An Information-Theoretic Approach to Resilience Assessment

レジリエンス性評価のための情報理論的アプローチ

ポリ　ジョナサン	東京大学		Jonathan POLI	University of Tokyo
出町　和之	東京大学		Kazuyuki DEMACHI	University of Tokyo

Rapid response processes are essential for enabling resilience, which provides a potential target for assessment. The response cycle is considered in three phases: (1) information transmission from system to Decision-Maker, (2) information processing, and (3) feedback action on the system. This research focuses on the first of the three phases, information transmission, and proposes a novel resilience assessment using an Information-Theoretic approach. From this point of view, resilience is considered to be the ability for a system to successfully communicate its state to the Decision-Maker in the face of unreliable information, particularly in the sudden degradation of transmission capability resulting from a disruptive event. A newly proposed metric of system resilience is the convexity in the curve of information entropy as a function of information loss. Convexity of the curve indicates the ability for a system to withstand disturbances without losing the ability to communicate and the ability to recover quickly if transmission is sufficiently compromised.

Keywords: Resilience, information entropy, Information-Action Cycle, communication, response, nuclear power plant systems

1. Introduction

Fukushima Daiichi showed that typical margins of safety can no longer be trusted to ensure that a system will remain in a safe condition. For situations beyond the predefined safety margin, Resilience offers a potential answer for ensuring system integrity and is primarily defined as the ability for a system to quickly recover from a disturbance. However, regarding a general framework for the means by which a system recovers quickly from a disturbance, attempts to agree on a suitable set of criteria have, on one hand, borne fruit, but on the other hand, the fruit has still not ripened. At this point, resilience still needs quantifiable criteria if it is to be used for engineering purposes. We propose to provide an alternative approach at resilience, starting with a more general interpretation of an underlying process a resilient system employs, and propose a novel metric of system resilience regarding one particular facet of the process, communication. Our hope is that the interpretation presented here contributes to a more solid foundation upon which to engineer resilience into systems.

Corresponding Author: Jonathan A. Poli、
〒113-8656、東京都文京区本郷 7-3-1 工 8 号館 802 号
E-mail: jpoli43@gmail.com

2. Resilience and Information

A fundamental process that allows resilience to emerge is illustrated in Figure 1. The system possesses information that

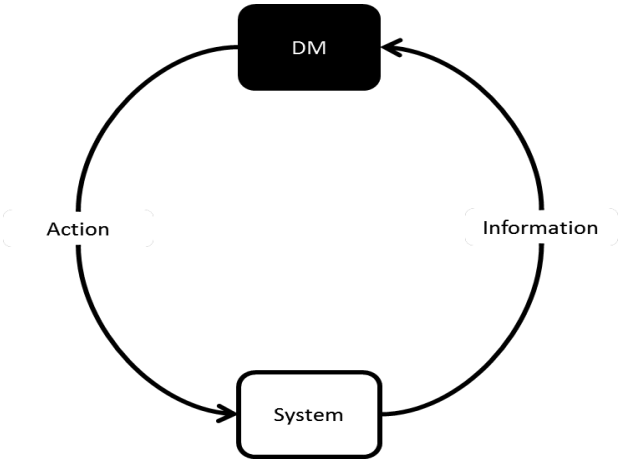

Fig.1 Illustration of Information-Action Cycle (IAC)

is extracted by a Decision-Making (DM) entity in order to be processed into an action and redirected back into the system. The action induces new information which is again extracted by the DM, and the cycle repeats itself in order to maintain a stable

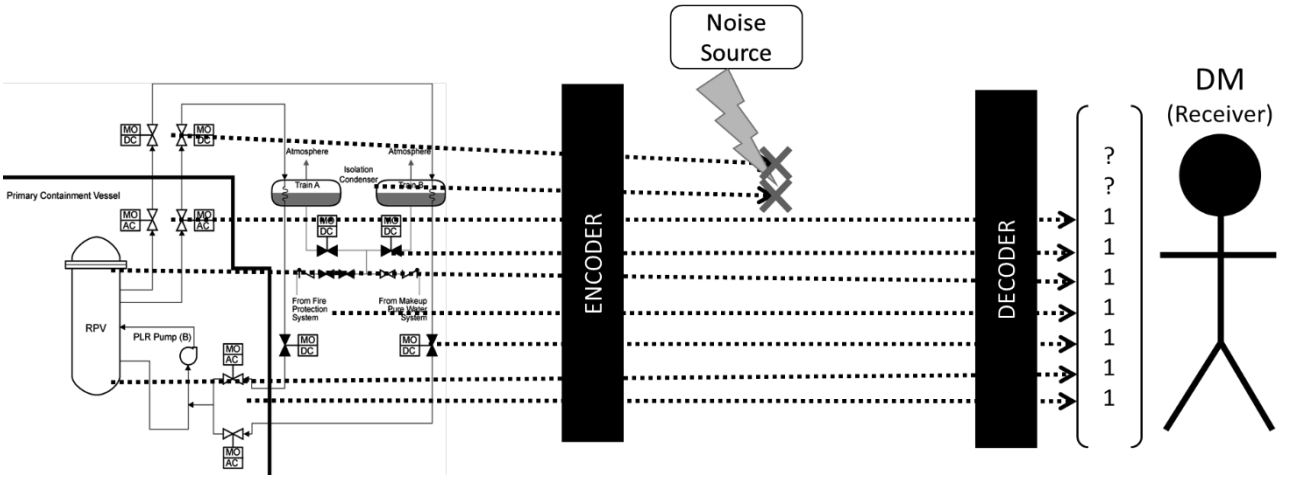

Fig. 2 The IC system and DM depicted as a communicating system. The sources of information each transmit their unique bit of information to the DM; the message consists of the collection of binary digits. The IC system shown to the right was excerpted from [3].

state and return to the stable state if perturbed sufficiently. This Information-Action Cycle (IAC) is a generalization to encompass feedback and feedforward processes, providing the possibility of being robust to, gracefully extending response capabilities during, and rebounding from a disturbance as well as sustaining adaptability over time [1]. Conversely, if the action or information processes break down, then the system loses the ability to recover to its stable state. One important distinction regarding the nature of resilience in this regard, *the cycle itself is not the indicator of a resilient system, rather the ability to maintain the IAC marks resilience*. The reasoning stems from the assumption that the cycle will invariably break down from some surprise combination of events; a resilient system, thus, must be able to cope with the consequences.

The IAC offers an underlying process that can be assessed for system resilience. In particular, the focus of this research is on the degradation of the information transmission step from the system to the DM. Most resilience assessments focus on the action step and although some have recognized the importance of information flow in the safety of the system, the phenomenon has yet to the treated formally as a resilience indicator.

3. Methodology for Calculating the Changes in Information of a System

The content thus far has been detached from any specific field in order to emphasize its general applicability. However, the methodology presented here will focus on the Isolation Condenser (IC) passive cooling subsystem from Fukushima Daiichi Unit 1. For this section, the IC system serves as an illustrative example.

Although the IAC depicts the DM extracting information, this step can be considered in the reverse and have the System communicating information to the DM. Reinterpreting the IAC in this manner allows for the information step to be characterized as a communicating system. The advantage lies in the ability for mathematical tools from Information Theory to be available as assessments for resilience.

Figure 2 depicts the IC system and DM in the type of communicating system based on the concept by C.E. Shannon in his seminal work [2]. The sender is the IC system which encodes the information on a set of sources that are transmitted through respective mediums; these pieces of information are decoded into a message that is finally interpreted by the receiver (DM). The IC system is a passive cooling system complete with two steam pathways, four valves per train, and two separate cooling tanks. The steam generated in the reactor pressure vessel (RPV) travels up into the cooling tanks, condenses, and passively flows back into the RPV. The heat removed converts the water in the condenser tanks to steam, which exits the plant as exhaust, the atmosphere being the ultimate heat sink. The collection of each component's operational state defines a unique state of the IC system.

The DM does not know the state of the IC system without the aid of instrumentation and control systems, or other cues that provide the status of the various components; examples from the IC system are valve closures, temperature sensors, water level

indicators, etc. provide information regarding which state the IC system is in. Each component is an information source that has a corresponding binary digit to indicate its operational state; 1 indicates operational whereas 0 indicates nonoperational. Thus, the collection of sources delivers a message, a sequence of binary digits that indicate the operational state of the IC system. For example, a sequence of all 1's corresponds to a fully operational IC system.

The length of the message is the number of sources located on the system, defined as N. Let V be total set of possible complete messages, messages which are transmitted by the system without any missing digits, the total messages is $V = 2^N$, the number of permutations of binary combinations N digits long. However, messages can be sent with a number of digits < N, which indicates that the message is incomplete. This situation arises from the "Noise Source" which serves as any disturbance that prevents a source or sources from communicating their part of the message (Fig. 2). Let U be the total set of possible messages complete or partial, thus, $V \in U$. At some time t, message $u \in U$ is sent by the system and received by the DM. The integer, $1 \leq i \leq N$, labels the position within u of digit i. Upon arrival at the DM, if $u \in V$ then the system state is fully defined and the action can be implemented readily. The assumption here is that complete messages indicate that the IAC is functioning. However, if $u \notin V$, then the message is incomplete. Let M be the number of missing digits in u, which leaves 2^M possible messages to account for, each missing digit doubles the possibilities.

Let Θ be the set of possible messages resulting from partial message u and $1 \leq k \leq 2^M$ label message k in Θ. At this point, each message is assigned a probability, π_k, of being the true message sent by the system. The uncertainty of the actual state of the system, denoted as random variable X, based on u is calculated by the information entropy [2]:

$$H(X) = -\sum_k \pi_k \log \pi_k \tag{1}$$

This equation provides the uncertainty, in units of bits, of the system state upon arrival of u. In the case where no other information is used to help differentiate between the possible messages in Θ as to which message is more likely the true message sent by the system, for all k in Θ, $\pi_k = 1/2^M$. The resulting plot of H is shown in Figure 3. The relationship is linear, each missing digit leads to a proportional change in the uncertainty. This means that for each missing digit, the DM must make a decision as to which message is the correct state of the system. In other words, partial messages inhibit the DM from being able to readily employ action, uncertainty reagarding system state inhibits the IAC. Hence, the linear trend is indicative of a system without resilience, if any information is lost the uncertainty increases.

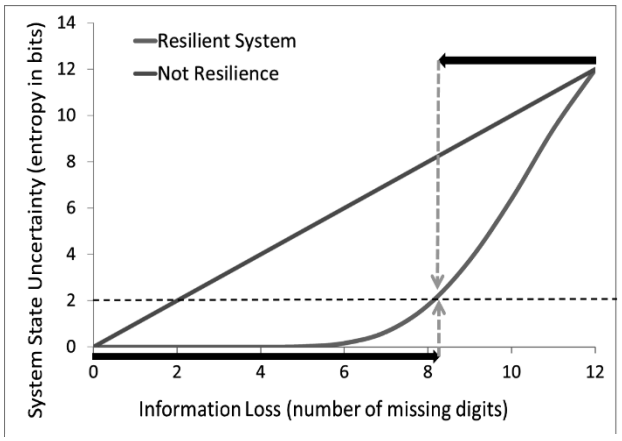

Fig. 3 Plot of system state uncertainty as a function of information loss. Information loss is represented by the missing digits in a message while uncertainty is measured by information entropy. The horizontal dashed line indicates a theoretical uncertainty threshold before the IAC breaks down.

4. New Resilience Metric Based on Information Entropy

However, a resilient system must maintain the IAC and manage the uncertainty such that the information transmission does not fail. One option is to quickly retrieve the missing information, a consideration which will be discussed later. However, take a scenario where the missing information is irretrievable, the possibility of directly reconstructing the message is no longer available. A resilient system then still must manage its uncertainty regarding system state. In order to do this, the π_k need to be adjusted in some manner that allows the DM to differentiate the most likely message being the sent by the system among all the possibilities. Thus, a resilient system must utilize the available information, the known sources, in order to alter the π_k.

In order to utilize the available information, the relationship between the information sources needs to be determined. For now, let the relationship between sources be defined as α, where

$-1 \leq \alpha \leq 1$. If $\alpha < 0$ then the two sources have an opposite relationship as in 1 and 0 or 0 and 1 rather than a parallel relationship when $\alpha > 0$. Each missing digit then has a probability of being 1 or 0, denoted as p and q respectively, which is based on α and the number of known sources s. When $u \in V$ arrives at the DM, for each missing digit, p_i is calculated by:

(2)
$$p_i = \frac{1}{1 + e^{-\alpha s}}$$

The probability $q_i = 1 - p_i$. Each message in Θ has a unique permutation of 1 and 0 combinations regarding the missing digits and the new π_k is calculated by multiplying the associated p_i and q_i combination.

The resulting curve is the convex curve in Figure 3 and is the novel indicator of resilience we propose. Convexity indicates that the system can resist changing its levels of uncertainty for a large domain of missing digits. At some point, however, the change in uncertainty happens rapidly and converges towards the linear trend. However, an alternative interpretation is that a resilient system at a high state of uncertainty can recover rapidly as well. Figure 4 shows the change in convexity resulting from changes in α. This indicates that the relationship between sources is a source of resilience. A discussion of the calculation of and meaning behind α is discussed in Appendix A. Thus, the convex behavior of the information entropy curve as a function of information loss in number of missing digits in a message models the behavior of a resilient system in regards to information transmission.

4. Conclusions and Discussion

Given the importance of information in maintaining the IAC, a resilient system is able to successfully communicate its state to the DM regardless if there is missing information or not. By utilizing the available information and the relationship of the known sources to the unknown sources, a system can maintain communication between system and DM when information is irretrievably lost. The information entropy of the communication between the system and DM offers a way to quantify the uncertainty of the system state in the perspective of the DM based on the incoming messages u. Low entropy levels correspond to successful information transmission and are desired for maintaining the IAC.

However, if too many sources are lost, the uncertainty inevitably increases and the system must have action measures in place in order to recover the lost information. Thus, a resilient system possesses both the ability to utilize available

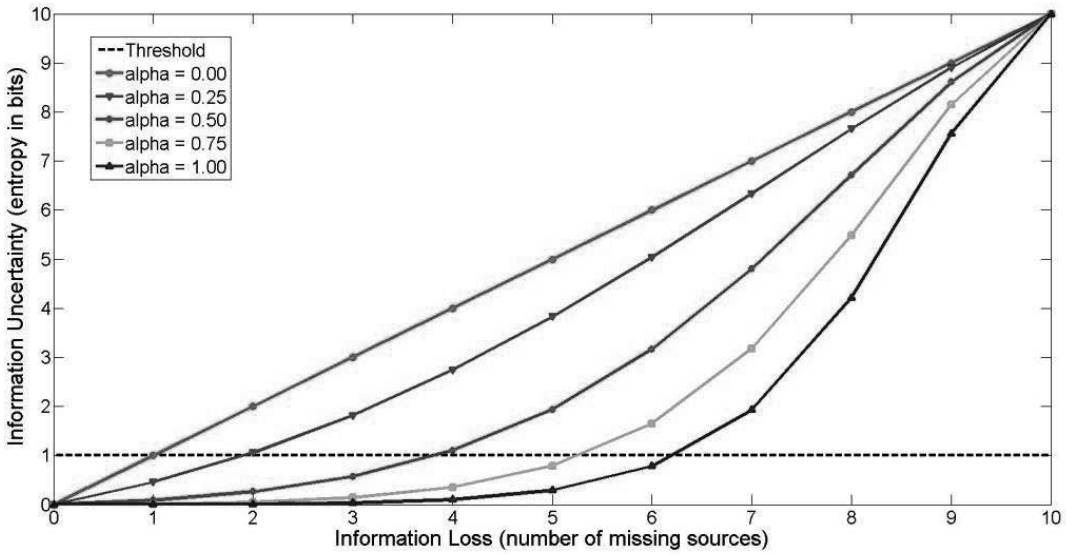

Fig.4 The information entropy curve as a function of information loss with increasing alpha values. As alpha increases, so does the convexity of the curve indicating that the more interconnected the information sources are, in the sense of being able to extract information of a missing source from a known source, the better equipped that system is to maintaining information transmission in the face of disruptions.

information effectively and the ability to recover lost information. In other words, both action and post-processing of available information is necessary to realize resilience. A possibility suggested by the interpretation of α as the relationship between two sources is that all sources are not equal in importance for information regarding system state. This study assumed that all the sources had equal α values, but this assumption is not realistic given that some sources possess more information about the state of the system than others. The relationship factor α indicates the importance of that particular source in containing information about other missing sources. Thus, sources with α values of greater magnitude are more essential in maintaining the information step in the IAC and deserve priority in recovery action as well as protection from loss.

Future work entails calculating the different α values for each source in order to identify critical components. This can be done using real time-series data or simulation data of the system in question. At present, a simulation model of the IC system is being constructed using Modelica, an open source language designed for modeling physical systems [4]. In conjunction with premade components from the ThermoPower library [5, 6], the goal is to have a representative model of the IC system. Using the model, various system dynamics can be modeled and used for the calculation of α. Furthermore, the ability to reconstruct messages using available information can be tested as well. The hope is to have the ability to understand the changes in uncertainty regarding the system state with losses in various information sources, defining critical pathways for uncertainty management. Once the information transmission capabilities are understood, the action protocols can be further elucidated in a complete management of the IAC, providing strides towards engineering a resilient system.

Appendix A: Calculation and Interpretation of α

The relationship between sources is a particular one where the information about one source is used to provide information about another. Another way to interpret this more rigorously is given missing source A, how much information is contained in known source B regarding A. The relationship is known as the mutual information between A and B and is calculated by:

(A1) $$I(A;B) = \sum_a \sum_b p(a,b) \log \frac{p(a,b)}{p(a)p(b)}$$

Where $p(a)$ and $p(b)$ are the marginal probabilities of A and B respectively, and $p(a,b)$ is the joint probability of A and B. This measures in bits the amount of information B possesses about A. If $I(A;B) = 0$, then the two sources are completely independent. To calculate α, the mutual information is divided by the entropy of the missing source [7]:

(A2) $$\alpha = \frac{I(A;B)}{H(A)}$$

Mutual information measures the reduction of uncertainty of A by knowing B. Divide that by the uncertainty of the individual source A, α is the departure from complete uncertainty regarding information coming from A. Dividing by the entropy of A offers two other advantages as well. One, α is dimensionless and can be used in the probability calculation. Two, by dividing by the entropy of A, the maximum value that $I(A;B) = H(A)$, thus normalizing the magnitude of α between 0 and 1. The marginal and join probabilities of A and B are determined using real-time or simulation data.

References

[1] Woods, David D. "Four concepts for resilience and the implications for the future of resilience engineering." *Reliability Engineering & System Safety* 141 (2015): 5-9.

[2] Shannon, Claude E. "A Mathematical Theory of Communication." *The Bell System Technical Journal* Vol. 27, 1948, pp. 379–423.

[3] American Nuclear Society. *Safety System Descriptions for Station Blackout Mitigation : Isolation Condenser, Reactor Core Isolation Cooling, and High-Pressure Coolant Injection*. 2011.

[4] The Modelica Association 2015, https://modelica.org/

[5] Casella, F., Leva, A., Modelica open library or power plant simulation: design and experimental validation. *Proceedings of the 3rd International Modelica Conference*, Linkoping, Sweden, 2003, 41-50.

[6] ThermoPower library home page. Available online at http://www.elet.polimi.it/uploads/casella/thermopower/.

[7] Gu, Guofei, et al. "Measuring intrusion detection capability: an information-theoretic approach." *Proceedings of the 2006 ACM Symposium on Information, computer and communications security*. ACM, 2006.

.