E-Government Infrastructure II: Survivability Issues and Challenges

James B D Joshi*, Suroop M. Chandran*, Aref Walid#, Arif Ghafoor@,

*School of Information Sciences,
University of Pittsburgh
#Department of Computer Science, @School of
Electrical and Computer Engineering; Purdue University

INTRODUCTION

An electronic government (e-Government) is essentially an amalgam of interconnected heterogeneous information systems belonging to both government agencies and public and private sectors with a goal of modernizing the government's highly fragmented service-centric information infrastructure by improving information flow and the decision-making process (Joshi, 2001a). The e-Government environment also embeds the nation's critical infrastructures, that are required for providing the nation's basic services to the citizens (PDD, 1998), such as energy, telecommunications, banking and finance, and transportation facilities. The intricate connectivity of systems and their increasing dependence on IT dramatically magnifies the consequences of damages resulting from even simple system faults/accidents and intrusions, as well as natural events (fire, earthquakes, etc.), also collectively called *disruptions* (Ellison, 1997). A key challenge for such an infrastructure is to ensure continuous service availability to prevent financial losses, loss of prestige, endangerment of citizens' lives, and disturbances in national socio-psychological structures adversely effecting governance and democracy (Ellison, 1997; Gibbs, 1994; Moore, 2001). While it is essential that the e-Government infrastructure is resilient to disruptions, an even bigger concern is the protection of critical infrastructure components within the e-Government. In essence, the e-Government infrastructure should have the capability *to provide services in a timely manner, irrespective of disruptions*, a capability known as *survivability*.

E-GOVERNMENT SYSTEMS SURVIVABILITY

The e-Government survivability infrastructure should support both the intricate interdependence of government programs at different levels and between government and the private/public sectors, and address the need for continuity of its services in presence of disruptions. While such disruptions are inevitable in an e-Government, key to its success lies on the effectiveness of mechanisms for detecting and responding intelligently to disruptions, which is a daunting challenge. Intelligent distributed capability is required to detect and counter both structured and unstructured disruptions that can be either in the

form of intrusions or faults. Intrusions refer to the illegal access to a system by an intruder, whereas faults refer to the causes of physical failure of a system. Intrusions can be detected with the help of intrusion detection systems (IDS). IDSs report *anomalies* in behavior or recognize intrusion *signatures*. Faults can be detected but more importantly, methods for fault tolerance have to be implemented in the system. Fault tolerance is the ability of a system to withstand physical failure.

A survivability system needs to employ a combination of intrusion detection/prevention and fault tolerance methods. Separation between faults and intrusions, which have been hitherto studied separately, does not leverage the synergy existing between the two areas. This increases the overall cost of deploying measures against them, as well as the complexity of the overall system. Newly emerging coordinated, distributed intrusion detection techniques, coupled with data mining or stream mining techniques show promise in improving the survivability capability of a large infrastructure like that of an e-Government system by facilitating real-time detection of and responding to disruptions.

Disruption Categories for E-Government Systems: Disruptions to e-Government services can be divided into two categories – *cyber disruptions* and *critical infrastructure disruptions*. Cyber-disruptions include cyber-terrorism, like NIMDA and the Code Red worms, and information warfare. Potential "*info weapons*" that can be used to launch an attack on an e-Government include computer viruses, logic bombs, worms, Trojan horses, etc. (Alexander, 1999; Denning 2001; Garfinkel, 1997). Various attacks on systems include denial of service attack, virtual sit-ins and blockades, rootkits, etc. (Denning 2001). The attacks using these malicious tools range from simple hacktivism, which refers to active hacking activities with the intent to disrupt normal operations but not causing serious damage, to the more damaging *cyber-terrorism* and *information warfare* (Alexander, 1999; Denning 2001), which have become growing concerns post 9/11 era. Information warfare refers to the large scale malicious activities launched by independent individuals or attackers hired by terrorists or belonging to rival countries. Cyber-terrorism is a more dangerous form of cyber-disruptions that can cause severe damage to the nation's systems (Denning, 2000). Even a simple, hour-long coordinated hacking activity that affects the country's air traffic system, a critical infrastructure, can have very drastic consequences for government operations. In a few years the cyber-threats to the country is expected to be worse than the physical threat (Alexander, 1999).

Critical infrastructure disruptions could be some malicious attack, accident or disaster causing critical infrastructure malfunction, which becomes a national concern. Protection of critical infrastructure is an important issue, because any disruption in their functioning would cause nation-wide chaos, for instance, the North-East Blackout of 2003 in the United States and Canada - a power failure over the Northeastern regions of the United States and Canada in 2003 that caused many systems dependent on the electrical grids to fail disastrously. The damage was estimated at almost US$ 5 billion (Anderson, 2003).

Table 1 shows various threat levels and the criminal intent behind them (Alexander, 1999). At the highest level, we see national security threats, which are essentially aimed at the nation's critical infrastructures. Threats common to both government and non-government agencies include cyber-terrorism and e-espionage. Finally, there are frequently occurring hacking incidents that can create huge losses within an e-Government

environment. An alarming issue is the lack of awareness and ability to identify cyber-threats. Newer spamming and phishing attacks make survivability function more difficulty to implement (GAO, 2005).

Table 1. Threats and their intent (Alexander, 1999)

Threat level	Actor	Intent
National security threats	Information Warrior (Cyber-soldier)	Reduce decision making capability at the national level, National chaos and psychological terror
	National intelligence (Cyber-spy)	Information leakage for political, military and economic advantages
Shared threats (government & Private sector)	Cyber-terrorist	Visibility/publicity, chaos, political changes
	Industrial espionage	Competitive advantage
	Organized crime (Cyber-crime)	Revenge, retribution, monetary gain, institutional/political change
Local Threats (Hacktivism)	Institutional hackers	Monetary gain, thrill/challenge, publicity/prestige
	Recreational hacker	Thrill, challenge

At present, there is no nationally coordinated defense and survivability capability to detect and counter strategic and well-coordinated act of cyber-terrorism against the nation and to ensure the continuity of e-Government services under cyber-siege. The US National Infrastructure Protection Center (NIPC) is a program started by the Clinton administration in 1998 with an intention to maintain public and private sector infrastructure from disruptions of any sort and perform vulnerability checks regularly as preventive measures. Other nations such as Canada (PSEPC) and New Zealand have also taken to emergency preparedness and critical infrastructure protection. The Critical Infrastructure Protection project focuses on the impediments to the security and protection of the assets and addresses *public-private cyber-security cooperation, industry-academia consortium, knowledge management long-term high risk cyber-security research.*

AN ADAPTIVE E-GOVERNMENT INFRASTRUCTURE SURVIVABILITY FRAMEWORK AND ITS CHALLENGES

The key e-Government survivability challenge is to synthesize a unified adaptive survivability framework (ASF) by integrating the best of breed, synergistic techniques in the fields of vulnerability analysis, intrusion detection, containment and response, and fault tolerance. In particular, these techniques for network, system or application layers need to collaboratively work to generate a survivability framework that provides the following capabilities:

1. Efficient diagnosis of disruption, pinpointing the cause and determining or predicting its impact on the system. The diagnosis should provide support for choosing the best possible mechanisms for timely prevention, control, and recovery from a single, or multiple concurrent disruptions.
2. Isolation of the effect of an impending or ongoing disruption, facilitating quick recovery and high availability of system functionality while the system is being disrupted.
3. Utilization of prior knowledge in adapting to faster, more effective and economical methods of disruption tolerance. Adaptation also occurs in response to unpredictable environmental conditions.

In this section, we present such an ASF for a generic system. A nation-wide extension of its key functionality is crucial to create an overall coordinated e-Government survivability capability. The key component of the ASF is the Adaptive, Disruption

Detection, Response and Recovery (AD^2R^2) module, (see Figure 2) which is responsible for correlating the events across the three architectural layers to efficiently detect the disruptions in the system and respond to them.

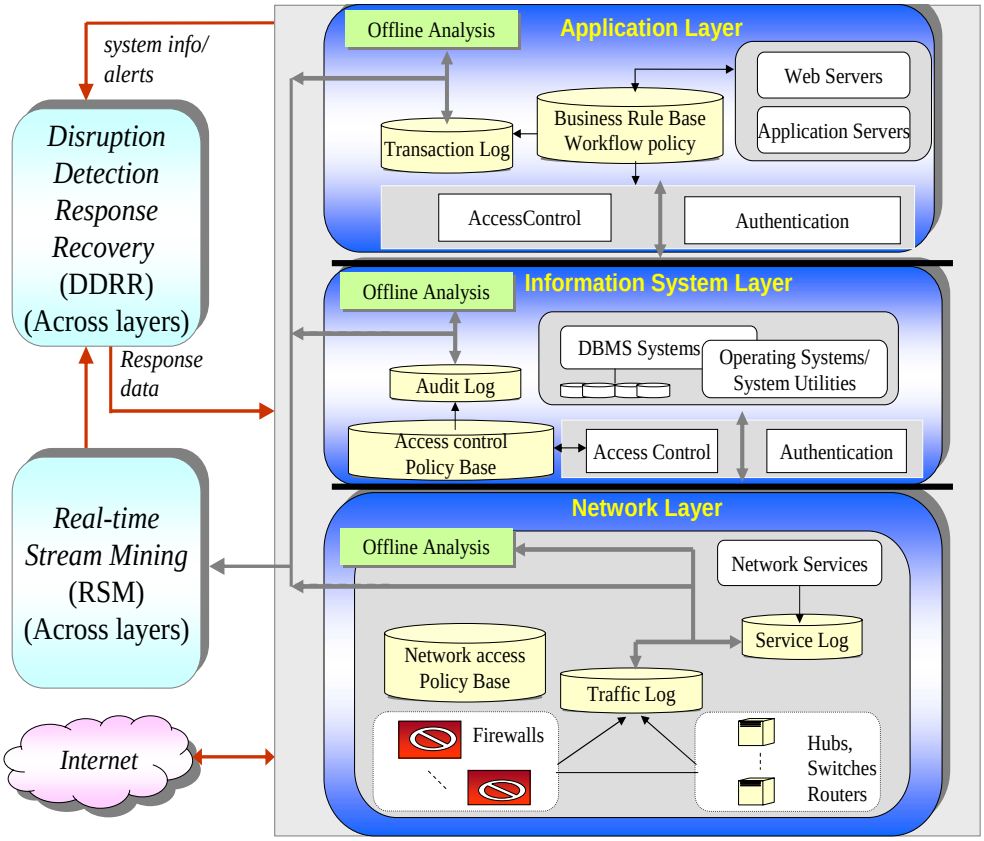

Figure 1. The Adaptive Survivability Framework

The three layer architecture captures a general separation of components within an application environment and include information resources (databases, files), utilities (SSH, Web service, FTP), computational resources (CPUs), and communication links (Network card, routers, switches etc) organized at different layers. These components interact with each other to perform various activities, represented as transactions or events. Figure 2 depicts the functional architecture of the AD^2R^2. The thick arrows indicate the information exchange between each module, whereas the thin arrows indicate inter-module information exchange. The Disruption Detection Module (DDM) analyzes system data to identify ongoing/impending disruptions with the assistance of the Disruption Classifier Module (DCM). A set of predictive parameters are used by the Disruption Diagnosis Module (DDiaM) to construct the containment boundary around components that have been affected by disruption. The Disruption Recovery Module (DRM) is responsible for response and recovery actions. The Coverage Computation Module (CCM) computes coverage for each of the four phases, and thus facilitates the determination of the efficiency and efficacy of the individual modules as well as the entire system.

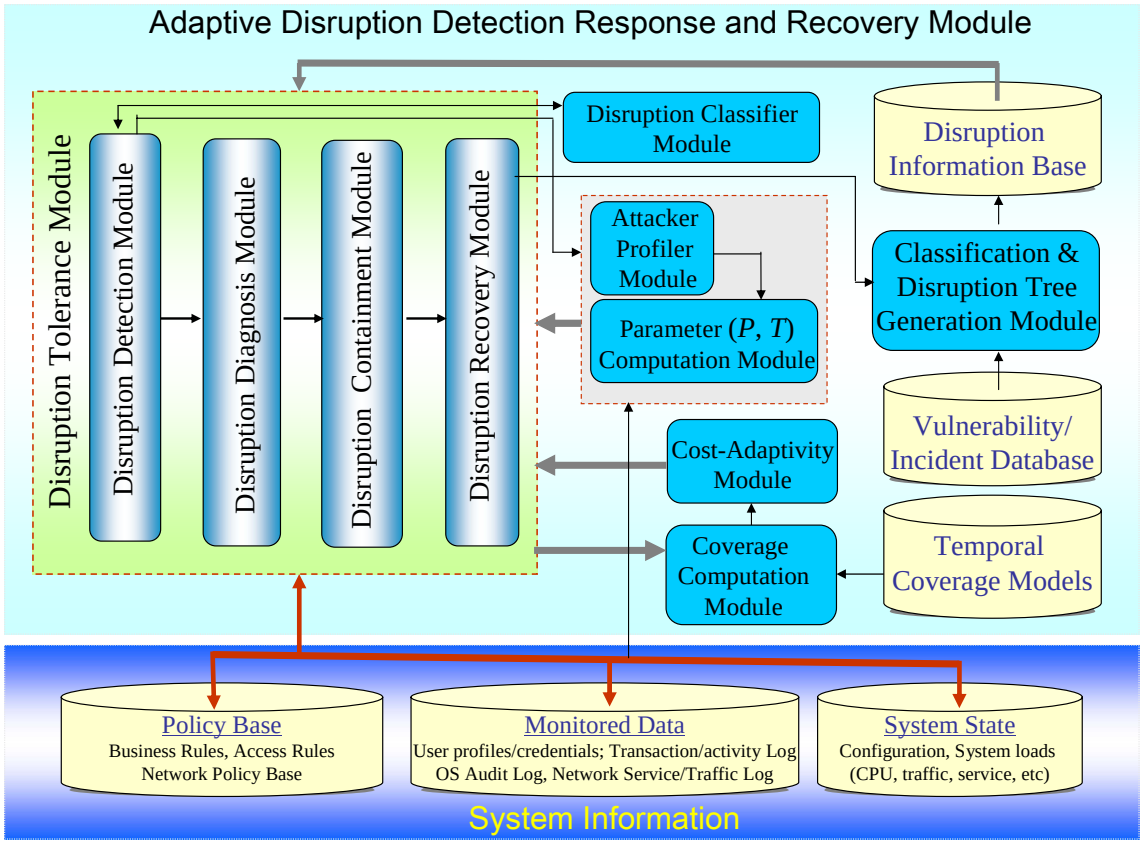

Figure 2. The Adaptive Disruption Detection, Response and Recovery Module

Disruption Information Base

A key challenge for the ASF entails the identification of taxonomic features of general classes of disruptions and representing them efficiently, in a disruption information base (DIB). Various publicly available databases related to security vulnerabilities and incidents, such as those maintained by the US-CERT, CERIAS-VD (Meunier & Spafford, 2002; Song et al., 2000), and BUGTRAQ, are available that contain volumes of vulnerability information that needs to be properly used to generate a knowledge base of disruptions. These databases provide important information related to known software vulnerabilities, such as the impact of the exploitation of a flaw, the types of objects that are directly or indirectly affected by the exploitation, and the fixes that are available. Existing characterization of attacks and intrusions are ad-hoc, unstructured and restricted. For example, a denial of service attack does help us to understand the nature of the attack, however, what impacts it has on different parts of the target system is vague. Furthermore, a unified classification of faults and intrusions is needed.

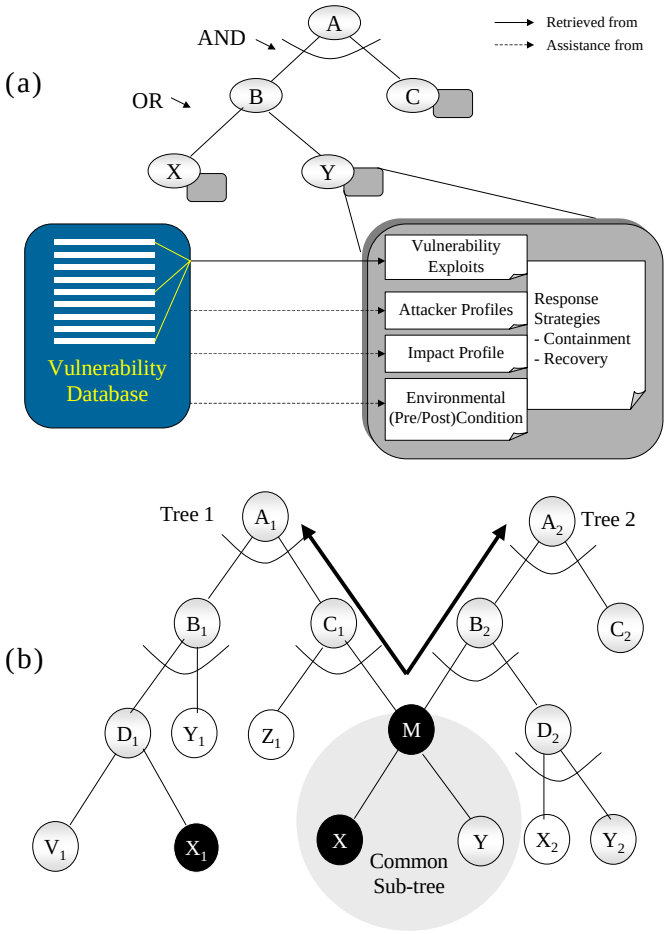

Figure 3. A disruption tree (a) with associated information (b) with a common sub-tree

A promising approach is the representation of the disruption classes using AND-OR trees (Schneier, 2000). The root of a disruption tree represents an ultimate attack goal or a major system failure and its children represent different disruptions that must collectively (*AND*-decomposition) or alternatively (*OR*-decomposition) occur for the major disruption to occur. For example, Figure 3(a) depicts a simple attack tree for attack type A. The model can be expanded to include two key parameters that are associated with each node in the disruption tree, and are computed by the Parameter Computation Module (PCM):

- $P_s(d|D)$ - *probability* of occurrence of disruption d given the occurrence of disruptions D, and
- $T_m(d|D)$ - propagation time of d given the occurrence disruptions D.

Computation of P_s and T_m is based on monitored real-time data. The trees need also be augmented with the following information: (1) system vulnerabilities, (2) attacker profile, (3) system state, (4) impact profile, and (5) response strategies. In particular, a crucial issue is an efficient modeling of attacks and attackers, which is the function of the Attacker Profiler Module, to capture attack objectives and strategies. Determining efficient and cost-effective countermeasure against a disruption will depend on the P_s and T_m values associated with the disruption node and the cost the response strategies. In

scenarios where two disruption trees have a common node, (e.g., node M in Figure 3(b)), choosing the response strategy will involve more complicated decision making. For a robust ASF, it is crucial that the DIB contain detailed information about response strategies for each disruption class.

Most existing classification schemes focus primarily on classifying security vulnerabilities. Attacker modeling has been addressed in (Ellison et al., 1997; Moore et al., 2001; Avizienis et al., 2004), but no formal model has been proposed. Some model checking approaches have been used to generate attack/fault trees (Moore et al., 2001; Schneier, 2000). However, these models rely on exhaustive knowledge of system states and have serious complexity problems.

Adaptive Disruption Detection, Response and Recovery

Development of a real-time collaborative disruption detection framework that analyzes monitored data and the disruption trees is another significant challenge for an ASF. Multiple detectors should be employed at different architectural levels and their results correlated to detect disruptions accurately. Traditional intrusion detection techniques lack real-time capabilities essential for timely response in emerging systems because of the need to analyze a huge amount of log data. In such a case disruption management requires real-time continuous monitoring and intelligent analysis of event data streams to detect impending/ongoing disruptions.

Real-time Mining of Transactional Audit Trail for Disruption Detection: Stream mining, a newly emerging data analysis and mining paradigm, encompasses mining potentially infinite amount of continuously arriving data at a variable rate. Mining event data streams poses many new challenges as such streams differ substantially from traditional data. While traditional databases assume that exact mining results can always be obtained as data elements are synchronized, streams data is often lost, stale, or intentionally omitted for processing reasons and so mining results must be computed with incomplete information. Furthermore, traditional databases assume that applications require no real-time services whereas event data streams are real-time by nature. Therefore, the challenging task of mining such streams is to develop online incremental data mining techniques that operate on incomplete data, or alternatively, data summaries, yet that can obtain “good'' results. Research in mining data can significantly enhance real-time detection.

Disruption Diagnosis and Containment: The goal of the DDiaM is to identify and enforce a *disruption containment boundary* (DCB) so as to prevent the propagation of disruptions. Once the DDiaM identifies that one or more of the sub-goals in the disruption sub tree have been achieved, it queries the DIB to determine the propagation time for a higher-level goal and the post-condition of the particular phase of the disruption to construct the DCB. DCBs may be composed of services, a node, a layer within a node, a component/sub-component within a layer. The diagnosis will also obtain the latency of the detection mechanism that triggered the diagnosis phase.

Response and Recovery: This phase recovers the system from the effect of the disruption and/or initiates counter-measures, and prevents future disruptions. Response may involve downgrading the trust level of the system. Vulnerability prevention in an

automated manner is difficult. The recovery infrastructure also provides feedback to the disruption tree. If the suggested recovery from the current DIB is not successful, alternate recovery strategies are attempted and the node in the disruption tree is re-labeled with the successful recovery strategy. A key challenge is handling multiple simultaneous disruptions as it requires analyzing the dependencies among different disruption classes.

Cost-based Adaptability and Coverage Computation: Another key challenge is the development of various cost-based recovery strategies that can be used to recover the system's normal or good state after a system disruption has been detected. These techniques will aim to achieve an acceptable level of survivability in presence of ongoing disruptions. It is crucial to develop cost-based models for adaptively guiding the response to system disruptions, based on risk analysis. The CC Module computes coverage for each of the four phases facilitating the determination of the efficiency and effectiveness of the different modules. The CC module needs to be augmented by cost-based models to adaptively guide the response capability of the system. The key challenge is ensuring that acceptable values of key metrics such as survivability are achieved at the desired cost.

Several IDSs exist in the literature, such as in (Habib, et al., 2002; Kerschbaum, et al., 2000; Kerschbaum, et al., 2001; Kumar & Spafford, 1994; Reynolds et al., 2002). Some work on isolation of sub-systems targeted by attacks through cutting the connections to them has been shown in (Reynolds et al., 2002). There have been several systems that provide adaptive fault tolerance in distributed systems through a middleware layer. Chameleon (Kalbarczyk et al., 1999), and AQuA (Ren et al., 2003) are two recent examples that focus on tolerating different classes of faults. Some existing adaptive intrusion detection systems include (Ragsdale et al., 2000; Hinton et al., 1999).

C O N C L U S I O N S

In this paper, we have presented the e-Government infrastructure survivability challenges and motivated that the success of an e-Government system is dependent on how resilient it is to the continuous onslaught of intrusions and faults, as any disruption can have severe impact on national security and effective governance. It is crucial that an e-Government system has a coordinated, adaptive capability to analyze, diagnose and timely respond to impending/ongoing infrastructural disruptions. We have presented an adaptive survivability framework for a generic system that integrates the synergy among different technologies to synthesize coordinated and efficient defense capability. Such a framework needs to be implemented to create a dependable, secure and survivable e-Government infrastructure.

R E F E R E N C E S

Alexander, Y. and Swetnam, M. S. (1999). *Cyber Terrorism and Information Warfare I, Assessment of Challenges*, Oceana Publisher Inc./Dobbs Ferry, New York.

Anderson P. L., Geckil, I. (2003). Economic Impact of the 2003 Blackout. *www.andersoneconomicgroup.com.*

Avizienis, A., Laprie, J.-C., Randell and B., Landwehr, C. (2004). Basic Concepts and Taxonomy of Dependable and Secure Computing, *IEEE Transactions on Dependable and Secure Computing*.

Bagchi, S, Srinivasan, B, Whisnant, K, Kalbarczyk, Z. and Iyer, R. K. (2000). Hierarchical Error Detection in a Software Implemented Fault Tolerance (SIFT) Environment. *IEEE Transactions on Knowledge and Data Engineering*. 12 (2).

Bagchi, S., Kalbarczyk, Z., Iyer, R. and Levendel, Y. (2003). Design and Evaluation of Preemptive Control Signature (PECOS) Checking for Distributed Applications. *IEEE Transactions on Computers*.

Denning D., (2000). Cyberterrorism. Testimony before the Special Oversight Panel on Terrorism Committee on Armed Services U.S. House of Representatives.

Denning, D., (2001). Activism, Hacktivism, and Cyberterrorism: The Internet as a Tool for Influencing Foreign Policy, *Internet and International Systems: Information Technology and American Foreign Policy Decisionmaking Workshop*.

Ellison, R. J., Fisher, D. A., Linger, R. C., Lipson, H. F., Longstaff, T. and Mead, N. R. (1997). Survivable Network Systems: An Emerging Discipline. *Technical Report CMU/SEI*-97-*TR*-013 Software Engineering Institute.

Government Accountability Office [GAO]. (2005). Information Security. Emerging Cybersecurity Issues Threaten Federal Information Systems. *GAO Report to Congressional Requesters.GAO-05-231*.

Garfinkel, S. and Spafford, E. H. (1997). *Web Security and Commerce*. O'Reilly and Associates, Inc., Sebastapol, CA.

Gibbs, W.W. (1994). Software's Chronic Crisis. *Scientific American*.86.

Habib, A., Hafeeda, M. M. and Bhargava, B. (2002). Detecting Service Violations and DoS Attacks," Technical Report *CERIAS TR 2002-15*, Purdue University.

Hinton, H., Cowan, C., Delcambre, L. and Bowers, S. (1999). SAM: Security Adaptation Manager. *Annual Computer Security Applications Conference* (ACSAC), Phoenix, AZ.

Joshi, J. B. D., Aref, W. G., Ghafoor, A., and Spafford, E. H. (2001a). Security models for web-based applications. *Communications of the ACM*. 44(2), 38-72.

Joshi, J. B. D., Ghafoor, A., Aref, W., and Spafford, E. H. (2001b). Digital Government Security Infrastructure Design Challenges. *IEEE Computer*. 34 (2), 66-72.

Kalbarczyk, Z., Iyer, R. K., Bagchi, S. and Whisnant, K. (1999). Chameleon: A Software Infrastructure for Adaptive Fault Tolerance. *IEEE Transactions on Parallel and Distributed Systems*. 10 (6). 560-579.

Kerschbaum, F., Spafford, E. H. and Zamboni, D. (2001). Embedded Sensors and Detectors for Intrusion Detection. *Journal of Computer Security*.

Kerschbaum, F., Spafford, E. H. and Zamboni, D. (2000). Using embedded sensors for detecting network attacks. 1st *ACM Workshop on Intrusion Detection Systems*.

Kumar, S. and Spafford, E. H. (1994). An Application of Pattern Matching in lntrusion Detection. *COAST TR* 94-07, Department of Computer Sciences. Purdue University.

Landwehr, C. E. ,Bull, A. R., McDermott, J. P. and Choi, W. S. (1994). A Taxonomy of Computer Program Security Flaws, with Examples. *ACM Computing Surveys*, 26 (3). 211-254.

Meunier, C. and Spafford, E. H. (2002). Running the free vulnerability notification system Cassandra. *CERIAS TR* 2002-34, Department of Computer Sciences, Purdue University.

Moore, P., Ellison, R. J. and Linger, R. C. (2001). Attack Modeling for Information Security and Survivability. *CMU/SEI*-2001-*TN*-001, Software Engineering Institute, Carnegie Mellon University.

Ragsdale, D. J., Carver, C. A., Humphries, J. W. and Pooch, U. W. (2000). "Adaptation Techniques for Intrusion Detection and Intrusion Response Systems. *IEEE International Conference on Systems, Man, and Cybernetics*. 2344-2349.

Ren, Y. (J.), Bakken, D. E., Courtney, T., Cukier, M., Karr, D. A., Rubel, P., Sabnis, C., Sanders, W. H., Schantz, R. E. and Seri, M. (2003). AQuA: An Adaptive Architecture that Provides Dependable Distributed Objects. *IEEE Transactions on Computers*. 52 (1). 31-50.

Reynolds, J., Just, J., Lawson, E., Clough, L, Maglich, R. and Levitt, K. (2002). The Design and Implementation of an Intrusion Tolerant System. *International Conference on Dependable Systems*

Schneier, B. (2000). *Attack Trees. Secrets and Lies*. John Wiley and Sons, New York. 318-333.

Song, G., Mandujano, S. and Meunier, P. (2000). CERIAS Classic Vulnerability Database User Manual. *CERIAS Tech Report* 2000-17, Purdue University.

Terms and Definitions

Disruption. A system failure or intrusion that prevents the normal operation of a system.

Critical Infrastructure. An infrastructure containing systems, assets and services which a country's economy and society depend on.

Survivability. The ability of a system to maintain its functionality even when disruptions occur.

Disruption Signature. A sequence of events by which a disruption can be recognized.

Intrusion Detection Systems: Systems that can identify ongoing/impending system intrusions by recognizing intrusion signatures or anomalous system activities.

Disruption Containment. Controlling the propagation of the effects of a disruption to other parts of the system.

Stream Mining: An emerging data analysis and mining paradigm that can be applied to potentially infinite amounts of continuously arriving variable rate, real-time, asynchronous event data streams.