Resilience Engineering: State of the Art and Research Challenges

Kazuhiro Minami and Hiroshi Maruyama
Institute of Statistical Mathematics

July 2, 2012

Abstract

Many social infrastructures fail in an unexpected way, and thus it is important to make a system resilient such that it can recover from various damages in a dynamic and flexible way. In this paper, we summarize major research results on systems resilience in the literature and present future research challenges crucial to establish a solid foundation of resilience engineering.

1 Introduction

After the 3.11 earthquake, many people realized that there are events that cannot be reasonably anticipated. These "unexpected" events occur as an outside of the anticipated envelope (e.g., Tsunami of 14m high vs the anticipated max of 5.7m), or something completely unheard of (e.g., Tokyo subway gas attack in 1995). We recognize that these unexpected events do happen, but because they are "unexpected," we cannot prepare for the event and protect our systems. The only thing we can do is to give resistance that contain damages form the event locally and recover from that damage as quickly and as inexpensively as possible. We call this combination of resistance and recovery the *resilience* of the system.

Many researchers have recognized the importance of establishing a new research discipline concerning the resilience of complex systems to provide a set of general principles for building resilient systems in various fields. Although we have seen many examples of seemingly resilient systems in various fields, such as biology and computer science, researchers have not agreed on a common definition on resilience yet and it is thus not clear how we should adopt a strategy effective in one domain to systems in another. Therefore, we set out to establish a new research discipline what we call "resilience engineering," which provides a unified design principles for building resilient systems.

In this paper, we provide a brief summary of important research results addressing issues on systems resilience in various fields. We consider resilience as the combination of resistance and recovery abilities, and introduce strategies or mechanisms for achieving each property. Our goal here is not to give an extensive list of research projects studying the resilience of systems, but to introduce the reader to a new exciting research filed of systems resilience by giving several interesting research ideas. We finally describe fundamental research questions in this research field and discuss possible research directions. We refer the reader to our companion paper [10] for the research agenda of our research group.

2 Related Research

In this section, we cover several research topics related to resilient engineering. We first introduce the definition of resilience and describe several techniques for building resilient systems.

2.1 Definition of resilience

The concept of resilience appears in various disciplines ranging from environmental research to materials science and engineering, psychology, sociology, and economics. In this paper, we consider Bruneau's definition [6], which could give us a precise quantitative metrics of resilience. Bruneau considers a situation where a system's quality degrades abruptly at time t_0 due to some unexpected event and fully recover to the original state at time

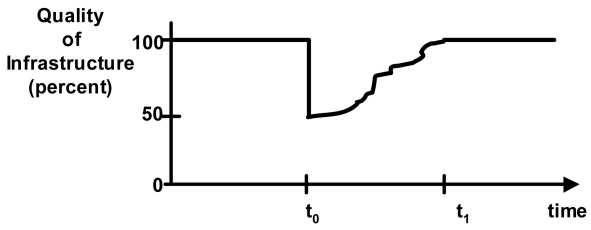

Figure 1: Bruneau's definition of resilience

t_1, as shown in Figure 1. If we denote by $Q(t)$ the quality of the system at time t, the resilience of the system is measured as follows:

$$\int_{t_0}^{t_1} [100 - Q(t)]dt$$

As the measured triangle area gets smaller, the system becomes more resilient. That is, there are two dimensions concerning the resiliency of the system.

- Resistance (reduced service degradation from failures at time t_0)
- Recovery (reduced time to recovery (i.e., time interval between t_0 and t_1))

We next describe representative strategies for achieving goals in each direction.

2.2 Resistance techniques

Redundancy is a common technique in engineering for increasing the availability of a system under the presence of various component failures. If a system has multiple components for the same functionality, the system can continue to provide a service when some of the duplicate modules fail. For example, RAID (redundant array of independent disks) uses multiple disks to realize a single logical disk such that the system still restore data when some of the disks fail.

Many complex engineering and biological systems are modeled as a network in which nodes are connected via links. Such network-based systems ensure the availability of services by maintaining redundant paths among any two components of the system in the network. The most successful example is a decentralized network management in the Internet. When a certain path becomes unavailable due to a failure of a router or a communication link, the ICMP protocol enables routers on the network to reconfigure their routing tables and establish alternate paths.

Barabasi [2] shows that network-based systems that possess the *scale-free* property are extremely robust against random failures of system components. The probability that any node in a scale-free network is connected to k other nodes is proportional to $1/k^n$. Thus, a scale-free network maintains a relatively large number of *hub* nodes with a degree that greatly exceeds the average, and those hub nodes make the network significantly robust against random failures of nodes and links. For example, any two nodes on the Internet can sustain connectivity between them even if as many as 80 percent of randomly selected routers fail. Therefore, to ensure the scale-free property in a network-based system is considered to be a good design principle to make the system resistant against random failures.

Diversity is another commonly applied principle for fault tolerance and it complements the weakness of *pure* redundant strategies against systematic attacks. Littlewood [9] considers the effectiveness of redundancy and diversity in computer security and argues that to provide redundant software components of the same implementation is not effective against a coordinated attack exploiting the same vulnerability; if one component fails with a certain attack, the attackers can easily compromise the other components in the exactly same way. Therefore, it is important to provide multiple different versions of the same subsystem. Littlewood discusses several strategies, such as separation, forced diversity, and tailored diversity, to ensure diversity among software components of the same functionality.

In the context of evolutionary biology, Kimura's neutral theory [7] claims that the great majority of evolutionary changes at the molecular (DNA) level are caused not by Darwinian selection but by random fixation of selectively neutral mutants. Such biological diversity at the DNA level could possibly help species survive environmental changes or sudden catastrophic events since the diversity of the species ensure that there exist some population group whose members possess qualities suitable to the new environment.

2.3 Recovery techniques

Recovery, which is usually an expensive strategy for physical systems, has been studied in computer systems since rebooting a computer system is a relatively cheap process.

Patterson [5] introduces the notion of Recovery-oriented computing (ROC), which claims that we should focus on Mean Time to Repair of a system rather than Mean Time to Failure. ROC takes the perspective that hardware faults, software bugs, and operator errors are facts to be coped with, not problems to be solved. Patterson suggests a technique of rebooting only some modules of the whole system to reduce the penalty of rebooting the whole system. He also mentions that to rebooting modules periodically is an effective way to removing latent errors whose accumulation could eventually lead the system to a fatal failure.

Checkpointing [8] is a common recovery techniques in database and distributed systems communities. Logging with checkpoints ensures that a system can restore a previously *consistent* state after a failure.

3 Research challenges

In this section, we discuss several possible research directions addressing open issues to achieve the vision of resilience engineering.

3.1 Centralized vs. decentralized

Many complex systems consist of numerous components interacting with each other in a decentralized way, and to modularize a large system into smaller independent components seems to be a good design principle in order to contain a damage from a failure in a limited area.

However, Bak [1] shows that many decentralized systems that are modeled based on cellular automaton naturally reach a critical state with *minimum* stability without carefully choosing initial system parameters and that a small disturbance or noise at the critical state could cause *cascading* failures of the system leading to a large disaster, such as Northeast blackout of 2003.

Although subsequent research shows that there are various natural or artificial systems, such as earthquakes, DNAs, and stock exchange prices, which can be explained well with the notion of critical points, there is very little research about avoidance of critical points in complex systems. In ecological biology, to perform small destructions to an environment is known to improve the sustainability of the ecological system, and we might need to have such centrally coordinated interventions to a decentralized system in order to avoid critical points. We plan to investigate such tradeoffs between centralized and decentralized approach in the future.

3.2 Dimensions of resilience

As we discuss in Section 2.2, a system based on a scale-free network is robust against component failures when the connectivity among components correspond to service availability. However, when we consider a containment of a spreading virus through a network, such connectivity becomes a vulnerability of the system. We need to investigate whether there is a common property of resilience for various requirements or we need a way to dynamically switch a "resilience" mode of the system.

The conflict of resilience requirements appears in evolutionary biology. There are two possible ways to consider the resiliency of a biological species. One is the resiliency of the species in the process of evolution. We consider that some species are resilient if its descendant survives in the future generations. The other is the resiliency of an individual in the species, where we consider an adaptability of an individual to its environment during its life span.

Each individual of the species could have a conflicting requirement from that of the species as a group. Although group evolution theory [11] provides a coherent theory for resolving such conflict in evolutionary biology, we should explore design principles for resolving dimension of resilience in various domains.

3.3 Reasoning uncertainty

Assuming that all possible states and events of a given system are known in advance, the notion of K-maintainability [3] precisely defines the notion of resilience. We say that a system is K-maintainable if, for any non-normal state of the system, there exists a sequence of actions (i.e., events controllable by a system administrator) that move the system

back to one of the normal states within k steps. However, to analyze a system based on this definition requires us to know in advance all possible events, some of which could be totally unexpected. Therefore, it is not clear whether a model checking approach is applicable to evaluate the resiliency of a system with an incomplete specification. We, therefore, expect that reasoning techniques dealing with various uncertainty of a system model [4, 12] be a promising tool to explore this research space.

4 Conclusions

In this paper, we survey the state of the art of research related to resilience science and show that this research field is truly interdisciplinary. We also discuss some of the most fundamental research questions in this field.

Although our survey is far from an extensive list of relevant research, we hope that we convince the reader that systems resilience is an exciting new research field where researchers must address interesting and challenging problems by teaming up with researchers in other disciplines.

Acknowledgments

This research is supported by grant from the Transdisciplinary Research Integration Center of Research Organization of information and Systems in Japan.

References

[1] Per Bak, Chao Tang, and Kurt Wiesenfeld. Self-organized criticality: An explanation of the $1/f$ noise. *Physical Review Letters*, 59:381–384, Jul 1987.

[2] Albert-Laszlo Barabasi and Eric Bonabeau. Scale-free networks. *Scientific American*, 288:50–59, 2003.

[3] Chitta Baral and Thomas Eiter. A polynomial-time algorithm for constructing k-maintainable policies. In *Proceedings of 14th International Conference on Automated Planning and Scheduling*, 2004.

[4] Hei Chan and Adnan Darwiche. On the Revision of Probabilistic Beliefs Using Uncertain Evidence. 63:67–90, 2005.

[5] David Patterson et. al. Recovery Oriented Computing (ROC): Motivation, Definition, Techniques. Technical report, Berkeley, CA, USA, 2002.

[6] Michel Bruneau et. al. A Framework to Quantitatively Assess and Enhance the Seismic Resilience of Communities. In *Earthquake Spectra*, volume 19, 2003.

[7] Motoo Kimura. DNA and the Neutral Theory. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 312(1154):343–354, 1986.

[8] Richard Koo and Sam Toueg. Checkpointing and rollback-recovery for distributed systems. In *Proceedings of 1986 ACM Fall joint computer conference*, ACM '86, pages 1150–1158, Los Alamitos, CA, USA, 1986. IEEE Computer Society Press.

[9] B. Littlewood and L. Strigini. *Redundancy and Diversity in Security*, volume 3193 (9th European Symposium on Research in Computer Security, Sophia Antipolis, France – ESORICS '04) of *LNCS*, pages 423–438. Springer, 2004.

[10] Hiroshi Maruyama, Katsumi Inoue, Hiroe Tsubaki, Hiroshi Akashi, Hitoshi Okada, and Kazuhiro Minami. Systems Resilience. In *Forum on Information Technology*. Information and Systems Society, 2012.

[11] Martin A. Nowak. Five rules for the evolution of cooperation. 314:1560–1563, 2006.

[12] Chiaki Sakama and Katsumi Inoue. Abduction, unpredictability and garden of eden. In *Proceedings of Model-Based Reasoning in Science and Technology (MBR)*, 2012.