina, Univerzitná 1, 010 26 Žilina
Phone: (+421-41) 513-3301, Fax: (+421-41) 513-1515, e-mail: karol.rastocny| michal.foltan @fel.utc.sk

Summary Aim of the paper is to point out some of the factors that affect safety of control system. Already in initial stages of the control system life-cycle it is necessary to analyze the effect of each of these factors on safety of a designed control system. Following this analysis, it is possible to propose appropriate measures for achieving specified safety goals.

1. INTRODUCTION

The control system is expected to execute control task with required level of availability. There is a special group of safety-relevant control systems with an addition request, that their malfunction will cause no injuries, considerable material harms, environmental destruction or other unwanted results. Safety as system property depends on several factors, that are close-knitted and their effects on the safety must be considered right from initial phase of the system life-cycle [1]. The level of safety can be evaluated quantitatively and/or qualitatively, by evaluation of its primary attributes – integrity, availability and confidentiality. The used measure of preferring a single safety attribute depends on a particular application [2].

2. PRIMARY FACTORS AFFECTING SAFETY OF CONTROLLED PROCESS

2.1. Redundancy rate in data processing

Function of redundancy in the system is to ensure some kind of response to failure. From this point of view redundancy can serve the purpose of:

- fault detection;
- fault masking;
- fault recovery.

If the reason of use of redundancy in the system is the reliability parameters improvement, then it is called reservation and use of redundant resources in the system is specified with the term *backup*. The system backup parts are those, whose use would have been useless, if the other parts of system had not worked correctly. That means, the term redundancy is wider term as backup.

If the reason of using redundancy in the system is to ensure improvement of its safety parameters, then redundancy is used only for failure detection (e.g. on the base of comparing results of multiple identical solutions tasks).

With appropriate use of redundancy it is possible to achieve an improvement of reliability and safety parameters.

According to, which redundant resources were used, we know the following forms of redundancy:

- Hardware redundancy – reservation of construction parts on all system levels;
- Software redundancy – implementation of diagnostic programs (they are not necessary for control function performance), repeating and multiple calculation programs;
- Information redundancy – the coding utilization for failure detection or failure correction;
- Time redundancy – additional time demands improvement of calculation with relation to information and software redundancy, or extension of calculation time due to its repeating.

The use of one redundancy form generally implies the use of another form of redundancy. Therefore wider redundancy classification is often used:

- Space redundancy;
- Time redundancy.

Space redundancy is adherent to hardware resources, whose use can be evoked by hardware redundancy, but also e.g. software and information redundancy, because existence of both these redundancy forms is generally bound to hardware. Time redundancy is usually bound to software, information and time redundancy resources.

2.2. System reliability

Generally, system operation must be regarded together with:

- Failures caused by environmental effects;
- Systematic failures;
- Random failures caused by system aging.

It is possible to avoid the failures caused by environmental effects (completely or partially), if during system design there are used such technological and circuit solutions, that meet the boundary conditions of operation environment. Operation environment of the system must be selected in accordance with specification of system requirements. This kind of solution also eliminates possible violation of mutual independence of system elements caused by environmental effects. However, in data transfer the environmental effects cannot generally be ignored.

Systematic failures do not occur as a result of

system aging, but their occurrence is bound to the specific situation and state of the system. Essential part of systematic failures is bond to mistakes, made in system development (e.g. software errors). One of the most important activities in development of the safety-critical control system is definition of functional requirements. Specification of functional requirements must be unambiguous, understandable, complete, consistent and controllable. Therefore it is recommended to realize it on the basis of semi-formal and formal methods. These methods are generally supported by software tools, which can serve the purpose of an automatic code generation and in final consequence also simplify system verification and validation. Such methods and techniques aimed at minimisation of systematic faults in software considerably help to increase the functional system safety.

The application of measures to failure prevention can considerably minimize failure occurrence caused by environmental effects and systematic failures. Evasion of accidental failures caused by system aging is impossible, therefore they must be considered during system operation.

Intensity occurrence of critical failures of the *i*-th system element can be very difficult or impossible to estimate. Therefore pessimistic assumption is generally accepted, that each element failure is potentially critical. Such assumption is acceptable, because

$$\lambda_{iK} \leq \lambda_i , \tag{1}$$

where λ_{iK} is the critical failure rate of the *i*-th element. In some cases such a procedure can be applied that is based on definition of "critical failures-to-all failures" ratio using long-term experience. E.g. if we assume, that electronic elements are elements with symmetrical failure occurrence (probability, that the failure will have negative effect on system safety is approximately equal to probability, that the failure will not have negative effect on system safety), then this coefficient has the value 0,5.

2.3. Fault detection and negation

Fault detection and thereafter fault negation has essential significance to reinsurance of required level of system safety. Depending up requirements technical diagnostics can be in different forms (functional diagnostics, test diagnostics, periodical diagnostics, continuous diagnostics, …). From safety point of view it is important, that every potential hazardous (critical) fault is detected and negated in required time (considering these faults it is the test with full coverage).

Functional diagnostics is used to detect faults affecting function of the object. The diagnostic system does not affect object of diagnostic with special test signals, but is only analysing signals created by activity of diagnostic object. The diagnostic system realized in this way can be excluded from the process of safety analysis and therefore it can use standard methods and techniques of information processing. Such a form of diagnostics is usually insufficient for safety – critical control systems.

Measures used to negate faults can be effective only if the fault is identified. Therefore the safety critical systems generally also contain test diagnostics in addition to functional diagnostics. The diagnostic system implements specific (test) signals into object of diagnostics and analyses responses. In case, that such a diagnostic system is also used when the object is in operation (operation diagnostics), the test signals are not allowed to corrupt normal object activity. Test diagnostics is used to detect faults during operation, which are not immediately visible (masked faults) if the object is active, but by change of system state or by combination with other fault can lead to critical state. In this case, the test techniques must be exposed to safety analysis, because they can be as well source of faults.

Even if the process of system elements testing and their faults negation is generally specific for each system, it must be true, that

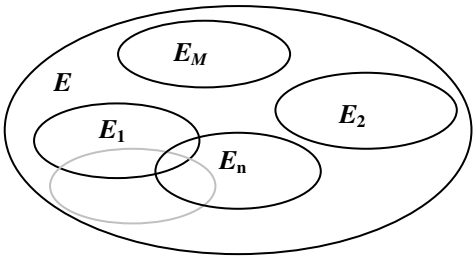
$$t_{ij} \leq t_{i0} \; pre \forall \; j \in \{1, 2, ..., m\}, \tag{2}$$

where t_{ij} is detection and negation time of the *j*-th fault of the *i*-th element, t_{i0} is the maximal allowed time to fault detection and negation of the *i*-th element and *m* is a number of faults of the *i*-th element. Time t_{i0} can be determined on the basis of failure rate of the *i*-th element and required system safety level.

Let the system dispose of faults detection mechanism whose diagnostic covering is $c \leq 1$. Then:

$$\begin{aligned} \lambda_{iM} &= \lambda_i.(1-c); \\ \lambda_{iD} &= c.\lambda_i ; \\ \lambda_i &= \lambda_{iM} + \lambda_{iD} ; \end{aligned} \tag{3}$$

where λ_{iM} is intensity of undetectable (masked) faults of the *i*-th system element and λ_{iD} is failure rate of detectable faults of the *i*-th system element.



Legend:
E - set of all faults
E_M - set of undetected faults
E_1 - set of faults detected with the 1st mechanism
E_2 - set of faults detected with the 2nd mechanism
E_n - set of faults detected with the *n*-th mechanism

Fig. 1. Set of system faults

Real systems generally dispose of different fault detection mechanisms that differ by diagnostic covering and/or fault detection time. It is clear, that some system faults can be covered by several faults detection mechanisms (Fig. 1).

According to Fig. 1 failure rate of the *i*–th element can be expressed as relation

$$\lambda_i = \lambda_{iM} + \sum_{j=1}^{n} \lambda_{i(j)} + \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} \lambda_{i(j,k)} + \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} \lambda_{i(j,k)} + \\ + \sum_{j=1}^{n-2} \sum_{k=j+1}^{n-1} \sum_{l=k+2}^{n} \lambda_{i(j,k,l)} + \ldots \quad (4)$$

where *n* is number of detection and negation mechanisms of element faults; $\lambda_{i(j)}$ is failure rate of the *i*–th element detectable **only** by the *j*–th mechanism; $\lambda_{i(j,k)}$ is failure rate of the *i*–th element detectable by the *j*–th **as well as** the *k*–th mechanism etc.

Multiple detection mechanisms in the system are implemented to increase system faults detection coverage. Each mechanism is characterized by different fault detection times. However, slower detection mechanism can not be used in full spread, because in case of fault occurrence, the fault is detected only if:

- The mechanism is able to detect this fault;
- Fault was not detected by faster mechanism.

Likewise intensity of system transition into safe state after fault detection and negation of the *i*-th element can be expressed by the equation

$$\delta_i = \sum_{j=1}^{n} \delta_{i(j)}, \quad (5)$$

where $\delta_{i(j)}$ is intensity of the *i*–th element transition into safe state after fault detection and negation of the *j*–th mechanism.

Generally, the possibility of fault detection mechanism failure must be also considered. Influence of this fact on system safety depends on particular system design. There is tendency to proceed in system design so that:

- Test diagnostics is independent (functionally and physically) from control circuits and then the effect of failure of test diagnostics on system safety can be uniquely proved, or
- Test diagnostics failure is identified by comparator mechanisms (fast diagnostics).

2.4. Service staff mistake

Even if we speak about control system safety, we have on mind safety of the controlled system. If it is a control system with continuous operation, than the control system failure is bound to two sorts of hazards:

- Primary hazards – transport process can be endangered due to faulty realization of the control function by the system;
- Secondary hazards – the system has protection mechanisms that get the system into the safe state after fault detection; safe state is characterized by partial or total reduction of control functions performance; in this case realization of control functions is taken over by the service staff; risk resulting from existence of these hazards will be proportional to probability of failure of this staff and time of system (or system part) failure.

Apparently intensity of service staff mistakes is essentially higher that intensity of hazardous system failures, especially if it is the fail-safe system. Intensity of personal mistake depends on many objective (number of actions, operation accomplishment time, complexity of action, …) and subjective (psychic condition, qualification level…) factors.

2.5. System availability

Though system safety integrity is dominant attribute of system safety, it's not the only one. Generally system safety can be conceptualized like pack of system attributes, where besides integrity especially system availability can be inserted [3].

Since coefficient of availability *A* is defined:

$$A = \frac{MUT}{MUT + MDT}, \quad (7)$$

where MUT is mean up time and MDT is mean down time, it is apparent that the higher availability of the system is, the lower risk of transport process will result from mistake of the service staff in emergency system operation.

Generally, coefficient of availability can be improved by technical measures (e.g. using high reliable components in design of safety-related systems, using diagnostic equipment that minimizes

time to identify failures) and organizational measures (e.g. optional location of service centres, improving qualification of maintenance staff, proper technical manuals).

2.6. System confidentiality

Confidentiality is system attribute that ensures, that system won't be misused by unauthorized subject. Term subject in this meaning does not include only persons, but also technical resources and software.

Confidentiality represents hierarchically ordered mechanism, that guarantees required competence level of rights (e.g. to write and read information) in a given part of the control system to human or machine. Confidentiality level can be expressed like stochastic process of successful or unsuccessful attempts to break it.

3. CONCLUSIONS

System safety integrity and system availability are close knitted, i.e. in case that required level of one or another attribute becomes not ensured this may prevent creating of the control system with required quality. Fulfilling safety integrity and availability requirements can be achieved only if the requirements for no-failure operation (error and fault effect on system functionality), maintainability (scheduled maintenance, identification and localization of fault states, system restoration after fault) operation (operation mode) and maintenance (human operator effect on effective maintenance, maintenance technique etc.) are fulfilled.

It's necessary to note, that strict safety requirements on safety-critical control system can not be proved only by tests or using practical results. To prove, that safety requirements are fulfilled and final risk is acceptable, only proper combination of qualitative and quantitative methods of failure effects analysis may be used.

This work has been supported by the Agency for science and technical support by financial support no. APVV - 20 - P00705.

REFERENCES

[1] ZAHRADNÍK, J. – RÁSTOČNÝ, K. – KUNHARD, M.: *Bezpečnosť železničných zabezpečovacích systémov*. EDIS – vydavateľstvo ŽU, 2004, ISBN 80-8070-296-9

[2] RÁSTOČNÝ, K.: Modelling of Failure Effects to Integrity of System. *AEEE No. 2 Vol. 3 /2004*. ŽU v Žiline, pp. 91-94, ISSN 1336-1376

[3] RÁSTOČNÝ, K. - JANOTA, A. - ZAHRADNÍK, J. - FRANEKOVÁ, M.: How to negate risk resulting from implementation of new functions into the existing safety-related systems. *FORMS/FORMAT 2004,* pp. 24-29, December 2-3 2004, Braunschweig, Germany, ISBN 3-9803363-8-7