Toward Modelling of Human Performance of Infrastructure Systems

Cen Nan[ac] and Wolfgang Kröger[b]

[a] Reliability and Risk Engineering Group (RRE), ETH Zürich, Switzerland
[b] ETH Risk Center, ETH Zürich, Switzerland
[c] Land Using Engineering Group (LUE), ETH Zürich, Switzerland

Abstract: During the last decade, research works related to modelling and simulation of infrastructure systems have primarily focused on the performance of their technical components, almost ignoring the importance of non-technical components of these systems, e.g., human operators, users. In contrast, the human operator of infrastructure systems has become an essential part for not just maintaining daily operation, but also ensuring the security and reliability of the system. Therefore, developing a modeling framework that is capable of analyzing the human performance in a comprehensive way has become crucial. The respective framework, proposed in this paper, is generic and consists of two parts: an analytical method based on the Cognitive Reliability Error Analysis Method (CREAM) for human performance assessment and an Agent-based Modeling (ABM) approach for the representation of human behaviors. This framework is a pilot work exploring possibilities of simulating human operators of infrastructure systems through advanced modeling approaches. The demonstration of the applicability of this framework using the SCADA (Supervisory Control and Data Acquisition) system as an exemplary system is also presented.

Keywords: Human Reliability Analysis, CREAM, Agent-based Modeling, Critical Infrastructure, SCADA

1. INTRODUCTION

Modern infrastructure systems, e.g., power supply, telecommunication and rail transport systems, are all large-scale, highly integrated, particularly interconnected and show complex behaviours. These systems are so vital to any country that their incapacity or destruction would have a debilitating impact on the health, safety, security, economics and social well-being [1]. The operators of these systems must continuously monitor and control them to ensure their proper operation [2]. These industrial monitor and control functions are generally implemented using an industrial control system (ICS), e.g., the SCADA system. The fundamental purpose of this type of systems is to allow its users (operators) to collect data from one or more remote facilities and send control instructions back to those facilities [3]. Most research studies on infrastructure systems, especially on this type of ICS, have taken an engineering point of view, which often underestimate the importance of their non-technical components, e.g., human operators [4, 5]. A number of studies have shown that human errors are major causes for accidents occurred in electric power, railway, aviation and maritime infrastructure sectors [6-8], highlighting the significance of examining the reliability of the human operators, which can be conducted using analytical methods and advanced modelling approaches.

2. RESEARCH STREAMS AND PROPOSED FRAMEWORK

Over the years, many Human Reliability Analysis (HRA) methods have been developed to analyse human performance in either qualitative or quantitative ways. Qualitative methods focus on the identification of events or errors, while quantitative methods focus on translating identified events/errors into Human Error Probability (HEP) [9]. The Technique for Human Error Rate Prediction (THERP), one of first generation HRA methods, is probably the most widely used technique to date [10]. THERP aims to calculate the probability of successful performance of the activities defined necessary for the accomplishment of a task. The calculations are based on pre-defined error rates (HEPs) and success is defined as the complement to the probability of making an error. Appropriate HEPs from a list around 100 factors are selected for a nominal assessment [11]. The results of the task analysis are represented graphically in a so-called HRA event tree that is a formal representation of the required sequence of actions. The use of the THERP causes limitations during human performance analysis since this method is focused on errors of omission and intends to characterize each operator action with a binary path (success or failure). Moreover, the representation of Performance Shaping Factors (PSFs) influence on human performance is quite poor and highly judgmental based on assessor's experiences [10, 12]. Success Likelihood Index Method (SLIM), another example of first generation HRA methods, is used for the purposes of evaluating the probability of a human error occurring throughout the completion of a specific task [13]. It is a decision-analytic approach, which uses expert judgment to quantify PSFs. Such factors are used to estimate a Success Likelihood Index (SLI), a form of preference index, which is calibrated against existing data to derive a final HEP. This approach is a flexible technique and able to deal with the total range of human error forms. SLIM is a subjective method and the choosing of PSFs is quite arbitrary. Another disadvantage of this approach is that there is a lack of valid calibration data [12]. A Technique for Human Event Analysis (ATHEAHA), one of second generation HRA methods, is designed to support the understanding and quantification of Human Failure Events (HFEs) [14]. This method is based on a multi-disciplinary framework that considers both human-centered factors and plant conditions creating operational causes for human-system interactions [10]. The human-centered factors and influences of plant conditions are dependent of each other, which are combined to create a situation in which the probability of making an error can be estimated. Such a situation is said to have an Error-forcing Context (EFC). The primary shortcoming of this technique lies in the fact that it is unable to produce final HEP meaning that the direct outcome of this analysis cannot be quantified [15].

CREAM (Cognitive Reliability Error Analysis Method) is one of the best known second generation HRA methods, which offers a practical approach to both performance analysis and error prediction [16]. This method presents a consistent error classification system integrating all individual, technological and organizational factors, which can be used both as a stand-alone method for accidental analysis and as part of larger design methods for interactive systems. In this method, human error is not considered to be stochastic, but shaped by different factors such as the context of the task, physical/psychological situation of the human operator, time of day, etc. One of the main features of this method is its integration of a useful cognitive model and framework that can be used in both retrospective and prospective analysis [17]. CREAM is capable of providing the estimated HEP that can be used as part of overall system analysis. Compared to other HRA methods, CREAM seems more promising as an option to assess human performance for several reasons. First, it represents a second generation HRA method with improved applicability and accuracy compared to most of the first generation methods. It is able to extend the traditional description of error modes beyond the binary categorization of success-failure and accounts explicitly for how the (performance) conditions affect the performance. Secondly, it is originally developed from the Cognitive Control Model (COCOM)[*] and also uses it to organize some of categories describing possible causes and effects on human action. Last but not least, CREAM can be used for performance prediction since quantified results can be provided as the final outcome. This capability especially makes the integration of the CREAM-based non-technical component model with other technical component models possible, which is a critical requirement for modelling infrastructure systems.

In recent years, a wide range of modelling approaches, e.g., Agent-based Modelling (ABM), Complex Network Theory (CNT), System Dynamic (SD), have been applied to represent technical components infrastructure systems. However, modelling efforts regarding the representation of the human behaviours remain on the adoption of classical analytical approaches, e.g., probabilistic modelling method, using a combination of fault and event tree techniques, making the analysis of the human

[*] COCOM models human performance as a set of control modes: strategic, tactical, opportunistic and scrambled and proposes a model of how transitions between these control modes.

performance in a comprehensive way particularly difficult. Furthermore, it is not an easy task to integrate this type of model with other technical system models in case all components (technical and non-technical) of an infrastructure system need to be considered. Among these approaches, the ABM seems more promising.

In this paper, a generic modelling framework is proposed and presented. The framework consists of two parts: First, an analytical method based on the CREAM for human performance assessment, which includes five working steps. In this method, a knowledge-based approach is developed in order to assess PSFs in a more efficient way. Second, an ABM approach for the representation of human behaviours. Within this approach, the human operators and uses of infrastructure systems are modelled as agents with capability of interacting with other agents, e.g., agents representing technical components. Using this advanced modelling approach, the human performance is able to be assessed and corresponding human error can be calculated in real-time dynamically based on current simulation environment, e.g., current time, simultaneous goals, etc.

3. A CLOSER LOOK AT CREAM

CREAM is derived from the method of COCOM, the purpose of which is to provide the conceptual and practical basis for developing operator performance models. In both methods, the cognition is regarded as not only an issue of processing input(s) and producing a reaction, but also an issue of the continuous revision and review of goals/intentions [18]. Therefore, the cognition should not be described as a sequence of steps, but rather a controlled use of available competence and resources [16]. The basic assumption of CREAM is that human performance is an outcome of the controlled use of competence adapted to the requirements of the situation, rather than the result of pre-determined sequences of responses to events. Four characteristic control modes are defined in the CREAM method : *scrambled control*, *opportunistic control*, *tactical control*, *and strategic control mode* [16]. Instead of PSFs, the method of CREAM uses CPCs (Common Performance Conditions) to determine sets of error modes and probable error causes. Total nine CPCs are proposed by Hollnagel: *adequacy of organization, working conditions, adequacy of MMI (Man-Machine Interface) and operational support, availability of procedures/plans, number of simultaneous goals, available time, time of day, adequacy of training and experience*, and *crew collaboration quality*. Various levels are also assigned to each CPC. For instance, three (CPC) levels are assigned to the CPC "working conditions": *advantageous*, *compatible*, and *incompatible*. The main difference between the CPCs and the PSFs is that the CPCs can be applied at the early stage of the analysis to characterize the context for the task as a whole, rather than a simplified way of adjusting probability values for each event. Therefore, the influence of CPCs is closely linked to the task analysis. Advantage working conditions such as the level "compatible" (CPC level) of "working condition" may improve the performance reliability, while disadvantage performance conditions such as the level "incompatible" may reduce the performance reliability. If the performance reliability is reduced, operators could fail more often. Relations between all nine CPC levels and their expected effects on the performance reliability can be determined based on author's general knowledge and experiences. In most first generation HRA methods, it is always assumed that PSFs are independent. This assumption raises concerns since even a cursory investigation is able to show that it is not possible that all PSFs are independent to each other. This concern has been taken into consideration by most second generation HRA methods. In the CREAM method, all the CPCs have influences on each other. For instance, the CPC "working conditions" (e.g., ambient lighting, noises from alarms, interruptions, etc) have direct impacts on both of "number of simultaneous goals" and "available time". Improved "working conditions" can be assumed to increase "available time" and decrease "number of simultaneous goals". It is very important to take these dependencies into account when applying the CREAM method (see[16] for more information).

4. FRAMEWORK PART 1-1: AN ANALYTICAL METHOD

Human error is defined as "*Any member of a set of human actions or activities that exceeds some limit of acceptability, i.e. an out of tolerance action (or failure to act) where the limits of performance are* *defined by the system*" in [19]. In our daily life, the human error is extremely common since everyone could commit at least some everyday. However, the human error has become a cause of great concern to the reliability of interactive infrastructure systems, since most these systems depend on the interaction with operators in order to maintain their appropriate function. A general analytical method based on the method CREAM is proposed in this chapter. To demonstrate the feasibility and applicability of the proposed framework, SCADA system is used as an exemplary system. This method can be divided into five working steps:

- Step 1: Constructing event sequence
- Step 2: Determining COCOM functions
- Step 3: Identifying most likely cognitive function failures
- Step 4: Assessing CPCs
- Step 5: Determining failure probability

In step 1, a task needs to be specified and corresponding event sequence can be constructed. In this case, a simplified task of general alarm handling is selected (see [20] for more details about introduction of the alarm handling). The overall operation of the task (task 0) involves four sequential subtasks. First, operators need to check whether or not the alarm monitor system is ready to work properly (subtask 0.1). The monitor system could include devices such as monitors, alarms, etc. Then operators start to keep checking the monitor system regularly to ensure that the new generated alarm will not be missed (subtask 0.1.1). If a new overload alarm is generated and sent by corresponding devices to the alarm monitor system, operators will be notified meaning that this identified alarm will be handled (subtask 0.1.1.1). Finally, a control command will be sent by operators (subtask 0.1.1.1.1).

In step 2, all possible COCOM functions need to be determined for each identified subtask. The model assumes that there are four basic cognitive functions: *observation*, *interpretation*, *planning*, and *execution*. Each defined typical cognitive activity can be described in terms of which combination of these four cognitive functions it requires. For example, the "monitor" activity involves "observation" as well as "interpretation". Therefore, all subtasks (cognitive activities) identified in step 1 are assigned with corresponding COCOM functions. Furthermore, it is important to determine a dominant function if the defined cognitive activity involves more than one COCOM functions. For example, subtask 0.1 (ensure the monitoring system is working) is assigned with COCOM activity "verify" that involves two COCOM cognitive functions: "observation" and "interpretation". Based on the description of the alarm handling task, this subtask involves more "observation" function and less "interpretation" function. In this case, the "observation" is the dominant COCOM function. Table 1 lists all possible cognitive functions defined for each subtask and one dominant cognitive function of each subtask is highlighted in red colour.

Table 1: Determination of cognitive functions

Subtask	Goal	Cognitive activity	Obs	Int	Plan	Exe
0.1	Ensure the alarm monitoring system is working	Verify	•	•		
0.1.1	Monitor overload alarm	Monitor	•	•		
0.1.1.1	Identify a new overload alarm	Identify		•		
0.1.1.1.1	Send command	Execute				•

Obs: observation, Int: interpretation, Plan: planning, Exe: execution

For each cognitive function, generic cognitive function failures have been defined in [16]. It is possible to use all pre-defined cognitive function failures for each cognitive activity. However, in order to make the CREAM more practical in use, one most likely cognitive function failure should be identified and used. This can be done based on the understanding and knowledge of the corresponding task in step 3. For example, three cognitive function failures can be defined for the subtask 0.1: 1) the observation of a wrong object, 2) the wrong identification made, and 3) the observation not made. According to the description of this task, it is more reasonable to assume that the possibility of missing an overload alarm is higher. Therefore, the third cognitive function failure can be identified as the most likely function failure for subtask 0.1.

Step 4, assessing CPCs, is the essential step among all 5 working steps, which is also the most challenging step. The purpose of this step is to examine and assess the CPCs under which the corresponding task is performed. Some of these CPCs can be easily assessed, e.g., the time of day (day time or night time depending on the time when the corresponding task is performed), the number of simultaneous goals, while the assessment of some CPCs can be difficult, e.g., the adequacy of organization, working conditions. In order to simplify the overall assessments of CPCs, it is necessary to assign some CPCs with a fix level. It should be noted that increasing number of CPCs with fixed levels will affect the output accuracy of the model. The effects of the CPCs on performance reliability can be quantified using the weighting factor. For instance, in the case where the expected effect is "not significant", the weighting factor is set to be 1. In the case where the expected effect is "improved", the weighting factor can be set to be less than 1 meaning that the final calculated HEP will likely be decreased. Lower weighting factor value indicates better performance. For instance, the weighting factor for level of “compatible” of CPC “working conditions” can be set to 1 and the weighting factor for level of “incompatible” can be set to 2.

To determine the CFP[†], each identified most likely cognitive function failure is firstly assigned with a nominal CFP, which can be conducted in step 5 using the information from [16]. Then, these nominal CFPs are adjusted considering the effects of the CPCs using weighting factors obtained from step 4. Table 2 lists the adjusted CFP for each subtask including best case scenario and worst case scenario.

Table 2 Adjusted CFPs for cognitive function failures

Subtask	Task step or activity	Nominal CFP	Best case scenario		Worst case scenario	
			weighting factor	**adjusted CFP**	**weighting factor**	**adjusted CFP**
0.1	Ensure the alarm monitoring system is working	0.07	0.2	0.014	9.6	0.672
0.1.1	Monitor overload alarm	0.07	0.2	0.014	9.6	0.672
0.1.1.1	Identify a new overload alarm	0.01	0.25	0.0025	6	0.06
0.1.1.1.1	Send the command	0.003	0.2	0.0006	9.6	0.0288

The final CFP can be obtained by choosing the maximum one from all calculated adjusted CFPs using the Equation 1:

$$\mathbf{CFP}_{final} = \boldsymbol{max}(\mathbf{CFP}_i), \boldsymbol{i} = \mathbf{1}, \mathbf{2}, .. \boldsymbol{n} \quad (1)$$

Where $\mathrm{CFP_i}$ represents the adjusted CFP value and n represents the number of values calculated. In the case of best case scenario, three out of nine CPCs have "improved" effects on the performance reliability and none of the CPCs have a "reduced" effect ($\sum improved = 3, \sum reduced = 0$). The corresponding control mode is "Tactical" and the probability interval is from 0.001 to 0.1. The calculated final CFP, shown in Table 2, is 0.014, which falls into the interval. In the case of worst case scenario, one out of nine CPCs have an "improved" effect on the performance reliability and three of the CPCs have "reduced" effects ($\sum improved = 1, \sum reduced = 3$). The corresponding control mode is "Scrambled" and the probability interval is from 0.1 to 1. The calculated final CFP, shown in Table 2, is 0.672, which falls into the interval.

5. FRAMEWORK PART 1-2: A KNOWLEDGE-BASED APPROACH FOR CPC ASSESSMENT

As mentioned above, step 4 (assessing CPCs) is the essential step of the proposed analytical method, which are challenged by following reasons. First, it is difficult to set a numerical threshold, by which

[†] Within the CREAM method, the final error probability is also referred as Cognitive Failure Probability (CFP) instead of HEP.

the corresponding level can be decided. Second, the assessment depends on the knowledge and experiences related to the specific task. Furthermore, many other issues could also have direct effects on the assessment of specific CPCs. The challenges could be solved easily for assessment of some CPCs. For example, the CPC 'Time of Day" can be assessed by examining the current time of the model assuming that "*if current time is between 8 am and 20 pm, then the CPC level is set to Day Time. If not, then the CPC level is set to Night Time*". However, it is not an easy task to assess some CPCs. For example, both the number of current simultaneous tasks and time left for operators to handle one task could have significant influences on the assessment of "available time". In order to assess this type of CPC, a knowledge-based approach using the fuzzy logic theory is proposed and developed. Fuzzy logic theory, first developed by Zadeh in [21], almost four decades ago, has emerged over last several years as a useful tool for modelling processes which are too complex or fuzzy for conventional quantitative techniques or when the available information from the process is qualitative, inexact or uncertain [10]. Fuzzy logic fills a gap between purely mathematical approaches and purely logic-based approaches. Instead of requiring accurate equations to model real-world behaviours, fuzzy logic is capable of accommodating the ambiguities of real-world human language and logic with its inference techniques. Fuzzy inference systems (FIS), developed based on fuzzy logic theory, have been successfully applied in fields such as automatic control, data classification, expert system, and decision analysis [22]. Unlike other regular mathematical systems, the FIS is related to the classes with unsharp boundaries where the output is only the matter of degrees. It is primarily about linguistic vagueness through its ability to allow an element to be a partial member of set, so that its membership value can lie between 0 and 1 [23]. Using the approach of FIS for the study of the HRA is also not a new concept. In 2006, a modelling application of CREAM methodology based on fuzzy logic technique has been developed by Konstandinidou and his colleagues [10], which can be regarded as a pilot application demonstrating the successful 'translation' of the CREAM into the language of fuzzy logic.

In order to demonstrate the applicability of the knowledge-based approach assessing CPCs, "available time" is used as an exemplary CPC. It is assumed that this CPC is mainly affected by two parameters:

- Time left: in the task analyzed using this model, each overload alarm must be handled in a predefined time period. If operators fail to process on time, the overloaded line will be disconnected automatically in order to prevent the thermal damage to the transmission line. In this task, it is assumed that the moderate overloads can be tolerated for up to 20 minutes [24, 25].
- Number of simultaneous goals: if there would be a number of simultaneous alarms, then the time to handle some of these alarms will be delayed.

Assessing "Available time" through a knowledge-based approach using a FIS can be conducted as follows:

Input:

1) *Timeleft***:** the remaining time of each alarm to be handled
2) *Simgoals*: the number of simultaneous alarms that is required for operators to handle

Output: The cognitive level: "Available time": *adequate*, *temporarily inadequate*, *continuously inadequate*

Membership Functions (MF):

The MF essentially embodies all fuzziness for a particular fuzzy set [26]. The shape of membership functions used for both input and output are triangular**.** Three MFs are selected for both inputs, with linguistic values: "insufficient", "sufficient", and "more sufficient" for input "Timeleft" and "fewer than capacity", "match current capacity", and "more than capacity" for input "Simgoals". The range for each MF is shown in Table 3 and the graph is shown in Figure 1. It should be noted that the membership functions defined below are based on the understanding and knowledge of the analyzed task.

Table 3 Ranges of MFs for both inputs

Input	insufficient	sufficient	more sufficient
TimeLeft (min)	<10	>6 and <16	>10
Input	fewer than capacity	match current capacity	more than capacity
Simgoals	<3	>1 and <5	>3

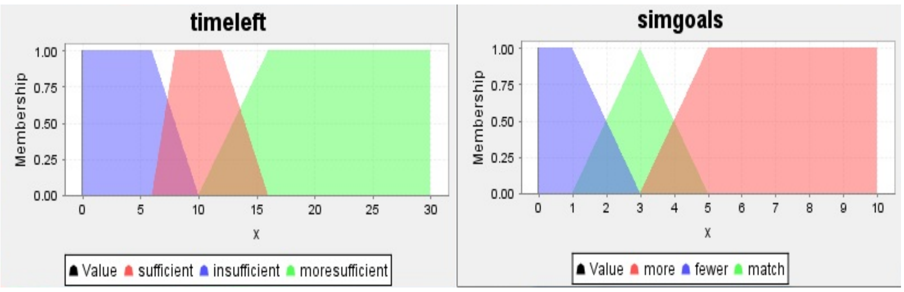

Figure 1 MF graphs of both inputs

Three output (consequence) functions are selected. The purpose of these functions is to determine the likelihood of the conclusion which is true, given a premise. The range for each MF is shown in Table 4 and MF graph is shown in Figure 2.

Table 4 The range of MF of output

Level of "Available time"	Continuously inadequate	Temporarily inadequate	Adequate
Consequence	<4	>2 and <6	>4

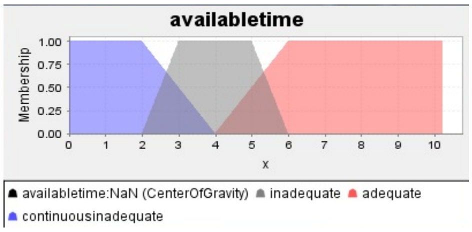

Figure 2 The graph of consequence of MFs

Rules:

Table 5 displays all fuzzy decision-making rules derived from knowledge base, developed based on the understanding and knowledge of the analyzed task. For example, the rule highlighted in the table can be read as "*If Time Left is sufficient AND the number of simultaneous goals is matching current capacity, then the level of "Available time" is set to adequate*"

Table 5 The rule table

Level of "Available time"		**Number of simultaneous goals**		
		Fewer than capacity	Match current capacity	More than capacity
Time left	Insufficient	Inadequate	Inadequate	Continuous inadequate
	Sufficient	Adequate	**Adequate**	Inadequate
	More sufficient	Adequate	Adequate	Inadequate

Defuzzification method:

Centre of Gravity (COG) method is implemented as the defuzzification method for combining all the consequences to make decisions, which is illustrated in the equation below. Basically this method calculates the weighted average of the centre values of the consequence membership functions (Equation 2).

$$u^{crisp} = \frac{\sum_i b_i \int \mu_{(i)}}{\sum_i \int \mu_{(i)}} \quad (2)$$

Where b_i denotes the centre of consequence membership and $\mu_{(i)}$ denotes the MF. In order to test the applicability of this knowledge-based approach, several test runs are performed. In the first test run, it is assumed that time left for operator to handle an overload alarm is 12 minutes and the number of simultaneous tasks is 2. Therefore, the inputs to the developed FIS are 12 for "Timeleft" and 2 for "Simgoals". The output of the FIS after the defuzzification is 7.24. All corresponding membership function graphs are shown in Figure 3. In this case, the level of "Available time" can be set to adequate. In the second test run, it is assumed that time left for operator to handle an overload alarm is 5 minutes and the number of simultaneous tasks is 4. Therefore, the inputs to the developed FIS are 5 for Timeleft and 4 for Simgoals. The output of the FIS after the defuzzification is 2.87. In this case, the level of "Available time" can be set to continuous inadequate.

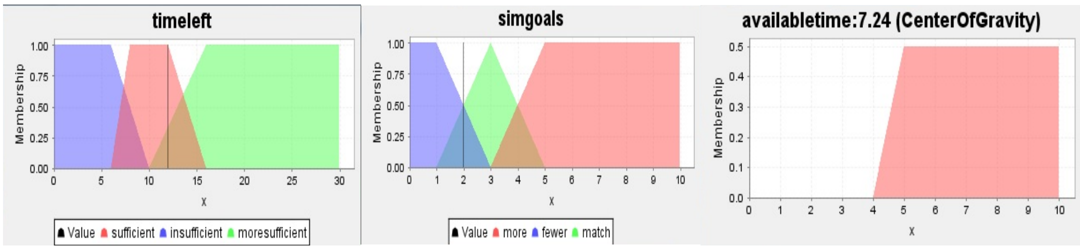

Figure 3 MF graphs of both inputs and the output for test run#1

One of advantages of integrating the approach of FIS into HRA lies in the fact that it provides a fundamentally simple way to handle complex problems without making itself exceedingly complex. It is straightforward, flexible, and easy to develop and understand. However, the FIS is a data-driven approach, meaning that the accuracy of the output depends on the quality of expert knowledge and experiences. Therefore, the membership functions, as well as developed rules, need to be carefully calibrated.

6. FRAMEWORK PART 2: MODELING HUMAN BEHAVIORS USING ABM

The ABM approach describes a whole system by its individual parts (bottom-up). Each component of the system is normally defined and modelled by an agent, capable to modify its own internal data (parameter and variable), its behaviours (function), its environment, and even adapts itself to environmental changes. An agent can be used to model both technical and non-technical components while different agents interact with each other directly or indirectly. One of the major advantages of this approach is the possibility to integrate various elements such as physical laws, Monte Carlo techniques, etc, into the overall simulation (see [27] for more details about the ABM).

In [24], a pilot human operator model is developed based on the ABM approach. The purpose of developing such a model is to assess the influences of human operator performance on the reliability of an Electricity Power Supply System (EPSS). The most critical shortcoming of this model is that the HEP is simply calculated by generating a random number between 0 and 1. Moreover, the model ignores the influences of the PSFs, restricting its applicability and accuracy. In order to overcome these shortcomings and be able to analyse the performance of the operator in a more comprehensive way, a further improved and agent-based human operator model is created using the proposed analytical method including the knowledge-based approach for CPC assessment. This model, developed as part of a SCADA model, is then integrated with an SUC (System Under Control) model in an experimental simulation platform that is built to assess interdependency-related vulnerabilities between two systems (SUC and SCADA)[‡] [28, 29]. During the simulation, if there is a request for the operator to handle an alarm, CPCs will be assessed automatically according to current simulation environment, e.g., time of day, simultaneous goals, etc, and corresponding CFP will be calculated as an input to other agents, e.g., MTU (Master Terminal Unit) agent from the SCADA model.

To simplify this assessment, it is necessary to make assumptions for following CPCs:

- working conditions (in control centre) are *compatible*
- the adequacy of organization is *efficient*
- the availability of procedures/plans is *acceptable*
- the adequacy of training and preparation is *adequate with high experience*
- the crew collaboration quality is *efficient*

[‡] It is assumed that the SUC and SCADA are parts of the EPSS.

This is the first effort to develop a human operator performance model that is capable of assessing CPCs dynamically using the ABM approach. Four CPCs are assessed and five CPCs are assumed to be fixed without further assessment due to limited data sources, which will affect the accuracy of output (CFP)[§] of this model. With the help of this model, several in-depth experiments have been developed for the identification and assessment of hidden vulnerabilities due to interdependencies, e.g., substation level single failure model experiment and small network level single failure mode experiment. The results from these experiments seem promising, highlighting the importance of human operator in the control center of the SCADA system. The lack of responses from human operators might not be the cause of failures of substation level devices, negative consequences caused by the failures of these devices could become worsen significantly. Figure 4 shows results from two case studies of the small network level single failure mode experiment. As seen from this figure, more components from both SUC and SCADA fail to function if performance of the human operator is assumed to be poor (see [28] for details and results of these experiments). This is only a pilot application demonstrating the possibility of assessing human performance using advanced modeling approaches. Motivated by these promising results, more experiments considering human operators as parts of the overall system are currently being developed, e.g., the experiment analyzing resilience related behaviors of infrastructure system.

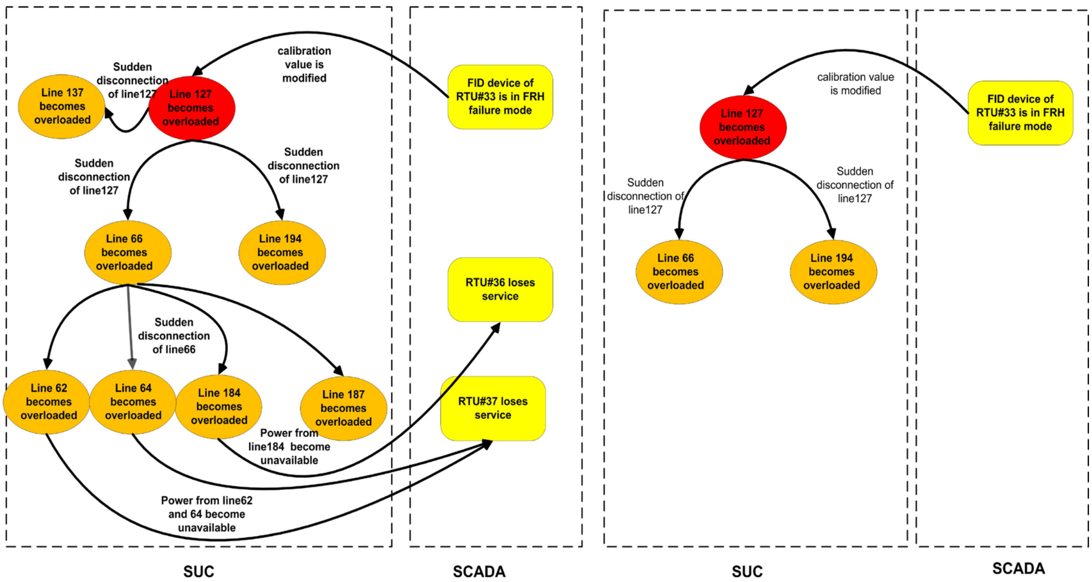

Figure 4 Affected components due to dependency between SCADA and SUC in two case studies [28]. Left: the human operator performance is assumed to be poor; Right: the human operator performance is assumed to be acceptable

7. OUTLOOK

HRA methods have been widely developed and improved during last several decades in order to provide a more applicable way to assess human performance. However, these methods are challenged by their inherent limitations, e.g., the lack of objectivity, inability to model tasks that consist of highly nested, concurrent cognitive activities, etc. These limitations hinder the possibilities of analyzing behaviors of infrastructure systems in a comprehensive way. In order to improve capability of current human operator model based on HRA method and take more contexts into consideration other than CPCs (e.g., emotion, learning ability, experiences, etc.), a conceptual agent-based hierarchy human model is proposed and currently under development, illustrated in Figure 5. This model consists of three levels. The upper level includes the components sensor and perception, which can be regarded as the input of the model. The information such as interactions with other agents, influence of environment, signal sent by technical components, predefined goals, are first processed by the sensor component. After that, the information will be further interpreted by the perception component. For

[§] The calculated CFP value in this case is between 0.0014 and 0.672.

example, if an alarm is received by the operators in the control room, it needs to be first received by the sensor and then interpreted by the perception in order to exact more detailed information, such as the severity of the alarm, etc. The middle level includes four components: physic status, emotion status, social status and cognition. These components contain parameters, state variables, states, rules, which can be used to determine states and behaviors of the agent. The lower level includes the components of behavior and actor, which can be regarded as the output of the model. The execution order of the agent action/behavior is determined by the component behavior, while the execution is carried out by the component actor. The information that has influence potentially on other agents/objects will be sent out by the component actor.

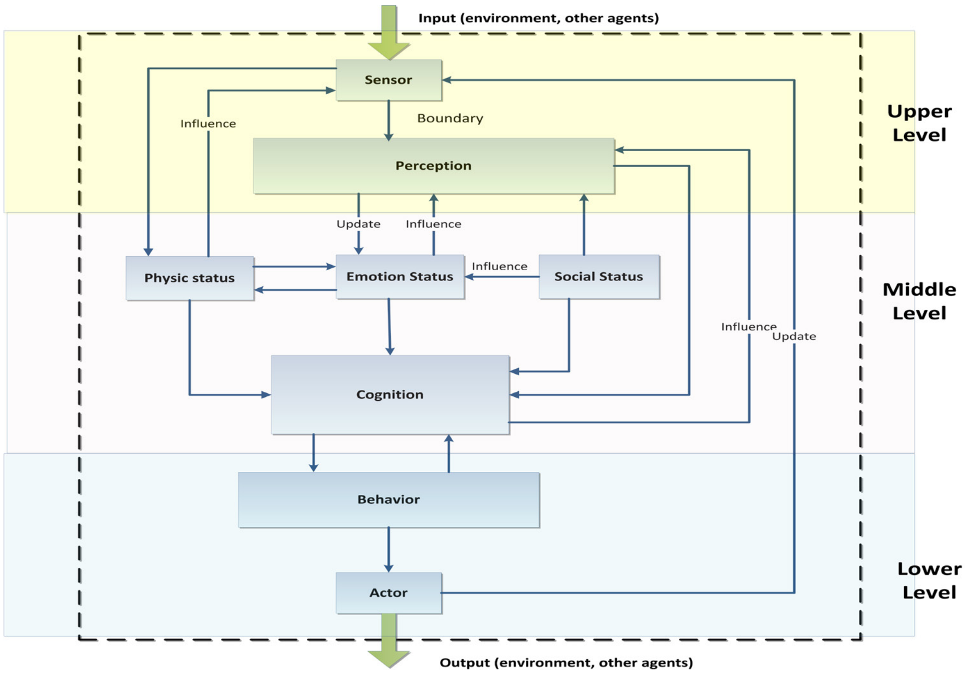

Figure 5 Overall structure of further improved agent-based human model

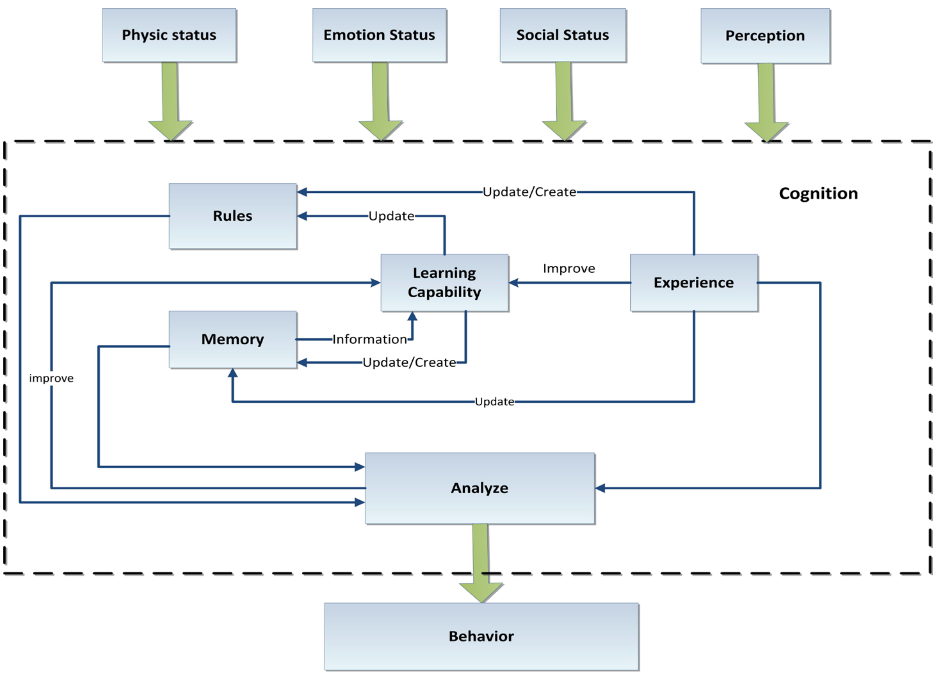

Figure 6 The structure of the cognition component

The cognition component is the most important component of the agent-based human model. The information received from outer environment and other agents will be mainly processed by this component and the corresponding behaviours will then be decided by this component. The overall structure of the cognition component is illustrated in Figure 6. Five subcomponents are included in it: rules, memory, experience, learning capability, and analyse. Compared to previous model, all the properties of an agent have been taken into account. Furthermore, this model is capable of making decisions according to previous experience, predefined rules, interaction with other agents, and outer environment.

8. CONCLUSION

Humans play an important role in the operations of vital engineered infrastructure systems. The lack of careful consideration of influences of these "non-technical components" of infrastructure systems often results in poor system performance and high costs. Their roles and impacts need to be strengthened by developing advanced approaches for human performance assessment that take more factors into consideration and focus more on the ways to analyze human behaviors efficiently in varying contexts. During the last decade, a number of research works have been developed and applied to analyze human performance and assess negative effects due to human errors, limited to errors of omission. Most of these works are based on the implementation of classical analytical approaches and seem not sufficient, which is the most critical shortcoming compared to research works focused on technical components of infrastructure systems. Full mapping of the complexity of infrastructure systems depends on continued development/ improvement of human performance models.

In order to explore the possibilities of adopting advanced modeling approaches for human performance assessment and bridge the gap between these two research communities, this paper proposes a generic modeling framework, including an analytical method for the performance assessment and an advanced modeling approach for the behavior representation, which is mainly focused on infrastructure systems, i.e. the electric power supply system. The analytical method is based on the second generation HRA method CREAM. In order to be able to assess CPCs more efficiently, a knowledge-based approach using the concept of Fuzzy Logic is proposed. The analytical method is further implemented as part of the human operator model, which is developed using the ABM approach.

The first results from the application seem promising, demonstrating the feasibility of the proposed modeling framework, not just due to its capability for representing the complexity of human performance of infrastructure systems, but also its modeling flexibility and adaptability. Thus, more application and simulation experiments based on this framework will be expected in the near future.

References

[1] W. Kröger, and E. Zio, *Vulnerable Systems*: Springer, 2011.

[2] V. M. Igure, S. A. Laughter, and R. D. Williams, "Security Issues in SCADA Networks," *Journal of Computers and Security,* vol.25, pp.498-506, 2006.

[3] S. A. Boyer, *SCADA supervisory control and data acquisition*, 3rd ed., Research Triangle Park: ISA, 2004.

[4] A. Ferscha, K. Zia, A. Riener, and A. Sharpanskykh, "Potential of Social Modelling in Socio-Technical Systems," *Procedia Computer Science,* vol. 7, no. 0, pp. 235-237, 2011.

[5] G. Baxter, and I. Sommerville, "Socio-technical systems: From design methods to systems engineering," *Interacting with Computers,* vol. 23, no. 1, pp. 4-17, 2011.

[6] C. W. Johnson, and C. M. Holloway, "A Longitudinal Analysis of the Causal Factors in Major Maritime Accidents in the USA and Canada (1996–2006)," *The Safety of Systems*, F. Redmill and T. Anderson, eds., pp. 85-104: Springer London, 2007.

[7] W. Kröger, and C. Nan, "Addressing Interdependencies of Complex Technical Networks," *Networks of Networks: The Last Frontier of Complexity*, Understanding Complex Systems G. D'Agostino and A. Scala, eds., pp. 279-309: Springer International Publishing, 2014.

[8] J. Wreathall, E. Roth, D. Bley, and J. Multer, *Human reliability analysis in support of risk assessment for positive train control*, 2003.

[9] J. Sharit, "Human Error and Human Reliability Analysis," *Handbook of Human Factors and Ergonomics, Fourth Edition*, pp. 734-800, 2012.

[10] M. Konstandinidou, Z. Nivolianitou, C. Kiranoudis, and N. Markatos, "A fuzzy modeling application of CREAM methodology for human reliability analysis," *Reliability Engineering and System Safety,* vol.91, no. 6, pp.706-716, 2006.

[11] Verein Deutscher Ingenieure, "Methods for quantitative assessment of human reliability," VDI 4006 Part 2, 2003.

[12] M. Kyriakidis, *Focal Report: A study regarding human reliability within power system control rooms*, Lab for Safety Analysis, ETH Zurich, Zurich, 2009.

[13] D. E. Embrey, P. Humphreys, E. A. Rosa, B. Kirwan, and K. Rea, *SLIM-MAUD: an approach to assessing human error probabilities using structured expert judgment. Volume I. Overview of SLIM-MAUD,* NUREG/CR-3518-Vol.1, 1984.

[14] M. Kyriakidis, *A scoping method for human performance integrity and reliability assessment in process industries*, ETH Zurich, 2009.

[15] J. Forester, A. Kolackowski, S. Cooper, D. Bley, and E. Lois, *ATHEANA User's Guide: Final Report*, U.S.Nuclear Regulatory Research, 2007.

[16] E. Hollnagel, *Cognitive Reliability and Error Analysis Method CREAM*: Elsevier, 1998.

[17] X. He, Y. Wang, Z. Shen, and X. Huang, "A simplified CREAM prospective quantification process and its application," *Reliability Engineering and System Safety,* vol.93, no. 2, pp.298-306, 2008.

[18] L. Bainbridge, *Building up behavioural complexity from a cognitive processing element*: London: University College, 1993.

[19] A. D. Swain, *Comparative evaluation of methods for human reliability analysis,* GRS-71, Institute for Reactor Safety, 1989.

[20] C. Nan, and I. Eusgeld, "Adopting HLA standard for interdependency study," *Reliability Engineering and System Safety,* vol. 96, no. 1, pp. 149-159, 2011.

[21] L. A. Zadeh, "Fuzzy logic," *Journal of IEEE Compute,* vol. 21, no. 4, pp. 83-93, 1998.

[22] R. L. Marcellus, "Evaluation of a nonstationary policy for statistical process control" pp. 89-94,1997.

[23] C. J. Harris, X. Hong, and Q. Gan, *Adaptive Modeling Estimation and Fusion from Data*, New York: Springer, 2002.

[24] M. Schläpfer, T. Kessler, and W. Kröger, "Reliability Analysis of Electric Power Systems Using an Object-oriented Hybrid Modeling Approach," in 16th power systems computation conference, Glasgow, 2008.

[25] W. R. Lachs, "Transmission-line overloads: real-time control," *IEE Proceedings,* vol. 134 (C), pp. 342-347, 1987.

[26] S. M. El-Shal, and A. S. Morris, "A fuzzy expert system for fault detection in statistical process control of industrial processes," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on,* vol. 30, no. 2, pp. 281-289, 2000.

[27] A. Tolk, and A. M. Uhrmacher, "Agents: Agenthood, Agent Architectures, and Agent Taxonomies," *Agent-Directed Simulation and Systems Engineering*, pp. 73-109: Wiley-VCH Verlag GmbH & Co. KGaA, 2010.

[28] C. Nan, I. Eusgeld, and W. Kröger, "Analyzing vulnerabilities between SCADA system and SUC due to interdependencies," *Reliability Engineering and System Safety,* vol.113, no. 0, pp. 76-93, 2013.

[29] C. Nan, W. Kröger, and P. Probst, "Exploring critical infrastructure interdependency by hybrid simulation approach," *Advances in Safety, Reliability and Risk Management*, pp. 2483-2491: CRC Press, 2011.