Transient Fault Tolerance and System Safety Enhancement Based on System Theory

Xiongfeng Huang, Chunjie Zhou, Yuanqing Qin, Ye Wang, Mingyue Yang

Key laboratory of ministry of education for image processing and intelligent control, Department of control science and engineering, Huazhong University of Science and Technology, Wuhan, Hubei, 430074, China

Email: sxduxhuangxf@163.com, cjiezhou@mail.hust.edu.cn, yuan_qing@163.com, wangye.happy@yahoo.com.cn, ymy7461@163.com

Abstract—Transient faults are hard to be detected and located due to their unpredictable nature and short duration, and they are the dominant causations of system failures, which makes it necessary to consider transient fault-tolerant design in the development of modern safety-critical industrial system. In this paper an approach based on system theory is proposed to tolerate the transient faults in tunnel construction wireless monitoring and control systems (TCWMCS), in which the effects of transient faults are expressed by dysfunction of interactions among software applications. After analyzing the dysfunctional interactions of the system by the operational process model and deducing the causes of dysfunction in the functional control diagram, a safety enhancement way is proposed for the designers, in which effective safety constraints are set up to tolerate the transient faults. The experiment evaluation indicated that the effects of transient faults could be exposed by the causal factors of dysfunctional interactions and system safety could be enhanced by the enforcement of appropriate constraints.

Index Terms—transient faults, dysfunctional interactions, wireless monitoring and control system, system approach, hazardous action, safety constraints

I. INTRODUCTION

Along with the development of national economy, the project of tunnel construction has been steadily increasing [1, 2]. Tunnel construction is a typical safety-critical underground infrastructure with narrow space, severe environment, high accident potential, and fickle structure. Problems of tunnel construction safety have been focused on. The monitoring and control systems based on computer network technology are generally applied to ensure the construction safety [3, 4]. To meet the special safety-critical requirements, transient faults must be tolerated in operation of TCWMCS. In system theory view, the effects of transient fault can be expressed by the dysfunction of interactions among software applications and the uncontrolled system failures generated by dysfunctional interactions could be thought as lack of pre-set constraints in early design stage. Because of the unpredictable nature and short duration, transient faults are difficult to be detected and located. Failure cause analysis technology based on system theory is employed to tolerate the transient faults by identifying the causal factors of dysfunctional interactions and enhance the system safety by setting up constraints in early stage.

At present, accident causation models based on event chains are used to reduce system failures, such as Failure Modes and Effects Analysis (FMEA) [5], Fault Tree Analysis (FTA) [6], Event Tree Analysis (ETA) [7], and Cause-Consequence Analysis [8]. These methods work well for identifying the causes of losses caused by permanent faults in relatively simple systems. However, they are limited in their capability to deal with the failures caused by transient faults in the more complex systems that were developed in modern industry. To deal with those failures, new paradigm is needed. Safety analysis based on system theory has grown up [9-11]. According to Leveson [9], system accidents arise from the interactions among components rather than individual component failure, and failures caused by transient faults can be thought as lack of constraints enforced on interactions. Leveson [9] has developed a pure systems-theoretic model of accident causation called STAMP (Systems-Theoretic Accident Modeling and Processes) based on system theory which takes systems as a whole and does not rely on redundancy are developed. And from the analysis of failure of the Titan IV B-32 mission by STAMP, dysfunctional interactions were assigned to be the dominant causations of failures. Takehisa Kohda et al [10] proposed an approach based on system control relations for Accident Analysis of Protective Systems. Laracy et al [11] applied STAMP to Critical Infrastructure Protection. The two cases showes the usefulness for identifying the causations by safety analysis approach based on system theory.

In this paper, we proposed a novel safety analysis approach based on system concepts, to find out the potential hazardous control actions of dysfunctional interactions in communication of TCWMCS, and to set up constraints for eliminating them. Firstly, we introduced this safety analysis approach, including the foundation of system approach, the dysfunctional interactions in control system, and detailed safety analysis processes of this
approach. Secondly, we applied this approach on safety analysis of communication in TCWMCS. Operational process model was established to describe the execution of the communication mission, and the dysfunction of interactions due to potential hazardous actions were identified by the system approach, eventually constraints on hazardous scenarios were set up to deal with the dysfunction of interactions in order to tolerate the transient faults and enhance the system safety. Finally, we compared the analytical ability about FTA (Fault Tree Analysis) method and System Approach, and a case of system safety enhancement by enforcing the constraints were discussed. So the effectiveness of System Approach was indicated by the results.

The rest of this paper is organized as follows: Section II describes the system safety analysis method. Section III describes the safety analysis of dysfunctional interactions in communication of TCWMCS under the proposed approach. Section IV describes the evaluation to show the efficiency of system approach. Finally, Section V concludes this paper.

II. SYSTEM APPROACH ON SAFETY ANALYSIS

Systems theory dates from the thirties and forties in last century and it was a response to the limitations of classic analysis techniques in coping with the increasingly complex systems being built [12]. In system concepts, system safety can be viewed as a control problem. Accidents result from interactions among components that violate the safety constraints—in other words, from a lack of appropriate control actions to enforce the constraints on the interactions [13]. The dysfunctional interactions are the interactions which violate the system safety constraints, which can be found out through the operational process models. The causes of dysfunctional interactions generally are not independent, so they should be induced by functional control diagram based on system theory. System safety can be enhanced by enforcing enough constraints according to the causes.

A. Foundation of system approach

The foundation of systems theory rests on two pairs of ideas: (1) emergence and hierarchy and (2) communication and control [13].

A general model of complex systems can be expressed in terms of a hierarchy of levels of organization, each more complex than the one below, where a level is characterized by having emergent properties. Emergent properties do not exist at lower levels; they are meaningless in the language appropriate to those levels. Hierarchy theory deals with the fundamental differences between one level of complexity and another. Its ultimate aim is to explain the relationships between different levels: what generates the levels, what separates them, and what links them. Emergent properties associated with a set of components at one level in a hierarchy are related to constraints upon the degree of freedom of those components. In a systems view of safety, the safety properties is emergent, which are controlled or enforced by a set of safety constraints related to the behavior of the system components. Safety constraints specify those relationships among system variables or components that constitute the non-hazardous or safe system states. Accidents result from interactions among system components that violate these constraints, that is to say, there is lack of enough appropriate constraints on system actions.

Regulatory or control action is the imposition of constraints upon the activity at one level of a hierarchy, which define the “laws of action” at that level yielding activity meaningful at a higher level. Hierarchies are characterized by control processes operating at the interfaces between levels. Control in open systems, which have inputs and outputs from their environment, implies the need for communication. In system theory, open systems are viewed as interrelated components that are kept in a state of dynamic equilibrium by feedback loops of information and control. Systems are not treated as a static design, but as a dynamic process that is continually adapting to achieve its ends and to react to changes in itself and its environment. For safety, the original design must not only enforce enough appropriate constraints on action to ensure safe operation or the enforcement of the safety constraints, but it must continue to operate safely as changes and adaptations occur over time.

B. Dysfunctional interactions among software applications

The dysfunctional interactions among software applications can be viewed as lack of safety constraints. Dysfunctional interactions exist in the interactive process of mission operation. An operational process is an organic active process through which a specified objective can be met. Operational process model is established to describe the execution of a mission. Fig. 1 shows the operational process model between two software applications in a control system. The software application is a communication entity. “Command/Inquiry” indicates the command or inquiry sent to the other software application. “Feedback” indicates the acknowledgement to the sponsor for the next action. The states of software applications are dynamically changing over time in an operational process.

Dysfunctional interactions existing in control loop are shown in Fig. 2. “Command/Inquiry provided early or later” and “loss of feedback or incorrect feedback” are classical dysfunctional interactions. These dysfunctional interactions should lead to system failures. During the early design stage, enough constraints must be set up to tolerate the dysfunctions.
The causes of the dysfunctional interactions are thought as lack of enough constraints, which should be exposed by functional control diagram. A system safety enhancement approach based on system concepts is proposed to expose them and enhance the system safety in the next.

C. System approach

This approach is developed to expose the dysfunctional interactions which could cause unacceptable events, such as loss of life, as well as advise system designers or maintainers to set up safety constraints which can be enforced on the interactions to tolerate transient faults and enhance system safety. The detailed analysis processes are described as follows, which involves four steps:

1) Establishing operational process model. The operational process model is an abstract model of a mission, which can be broken down into several operational processes. The operational process can be viewed as some control actions. In software-intensive system each of the operational processes is comprised of feedback, communicating between software applications, and an environment effect. The conceptual model is shown in Fig. 1. It is the fundamental of the safety analysis.

2) Defining system-level safety constraints. A system-level safety property can be viewed as an emergent property from the lower-level safety constraints. Design constraints are derived from design requirements. Dysfunctional interactions are viewed as lack of constraints. Constraints on dysfunctional interactions are low-level rules in compliance with the system-level constraints, which are the basis of the analysis of dysfunction.

3) Identifying hazardous actions. In the operational process model, three general categories of inadequate control actions were used to describe the dysfunctional interactions which would violate the safety constraints. The three general categories of inadequate control behaviors were “Not provided when it should be,” “Incorrectly provided,” and “Provided Too Early, Too Late, Out of Sequence.” The inadequate control actions were thought to be the potential hazard violating the safety constraints.

4) Setting up new safety constraints. To eliminate the potential hazardous actions, the causes must be identified. Functional control diagram [8] shown in Fig. 3 was used to identify the causes. In the diagram, the dysfunctional interactions could be listed in the control loop included inadequate control behaviors, delays, mismatches and disturbance etc. After the causes had been identified, safety constraints could be set up to eliminate or control them, such as “delay can not exceed 3 seconds” and “feedback must be correct”. Design features should be considered to enforce those constrains in system design stage.

III. SYSTEM SAFETY ANALYSIS AND IMPROVEMENT OF TUNNEL CONSTRUCTION WIRELESS MONITORING SYSTEM

Tunnel construction wireless monitoring and control system is a novel networked control system, in which wireless sensor networks are used to link the construction site and decision-maker. Because large quantities and multiple categories of software applications should be applied to meet the special safety-critical requirements in TCWMCS, dysfunctional interactions among them are the dominant causations leading to system failure which could be effected by transient faults. We used the proposed approach to analyze the dysfunctional interactions of communicaiton in TCWMCS. The rest of this chapter introduces the details.

A. Overview of the Tunnel Construction Wireless Monitoring and Control System

TCWMCS contains wireless sensor network, wired optical network, local supervisory centre and remote supervisory centre, whose structure is shown in Fig. 4. Wireless sensor network consists of sensor nodes, sink nodes and gateways. Sensor nodes collect tunnel environmental information from sensors and send them to sink nodes, as well as execute commands from the upper monitoring and control system. Sink nodes exchange messages with the gates, as well as fuse the messages which send to the TCWMCS that can decrease the total sending informations. Informations are exchanged between the Gateway and the upper monitoring and control system. Wired optical network are applied to connect the gateway and the local supervisory centre. Local supervisory centre deals with the environment informations in tunnel which are collected by the sensors, and it also sends those data to the remote supervisory centre. Control commands would be made by the local decision-maker and remote centre controller. Especially
when emergency occurs, the control commands are made by multi agents.

Transient interference such as electromagnetic interference could affect a short communication break, which could lead incorrect emergent decisions with resulting catastrophic hazards. These kinds of hazards could be thought as causation of system failure which caused by dysfunctional interactions as “command provided too late”. Enough constraints should be set up to eliminate them in system design stage. To identify the causal factors, we should establish an operational process model to describe the communication.

B. Operational process model in communication

To limit the complexity of the operational process model, the communicating process is split into hierarchical abstraction. Fig. 5 shows the operational process model of communication in TCWMCS. It is composed of three major components: upper monitoring and control system, local gateway and wireless nodes. The connection lines indicate the interactions. Upper monitoring and control system plays the decision-maker role in TCWMCS. Local gateway is the relay of communication. Wireless node is the terminal which collects the information of tunnel construction site and executes the commands from upper monitoring and control system.

Fig. 6 shows the operational process model inside wireless node. Major components inside the node include RF (Radio Frequency) Module, micro processor and sensors. RF module is the physical layer of the communication network. Micro Processor conducts the communication through the control of receiving tunnel information from sensors and sending tunnel information to upper level. Sensors are the nodes deployed in the construction site to collect tunnel environmental information and some actuators to enforce the control commands. Connecting lines between those components show control actions, information, and feedback.

In order to annotate some of those connecting lines with command actions, we reviewed the nominal command sequence during communication phase. The actions are listed as follows during communication phase:

1) Collecting environmental data
After the sensors collected the environmental data, the micro processor should converge those data from the sampling sensors.

2) Sending environmental data to RF Module
After micro processor received environmental data and processed, it sends “inquiry” signal to RF module. When acquired “Enable”, the processor sends environmental data to RF sets.

3) Sending environmental data to local gateway
Through WSN, RF module transmits environmental data processed by the micro processor. Then local gateway acquires from the RF channel.

4) Sending environmental data to upper monitoring and control system
Local gateway links upper monitoring and control system, and then sends environmental data through wired network.

5) Control command passed down to gateway
After acquired filed data, the upper monitoring and control system makes a decision and passes down the control command to local gateway.

6) Control command passed down to wireless node
   Local gateway passes down the control command to wireless node. Nodes treat the command and send it to many actuators.

7) Control command executed by the actuator
   After acquired the control command, the actuator executes.

C. System-level safety constraints

In TCWMCS, the safety of the communication between the sensor nodes and decision-maker is a critical requirement and it can be translated into the system-level constraints such as “the transmission of informations should be ensured in time and correctly”. The software applications in TCWMCS perform the communication. The communication failure caused by dysfunctional interactions could be thought as lack of safety constraints in system design stage.

According to the system-level safety constraints, the communication operations in TCWMCS were partitioned to tree interactions as follows: 1) Environmental data collected to the gateways; 2) Gateway sending data to upper monitoring and control system; 3) Upper monitoring and control system publishing inquiry and control command. All those interactions contains information provision control and information acquisition control, which are reflected by commands.

D. Identification of hazardous control actions

For each action, the causation of a failure would be identified by three categories of inadequate control behaviors: “Not provided when it should be,” “Incorrectly provided,” and “Provided Too Early, Too Late, Out of Sequence.” We identified three types of failures caused by dysfunctional interactions, which are underlined in Table I.

Each cell in the table describes what would happen if each action is executed inadequately. Each failure was assigned an identifier from (a) through (c). A failure (b), for instance, is that if the environmental information has not been provided due to transient faults, the upper monitoring and control system might receive incorrect tunnel information and make incorrect decisions, which is hazardous with no prevention mechanism. To failure (b), there are two actions and two kinds of potential hazardous control behaviors. Causal factors should be identified by employing the functional control diagram. An detailed analysis about failure (b) are introduced in the next.

E. Identification of Causes and Setting up system safety constraints

After the hazardous control action has been identified, design features must be used to eliminate or control it. To accomplish this goal more information about the cause of hazardous action are required and those informations could be provided by the functional control diagram. For an example of the analysis, we selected the hazardous control action (b) in Table I.

Hazardous Action (b) can be described as follows: Upper monitoring and control system receive incorrect environmental informations in tunnel and make incorrect decisions, which is hazardous with no prevention command being made.

A functional control diagram shown as Fig. 7 was established to identify the causal factors leading to hazardous control action (b), where t and x denote the

<table>
<thead>
<tr>
<th>Failure</th>
<th>Action</th>
<th>Potential hazardous control behavior</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Collecting environmental data to gateway</td>
<td>Not provided when it should be</td>
<td>upper monitoring and control system can not receive tunnel environmental data</td>
</tr>
<tr>
<td>b</td>
<td>Gateway transmit environmental data to upper monitoring and control system</td>
<td>Not provided when it should be</td>
<td>upper monitoring and control system acquire incorrect information about tunnel and should make incorrect decision</td>
</tr>
<tr>
<td>c</td>
<td>Collecting environmental data to gateway</td>
<td>Out of sequence/ disorder</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gateway transmit environmental data to upper monitoring and control system</td>
<td>Out of sequence/ disorder</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Collecting environmental data to gateway</td>
<td>Too early</td>
<td>the period of messages received by upper monitoring and control system are incorrect</td>
</tr>
<tr>
<td></td>
<td>Gateway transmit environmental data to upper monitoring and control system</td>
<td>Too early</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Collecting environmental data to gateway</td>
<td>Too late</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gateway transmit environmental data to upper monitoring and control system</td>
<td>Too late</td>
<td></td>
</tr>
</tbody>
</table>

Figure 7. Causal factors leading to hazardous control action (b).
sequence of the environmental data received and the environmental data, respectively. If \( t \) is incorrect, the correct \( x \) may be neglected, and then a fallacious decision would be made which would lead to accident. Scenarios were assumed to specify the reason why the accidents occur. For example, workers were in the tunnel, but the decision-maker thought they were not in. This hazardous scenario will lead to fallacious decision which is harm to the safety of workers. Some hazardous scenarios that could lead to hazard (b) are listed as follows:

1) Upper monitoring and control system component failures
   The auto component failed, but the back-up component or the supervisor acts incorrectly.
2) Supervisor mistakes in operations
   Supervisor makes a careless error to issue a command.
3) Abstract network model inconsistent
   A prevent command was passed down to the wrong terminal.
4) Wireless monitoring environment node failure
   The incorrect environmental informations send by the node were thought as correct informations.
5) The location node failure
   The location informations of constructors were wrong, but a decision made to deal with an emergency treat them as correct informations.
6) Wireless node R/F sets out of sequence
   Some packages were discarded because of wrong sequence due to the transient electromagnetic interference, but the messages was thought as correct messages.

Constraint as “appropriate protocol must be designed to ensure the sequence” is set up for the sixth scenario listed above. If the system was designed under this constraint in early stage, the transient fault such as electromagnetic interference should be tolerated in communication.

IV. EVALUATION

A. Efficiency of system approach

To show the efficiency of identifying the dysfunctional interactions, FTA [14] is used here to compare with system approach. FTA is an occupational safety model in system safety analysis. The comparison of the causes found by FTA methods and system approach is shown in Table II.

In this table, the ‘\( \checkmark \)’ indicates that hazardous action could be identified, the ‘×’ indicates that action could not be identified. We found that causal factors other than component failures such as process model inconsistency, causal factors with regard to “delay of command” and “delay of feedback” which could be caused by transient faults are not identified by FTA. The hazardous actions that had been identified by FTA could be also found out by the proposed approach and the causal factors in the control loop about dysfunctional interactions were not given by FTA.

The result shows that System Approach is more effective to found out the dysfunctional interactions among software applications and the transient faults could be exposed by the causal factors of dysfunctional interactions.

<table>
<thead>
<tr>
<th>Potential hazardous actions</th>
<th>FTA</th>
<th>System Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>*supervisory system component failure</td>
<td>( \checkmark )</td>
<td>( \checkmark )</td>
</tr>
<tr>
<td>*incorrect operation of supervisors</td>
<td>×</td>
<td>( \checkmark )</td>
</tr>
<tr>
<td>*inconsistency of the upper system controller abstract model and it real model</td>
<td>×</td>
<td>( \checkmark )</td>
</tr>
<tr>
<td>*Inquiry/Control Command missing/inappropriate</td>
<td>( \checkmark )</td>
<td>( \checkmark )</td>
</tr>
<tr>
<td>*Inquiry/Control Command delayed</td>
<td>×</td>
<td>( \checkmark )</td>
</tr>
<tr>
<td>*monitoring nodes failure</td>
<td>( \checkmark )</td>
<td>( \checkmark )</td>
</tr>
<tr>
<td>*data delayed in WSN</td>
<td>( \checkmark )</td>
<td>( \checkmark )</td>
</tr>
<tr>
<td>*other physical disturbance</td>
<td>( \checkmark )</td>
<td>( \checkmark )</td>
</tr>
<tr>
<td>*t, x parameter missing/inappropriate/delayed</td>
<td>×</td>
<td>( \checkmark )</td>
</tr>
<tr>
<td>*t, x parameter incorrect</td>
<td>×</td>
<td>( \checkmark )</td>
</tr>
</tbody>
</table>

B. A case study of system safety enhancement

From table II, the delay of feedback information is a potential hazardous action. A constraint as “real time of communication must be ensured” can be drew out. To realize that, many communication protocols [15, 16] have been developed based on FTA. However the deployment optimization of the WSN nodes is not considered by FTA but System Approach, which is an effective measure to improve the real-time performance of WSN. In this section, the simulation of optical node deployment based on GA (genetic algorithm) [17] is presented to show that the safety can be enhanced by ensuring the system real-time performance.

The performance of optimal specific deployment by GA algorithm was compared with general random deployment, such as [18], in terms of the end to end delay. To evaluate the performances of the networks properly, experimental parameters of the two protocols were set as the same. Our grid size was 2×2m2, the range of each terminal (routing) node was 100m (100m), the initial energy of each relay node was 5 J, and the GA parameters, such as population size, crossover rate and mutation rate, were 200, 0.8 and 0.2.
In the experiments: we used a 72-node network, with 50 terminal monitoring nodes, 20 routing nodes, and 2 gateways, for a tunnel construction of 5KM, which needed $10 \times 2500 = 25000$ grids. Size of a chromosome equaled to the number of grids. Crossover which was performed after one chromosome had been divided into 250 sectors which equaled to 250 grids having $10 \times 10$ size. The simulation was done 1000 times for each condition.

![Figure 8](image-url)

**Figure 8.** Evolution progress of the best individual.

The evolution progress of the best GA run is shown in Fig. 8, where both the fitness progress of the best individual found by the algorithm as well as the average fitness of the entire population at each generation are plotted. The fitness value is greatly enhanced after 125 generations due to the selection of the best fitness chromosomes to be used in the next generation.

![Figure 9](image-url)

**Figure 9.** An example of the optimal deployment.

After analysis of the best fitness chromosomes, we can find optimal deployment is that routing nodes must be laid along the middle line, and inside the cluster, terminal monitoring nodes closely surround the two routing nodes, in this case, they are just in the next grids. The number of cluster-member is dependent on the distance of each cluster to the base stations, which indicates more cluster-members in clusters closer to BS and less cluster-members in clusters farther from BS, as is shown in Fig.9.

![Figure 10](image-url)

**Figure 10.** End to end delay under optimal and general deployments.

Fig. 10 delineate the comparative performance in terms of end to end delay of monitoring network under optimal and general deployments. Results show that the end-to-end delay is also less than general deployment. These show the efficiency of the proposed deployment optimization strategy in reducing end-to-end delay. At the same time, the results indicate that system approach is a useful measure to enhance the system safety.

V. CONCLUSION AND FUTURE WORK

Transient faults are thought to be the grave threats in modern safety-critical industrial systems. In this paper, to tolerate the transient faults in TCWMCS, safety analysis and enhancement based on system concepts was proposed. The failures caused by transient faults were treated as lack of enough interactive constraints in the proposed approach. Using the operational process model, the interactions among the software applications were described. And following the functional control diagram, the causal factors of dysfunctional interactions were educed. Thereafter safety constraints could be set up to tolerate the transient faults from the hazardous scenarios, and the system safety could be enhanced by the enforcement of constraints. From the analysis of the communication of TCWMCS, we found that the constraints set up to eliminate some causal factors of dysfunctional interaction were the strategies to tolerate the transient fault and the system approach was an effective measure to enhance the system safety.

In our work, the safety constraints could only be recommended in the conceptual level. In the future, we will continue to work actively on the engineering safety analysis approach based on system concepts, including improving this approach in setting up more detailed constraints and implementing it in more practices.

ACKNOWLEDGMENTS

The author wish to give their sincere thanks to the editor and the anonymous referees for their valuable suggestions and helpful comments which improved the presentation of the paper.

REFERENCES


