A Treatment of Risk Estimation Error due to Uncertainty and Subjectivity

Kenji Kawaguchi a*, Hiroshi Sato b
aComputer Simulation & Analysis of Japan Co., Ltd., Tokyo, Japan
bNational Defense Academy of Japan, Yokosuka, Japan

Abstract: In response to the increasing needs to manage the risk of complex socio-technical systems, a number of different approaches have been investigated over the decades. In particular, Probabilistic Risk Assessment (PRA) has been successfully utilized for engineering systems and human errors while System Dynamics (SD) has recently emerged as a useful tool to analyze organization factors. As a natural consequence of having two different approaches with two different scopes, several researchers developed hybrid models by utilizing these approaches. However, in the hybrid models, the risk levels of nuclear power plants (NPP) fluctuate in a narrow range near the acceptable level, which did not well represent the situation of Fukushima Daiichi nuclear accident. Therefore, in order to fulfill the discrepancy between reality and the model, this study expands the hybrid model of PRA and SD by introducing risk estimation errors due to uncertainty and subjectivity. Indeed, uncertainty and subjectivity have been well recognized as the Achilles heel of safety regulations. The proposed model has the ability to describe how the risk of Fukushima Daiichi NPPs went further beyond the acceptable level due to the errors resulting from the uncertainty and subjectivity. Also, with Multi-Agent simulation, it represents the phenomenon where the subjectivity results in group thinking causing information cascade and mutually assured delusion (MAD). This model will enable us to combine PRA and control theoretic models without optimistic assumption. The proposed model will facilitate our understanding of the possible degree of difference between the true risk and estimated risk in the face of high uncertainty, and it also may contribute in the examination of effective ways to manage people related to risk estimations.

Keywords: PRA, Control Theory, Subjective Probability, Information Cascade

1. INTRODUCTION

While traditional risk assessment methods, particularly probabilistic risk assessment (PRA), focus mainly on engineering systems themselves, several accident reports and analyses have strongly ascribed the occurrences of accidents to organizational factors [1-2]. Accordingly, researchers have attempted utilizing new methodologies to understand risks of socio-technical systems in terms of organization factors, such as Viable System Model (VSM) [3] or Systems-Theoretic Accident Model and Processes (STAMP) [4]. Although these new models have different focuses, a common property among them is that an accident is lead by control flaws within the systems. In other words, these new models are based on control theoretic view (see [5] for more information). In control theoretic view, a system may be controllable if the system is observable or diagnosable [6]. Therefore, observability is considered as a central concept in engineering control theory and the theory provides mathematical tools to examine observability for engineering systems (illustrated in [7]). However, there is no way to discuss the degree of observability regarding ‘risk level’ or correctness regarding the assessment of ‘risk level’, which is ultimately what we want to take control of.

On the other hand, the new models based on control theoretic view have preferable properties to describe organization and software reliability. Therefore, in order to take the advantages of both PRA and control theoretic view, STAMP model was combined with PRA [8-9]. However, the lack of approaches to represent the correctness of risk estimation has resulted in an arbitrary assumption in the hybrid models that the risk estimation does not have any errors. That is, it has been taken for granted in the hybrid models that once risk level increases, regulatory entities in ‘effective’ social loop structures can detect the increased risk level, then identify real culprits of the increased risk, and take countermeasures to reduce the risk.

However, of course, risk level controlled in an organizational system should not be assumed to be observable. In fact, risk levels have been estimated with PRA for decades, but it is well recognized that the uncertainty and subjectivity in the estimations of PRA are inevitable [10]. Both rare occurrences of observable events and complicated mechanisms of the risk-related events make uncertainty so great that meaningful
estimations are impossible without the use of subjective beliefs. Therefore, there is no support to assume that estimated risk levels reasonably represent the true values. Discussing errors in estimated risk levels may have been avoided partly because it may have been assumed that the errors could be treated as noise and hence they have no impact on the results in a statistical sense. But, based on ideas from social science and probabilistic theory, it can be assumed that the use of subjective views can result in group thinking [11] causing information cascade [12] and mutually assured delusion (MAD) [13]. Also, high degrees of uncertainty can end up with serious model incompleteness [14].

Summarizing the above, although the new control theoretic models have advantages and are able to be incorporated with PRA, with lack of discussion to represent risk estimation errors, the current hybrid model needs to assume the correctness of risk awareness. However, Fukushima Daiichi nuclear accident demonstrated that this assumption can be wrong and risk estimation error has a huge impact on the true risk as predicted by original control theoretic view of an accident. That is, if we could have known the true risk level regarding the high tsunami, for example, we could have had the chance to control the risk under the acceptable level (this does not mean the accident could have been prevented with certainty). Therefore, it is important to develop a risk estimation error model to complete the incorporation of PRA and control theoretic model. This attempt naturally asks the properties of the risk estimation error, which is not covered by original control theoretic accident models, and thus sheds light on a new research topic to manage risk of interest.

2. REVIEW OF LITERATURE

As the area of system safety is a very broad interdisciplinary one, this paper restricts itself to discuss only the major methodological views developed to understand and estimate the risks in complex systems, based on which this study will develop a risk estimation error model. Therefore, it explicitly avoids explaining philosophical views regarding the nature of accidents, such as Normal Accident Theory [1], and heuristics employed in design phases to handle risks, such as defense-in-depth [16]. The historical developments of the philosophical and design views are outlined in Saleh's review paper [5] and major challenges in reliability engineering are explained from an overall point of view in Zio's concise paper [17]. Also, this study covers some aspects of social and behavioral factors in safety engineering systems. But, a more detailed explanation and the current activity for this domain can be found in the final report of the multi-disciplinary research supported by the US National Defense and its product in the form of a book containing its participators' papers [18].

In light of the above scope, this section discusses the previous major methods related to risk estimations. The appendix first reviews probabilistic risk assessment (PRA) as the most widely utilized risk analysis technique, and then it will discuss a new approach, namely control/system theoretic approach, to solve the problems of traditional PRA. We also discuss a way to incorporate PRA with the new approach.

2.1. Probabilistic Risk Assessment and Risk-Qualitative Simulations

In the history of the safety community, the risks of complex engineering systems were attempted to be understood and handled mostly by limiting the scope into only the engineering systems themselves. Accordingly, since WASH-1400 [19] successfully demonstrated the power of Probabilistic Risk Assessment (PRA) to assess the risk of engineering systems in both qualitative and quantitative manners, it has been adapted to various industries. Indeed, the areas where PRA is utilized today are very broad: transportation, construction, chemical processing, energy, aerospace, the military, and even financial project planning and management [20]. One of the chief reasons why PRA is so powerful and popular is its ability to answer the following questions, which are essential in determining risks: (1) what can go wrong? (2) how likely it is? (3) what are the consequences? Those questions were first posted by Kaplan and Garrick [21] in 1981 and are still cited in current technical reports, such as the one published by the National Aeronautics and Space Administration (NASA) in 2011 [22].

1 This paper refers reasoning or results that are used for risk assessment, but do not have risk value (frequency or probability) as risk-qualitative.
2 In this paper, the term ‘PRA’ is used to refer Level 1 PRA as it is usually the case in broad safety community while nuclear sector utilizes Level 2 and 3 PRA as well. See [2] for more information.
The power of a model, in general, comes from the simplification of reality [23] and the strength of PRA is not an exception. PRA abstracts reality based on the idea that a chain of events leads to an accident and thereby people can focus on cause-effect relationships to assess risks. Because of this simplification, a PRA model can be effectively represented by Boolean functions. A Boolean Function is a function $f: x^n \{0,1\} \rightarrow \{0,1\}$, meaning that the combination of states (0 or 1) with regard to the $n$ number of components $x$ is transformed to a state of interest (0 or 1). In PRA, $x$ represents a set of events, some combinations of which can cause an accident. By using Boolean functions as a core engine, Fault-tree (FT) and Event-tree (ET) are used as tools to represent Boolean functions and manipulate them with stochastic inputs.

While PRA is sometimes criticized for neglecting the effects of non-linear causal effects and positive feedback [24], it is more accurate to say that although PRA eventually simplifies these effects, it may consider them in the first step. In fact, in order to examine an ET, simulations with physics models where non-linear dynamics take place have been conducted. On the other hand, parameters in a FT can be obtained by stochastic or deterministic physics models where complex dynamics play a great role. External events, such as earthquake, hurricane, or flooding, are modeled and abstracted to generate inputs for FT and ET. For human error, experimental data and expert opinions with human cognitive models are used to generate inputs for FT [25]. Moreover, for external human induced hazard (terrorism), Multi-Agent Simulations (MAS) was recently utilized to obtain hazard curves [26] that have potential to be used as inputs for FT and ET in the future.

The linear-causal approximation of traditional PRA does not capture the interactions among hardware, software, human, and organization well. But, because a traditional PRA model is only a simplified model as discussed above, by changing the degree of the simplification, some researchers have developed new methodologies for this problem, which is now referred to under the term "dynamic PRA" [27]. In dynamic PRA, chains of events are not explicitly modeled, but naturally emerge as the consequence of interactions of physics models (and possibly human cognitive models which are influenced by organizational factors). While dynamic PRA is still under development today, another new approach is emerging and it has a different view from PRA as discussed in the next section.

2.2. Control Problem and System Theoretic Approach and Its hybrid method

While traditional and dynamic PRA focus on causation close in time and space to accident progressions, several researchers, especially those from accident investigation and social science, suggested that organizational factors and software should be treated in a broader view. This is chiefly because accident investigation reports usually claim that the real culprit of an accident is an organization that has already collapsed, and in some sense the accident was deemed to happen before the physical deviation of technical systems started. Indeed, many accidents have been analyzed by VSM, with which analysts have attributed accidents to structural vulnerabilities and gradual degradation of performance in organizations. In addition, for software failure, it is well recognized that software fails only because of its design since software always performs the way that was specified in the design phase, unlike hardware [24]. Therefore, causes of software failure are already incorporated into them far before an accident happens. This reasoning that an accident is already set before the occurrence of an initial event in PRA led to the need for a new approach.

A new idea satisfying the need has emerged based on the view that accidents happen due to the lack of constrains that keep systems under control. The idea is analogous to engineering control theory, but it is used here not only for engineering systems, but also for any safety-related systems including operating organization, designing sectors, and regulation communities. The view was first refined by Rasmussen [29] who speculated that decision makers can be seen as controllers and their control determines safety in all levels of the system, such as the legislation level, management level, or operation level. In this view, safety state is defined as the situation where controls in every level are functioning properly. Based on this point of view, in 2004, Leveson [24] developed a new accident model, called Systems-Theoretic Accident Model and Processes (STAMP), which has the preferable property to account for organizational factor and software reliability. A significant difference between PRA and this new model is that STAMP focuses on constrains instead of events.
STAMP was adapted for accident analysis [30], and the combination of STAMP and System Dynamics (SD) modeling proposed by Leveson [24] also got a lot of attention. This is because the combination enables us to conduct computer simulations to understand the mechanism regarding degradations and constrains. In particular, a demonstration of the STAMP-SD model was carried out, with the support of NASA, by an MIT team including Leveson herself [28]. Figure 1 (a) is adapted from the project report, and it illustrates how a STAMP-SD model works. Symbols used in the figure follow the rule of SD. By examining the essential parts that affects the safety level, figure 1 (a) can be simplified into figure 1 (b).

Criticism for this new approach rose for two points. First, STAMP-SD model does not include detailed technical systems compatible with those in PRA [9]. Second, it is unclear what "risk level" means in STAMP-SD model [8]. Accordingly, since incorporating the results of other simulations is a natural function of PRA as discussed above, PRA was combined with STAMP-SD model [8-9]. In the hybrid model, organizational structure was represented with STAMP-SD model and the results of the model were factorized with some arbitrary external functions (as in [8]) or Bayesian belief network (as in [9]) to generate inputs for PRA model. Then, the output of PRA model was used as current "risk level" for STAMP-SD model. The results of the PRA-SD (with STAMP view) models showed that risks of socio-technical systems fluctuate within very narrow ranges around the desirable level ("safety goal" in figure 1 (b)). This is because the SD models used for these examinations were essentially similar to figure 1 (b) and because once risk level decreases, the gap between "perceived risk" and "safety goal" is calculated to generate negative feedback.

However, in order for these models to reasonably represent real world situations, “actual risk” and “perceived risk” in figure 1 (b) need to be similar, but Fukushima Daiichi nuclear accident demonstrated that it may not be the case.

3. DESCRIPTION OF RISK-ESTIMATION ERROR MODEL

3.1. A Generic Framework of Risk-Qualitative Model with Control Theoretic View

A generic framework of risk qualitative model with control theoretic view is drawn here to make the relationship among existing models and the proposed risk estimation error model clear. The framework is
shown in figure 2 along with an explanation of its elements. “Feedback for Controller” (C), which is the focus of this paper, is indicated in red in both the diagram and the list. Note that any model with the feedback structure via (C) is itself a risk-qualitative model, and so is STAMP-SD model and the models in this paper. This is because to use the feedback structure, we need to assign “actual risk” value as we do so in STAMP-SD model (figure 1(b)), but we cannot possibly know the exact “actual risk” value. Hence, what we can discuss with control theoretic view is about relative risk. However, with some factorization, one may incorporate such risk-qualitative results into PRA to calculate quantitative risk in the end (i.e., results of PRA-SD model would be utilized to generate inputs for PRA).

A risk-related model should explicitly show the scope of the model so that missing parts can easily be realized. Having the generic view above would be helpful to express which part is neglected or focused on in a risk model. In fact, PRA-SD model in [8-9], for example, can be seen as a version of this generic model where “Feedback for Controller” (C) and “External Physics Model” (A-2) are implicitly neglected.

The focus of this paper is “Feedback for Controller” (C) and hence this paper will utilize simplified models for “Risk-qualitative Model” (A) and “Risk-quantitative Model” (B). In the following sections, “Feedback for Controller” itself is modeled in a simple form (section 3.2), and then it is elaborated further by utilizing MAS (section 3.3).

### 3.2. A Risk Model with Simplified Risk-Estimation Error

This section introduces a PRA-SD hybrid model with simplified estimation error. In this model, SD is used for “Organization Model” (A-3) and PRA is utilized for “Risk-quantitative Model” (B-1) in figure 2. The results of other “Risk-qualitative Models” (A-1 and A-2) are assumed to be incorporated into the PRA model. This simplified model does not consider how the risk estimation error itself emerges and is strengthen via human cognitive process and interactions of people. That is, it is assumed that there is only one risk estimator whose mean values of errors do not change over time.

To illustrate the error model for PRA-SD method, simplified PRA model was adapted by [31] with a modification of adding a basic event that represents model uncertainty (figure 3). Along with the simplified PRA model in figure 3, risk estimation errors $\xi$ are modeled for every parameter $\theta$ of basic event A, B, C, and D as in the following:

\[
\hat{\theta} = \theta \cdot \xi 
\]

\[
\xi \sim \text{LogN}\left(\ln\{\lambda \cdot \exp(-0.5\sigma^2)\}, \sigma^2\right) 
\]

Here, $\theta$ is the true value of the parameter, $\lambda$ is the mean value of the error variable $\xi$, and $\sigma$ is the standard deviation of the error variable’s natural logarithm. Based on the equations above, an estimated parameter can be seen as a random variable as in the following equations: $\hat{\theta} \sim \text{LogN}\left(\ln\{\theta \cdot \lambda \cdot \exp(-0.5\sigma^2)\}, \sigma^2\right)$ and $E(\hat{\theta}) = \lambda \theta$.

Thus, when $\lambda$ is close to 1.0, the mean of the estimated value becomes closer to the true value and when $\sigma$ is close to 0, the estimator’s confidence of the estimated value increases. Mosleh and Apostolakis [32] suggested two error models for expert opinions, which are additive and multiplicative error models. The model above is analogues to the multiplicative error model and one can easily modify it to additive error model.

This simple error model would be able to show how real risk and estimated risk can diverge from each other and what type of decision-making rules are efficient to minimize the diversions. It is expected that the error
leads three types of hazardous states: 1) total risk level is estimated correctly, but risk dominant events are misunderstood, 2) risk dominant events are reasonably understood, but the risks of most events are underestimated, and 3) both the consistency among risks of events and their quantitative values are not properly understood.

3.3. Risk-Estimation Error Model with Estimators’ Interactions

In reality, risk estimation error would come from interactions among people including risk estimators themselves. Figure 4 indicates an image of the models including interactions of people. As every person is not sure as to the true real risk level due to a great amount of uncertainty, in PRA for example, people try to utilize their own beliefs and other experts' opinions as subjective probabilities [20]. Therefore, risk estimators can be affected not only by cognitive effects, such as "social influence" [33], but also by formally expressed experts' opinions, in the forms of bounded rational reasoning.

In order to model cognitive effects due to human interactions, it may be a reasonable option to use the following simple model.

\[ E(\hat{\theta}_{i,t+1}) = (1-\alpha)E(\hat{\theta}_{i,t}) + \alpha \left( \sum_{k \in \Omega_i} E(\hat{\theta}_k)/N \right) \]  

(5)

where \( \Omega_i \) is a set that represents other estimators who have influence on estimator \( i \), and \( N \) is the total number of those estimators. This model is indeed equivalent to Bayesian updating with normal distribution [35] where the ratio of uncertainty regarding an estimator \( i \) to that regarding anyone else is \((1-\alpha)/\alpha\) times greater than others', or more simply, 2) each person becomes affected only by the means of others' beliefs in accordance with the degree of \( \alpha \), which is a similar view to the one in [33].

For the influence of formal opinion exchanges, though, the simplification above may not be adequate. This is because experts and analysts in the industries tend to explicitly express their uncertainties, assume log-normality for parameters, and use some formal procedures to update their beliefs based on experts’ opinions. In fact, there are three distinct approaches utilized for the use of experts’ opinions: Psychological Scaling models, Non-Bayesian Axiomatic models, and Bayesian models. As Bayesian model was noted as the most robust technique for the utilization of expert opinions [36], this paper borrows the idea of Bayesian theorem to represent the influence of opinion exchanges. According to subjective Bayesian theorem, an estimator \( i \) can update his belief \( \hat{\theta}_i \) with real data \( r \) and opinions of others \( \hat{\theta}_k \) by assuming independency among them as in \( p(\hat{\theta}_i | r, \hat{\theta}_1, ..., \hat{\theta}_N) \propto p(r | \hat{\theta}_1, ..., \hat{\theta}_N) p(\hat{\theta}_i) \).

Then, the problem becomes how to determine the term \( p(\hat{\theta}_i | \hat{\theta}) \) and \( p(r | \hat{\theta}_1, ..., \hat{\theta}_N) \). While the latter term may be well represented with a latent Bayesian model based on Binominal distribution [37], this paper treats real data and opinions equally and assumes both terms follow log-normal distribution. With this assumption, the mean of the updated belief can be expressed by

\[ E(\hat{\theta}_i) = \exp \left\{ w_i \mu_i + \sum_{k \in \Omega_i} w_k \mu_k + \frac{1}{2} \left( \frac{1}{\sigma_i^{-2}} + \sum_{k \in \Omega_i} \frac{1}{\sigma_k^{-2}} \right) \right\} \]

where \( w_i = \frac{\sigma_i^{-2}}{\sigma_i^{-2} + \sum_{k \in \Omega_i} \sigma_k^{-2}} \)  

(7)
Here, $\mu$ is the logarithmic mean and equal to $\ln\{\lambda \cdot \exp(-0.5\sigma^2)\}$. Also, $\Omega_i$ includes the real data as well as opinions.

Equations (5) and (7) represent how interdependency among people can affect beliefs of each estimator in terms of social influence and opinion exchange respectively. The connections of people $\Omega$ can be expressed by their relational network $\Omega_s$ as abstracted in figure 4 and also by previous assessments $\Omega_p$ that make positive feedback to safety estimators in next generations. The relational networks $\Omega_s$ can be modeled by any network topologies. Particularly, we may use three basic network topologies: 1) complete network where all people are connected to all others, 2) random network [38], and 3) scale-free network [39]. The reasons to use the former two network topologies are to make examinations of results easier and to generate some mathematical predictions. The last one, scale-free network would be used as a way to account for a common property of the real networks. On the other hand, who will impact the next phases as previous assessments denoted by $\Omega_p$ can be people who leave risk estimations (ex. in the form of reports).

The effect of people’s interaction is suspected to depend on network structures and population distributions. If decision makers can utilize all people’s opinions at each step, all people should be as independent as possible. In this case, as long as all people need to be connected in some way, scale free network would be the most risk tolerance topology according to the results of failure propagation models [40].

### 3.4. Designing Effective Estimators’ Network

In the previous section, it was discussed that the effect of interactions among people would make the risk estimation problem more serious. According to Condorcet’s jury theorem [41], an averaged answer from a group of independent people would be more reliable than that from fewer experts under some conditions. One key condition is that people in the group need to be independent from each other; otherwise they are susceptible to information cascade [42]. In social science, instead of using the statistical term ‘independency’, the term ‘diversity’ has been emphasized as an important property to prevent cascade failure [43]. Although those findings seem to indicate that reducing the degree of interaction is needed to diminish risk estimation error, it is not as simple as it may sound. The reasons are two fold: the degree of interdependency would not depend on the degree of connections within social networks, and more importantly, reducing the interaction itself may not be beneficial in reality.

For the first reason, it was shown that more connections among systems could be less contagious for failure [40]. Likely, the effects of social influence and opinion exchange may increase by reducing the connections among estimators and experts. For the second reason, the interaction among people would play an important role in reality and reducing it may result in more hazardous situations. While a group of independent people may be wise as a whole, each individual is unreliable because their opinions are based on only their own beliefs. In reality, we cannot utilize opinions and estimations from thousands of people and hence we want each individual analyst and expert to represent a reasonable opinion. Indeed, we want estimators and analysts to learn from each other so that small samplings of estimations would represent not only his or her personal opinions, but also a reasonable belief based on a number of points of views. Therefore, the social influence and the effect of opinion exchange are beneficial in this sense.

To deal with the complex problem due to the two reasons above, we would be able to design effective networks by having a direct objective function; Genetic Algorithm (GA) can be used with the objective function to minimize the risk estimation error. The way to apply GA for network designs would be able to be adapted by [40] where optimization of networks was carried out for consensus speed. With the results of the optimization, we would be able to examine whether network structures in reality share the qualitative properties of the theoretically preferable networks.

### 3.5. A Demonstration of Simple Risk Estimation Error Model

In this section, a simple situation where each person chooses a true risk value or a false risk value is considered to illustrate the importance of risk estimation error. With Logit model, the probability of each person selecting the true risk value can be expressed by $P_{true} = 1/[1 + \exp\{-E(T_i - F_j)/\beta\}]$. (Now, mean error $\lambda$ is expressed by the distance between $T_i$ and $F_j$, each of which follows Gumbel distribution with parameter
Here, $T_1$ and $F_2$ correspond $\lambda$ in equation (2) and (7), and represent degrees of beliefs on the true risk value and false risk value respectively. Also, $\beta$ corresponds $\sigma$ in equation (2) and (7), and expresses the variance of the beliefs. For each person, the original expected belief to select the true risk value $E(T_1 - F_2)$ is affected with equation (5) to account for a cognitive effect (social influence). All simulation experiments in this section were conducted with 500 trials for 1000 time steps, with complete network, and with 1000 agents (people related to the risk estimation) whose $T_1 - F_2$ are initially assigned in accordance with normal distribution with mean 0 (average people are uncertain for $T_1$ and $F_1$) and standard deviation 0.05. The ratio of the number of people selecting the true risk value is indicated by a symbol $R$ hereafter.

The maximum and minimum ratios of the number of people selecting the true value in the end of the simulation are shown in figure 5 along with the different degree of social influence $\alpha$. In this experiment, selections by people became stable before 1000 time steps. Thus, the figure illustrates that a group of people can become stuck in false belief, even though their initial beliefs were neutral. Moreover, the comparison between the left and right graph in the figure shows that the effect of social influence becomes lower with a higher variance of beliefs.

Figure 6 represents the different patterns of group thinking behaviors and table 1 summarize the scenarios that emerged in accordance with varying degrees of social influence $\alpha$ and of beliefs’ variance $\beta$. The pattern (1) is the situation where no people change their choices once they have made up their minds, while in pattern (2) a few people randomly change their choices, but it does not affect the overall trend of group belief. In pattern (3), random changes of choices by some people affect the whole groups way of thinking and the belief of the group itself becomes random. In pattern (4), people are separated into majority and minority (ex. majority believes it’s safe), but in some occasions, the relationship reverses (ex. majority believes it is unsafe). Without a survey of people’s selections in the real world, we cannot begin to guess where we can be in the table. But, Table 1 may imply the possibility of managing estimation errors by trying to change dependency between people related to risk estimations and their confidence.

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>0.0-0.2</th>
<th>0.4-0.6</th>
<th>0.8-1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0001</td>
<td>0.5 with (1)</td>
<td>0.3-0.7 with (1)</td>
<td>0 or 1 with (1)</td>
</tr>
<tr>
<td>0.1</td>
<td>0.5 with (2)</td>
<td>0.3-0.7 with (2)</td>
<td>0 or 1 with (2)</td>
</tr>
<tr>
<td>0.25</td>
<td>0.5 with (3)</td>
<td>0.2-0.4 or 0.6-0.8 with (4)</td>
<td>0 or 1 with (3)</td>
</tr>
<tr>
<td>1.0</td>
<td>0.5 with (3)</td>
<td>0.5 with (3)</td>
<td>0.5 with (3)</td>
</tr>
</tbody>
</table>

*value in the cells indicates approximated ratio of people selecting the true risk value and (1)-(4) represent patterns regarding the behaviour of the ratio that are shown in figure 6. For example, “0.5 with (1)” means that the ratio became stable with 0.5 with pattern (1).

4. CONCLUSION

In this paper, risk estimation error models were introduced to complete the hybrid model of PRA and STAMP-SD. It was shown that the majority of people can become stuck in believing a false risk level in the face of uncertainty, and this may explain one aspect of Fukushima Daiichi nuclear accident. On the other
hand, this paper illustrates how the ways for a group to become stuck in a false belief depend on the degree of social influence and of the belief’s variance. This would imply the possibility of managing such phenomenon. This paper also discussed that GA would be able to be applied to search for near optimal network topologies for varying situations, based on which we may be able to discuss good qualitative properties regarding the relationship of published reports, analysts, experts and a regulatory committee.

It was explained by an audit practitioner [1] that considering all dimensions even within the organizational aspect, such as structural or behavioral ones, is important to understand an accident. In this sense, this study would reveal a new dimension that needs to be understood and handled if we are willing to minimize risks of socio-technical systems as much as possible.

Today, the risks of socio-technical systems are closely related to our lives. The West Coast blackout on August 10, 1996 and Fukushima Daiichi nuclear accident in 2011 demonstrated that misunderstanding the risks regarding socio-technical systems can result in serious disasters. To deal with the risks, people have tried to predict the hazardous events and assess the risks by utilizing knowledge as to the systems. However, there are always some unknown aspects of the systems and uncertainties in our predictions, which can result in risk estimation error. The proposed model will improve our understanding of the accident in terms of control theoretic view, by accepting the existence of uncertainty and subjectivity in our estimations of the risks.

References