INTRODUCTION

What is interesting is that all of these attacks occurred despite the wide deployment of security technologies: 95 percent have firewalls, 61 percent an IDS, 90 percent access control of some sort, 42 percent digital IDs, etc. Clearly the technologies are not working.

— Bruce Schneier, CEO Counterpane Systems, referring to the published results of the 2001 CSI/FBI Computer Crime and Security Survey

If the components of security assurance are broadly defined in terms of capabilities (IT hardware and software products) and operations (systems management and process), this approach endeavors to quantify the operations component. That is, by quantifying the probability of secure operating practices, the risk evaluation of “cyber damage” will be more accurate.

An oft-quoted mantra of security professionals holds that “security is a process, not a product.” Poor systems management and operational practices are an enabling factor in many compromises. Yet, efforts to measure systems security and quanti-
fy the risk of cyber damage have focused on the information security (IS) technologies. If the crux of the problem with security technologies is misuse/mismanagement as a variable disconnected from the product design (e.g., the problem with firewalls is not how they are built, but how they are operated), there will continue to be a dearth of meaningful metrics and tools for quantifying actual operating practices. Without quantitative data to support qualitative determinations of reasonable operating practices, the effectiveness of systems management procedures will continue to be unpredictable, and attempts to significantly improve security will be largely ineffective.

This article advances the novel concept of (1) employing forensic tools and techniques to examine computer systems and correlate the data artifacts to find patterns of systems operations; and (2) incorporating these forensic techniques into a methodology that can be applied to a diversity of computer systems comprising the private, commercial, and civil infrastructures. Measuring the implementation of best practices according to evidence left by the operators’ activities will quantify the reality of human computer management, which can be applied to render more meaningful criteria for product certification. This, in turn, will provide an incentive to better align IT capabilities with operational reality.

Attempts to narrow the disconnect between IT capabilities and operations are perhaps misguided. For example, the Center for Internet Security has developed a set of Security Benchmarks that rate a site on a scale of 1 to 10. The scientific underpinnings for such ratings are scant, and security leaders admit that the vulnerabilities that enable security exploits are generally unknown and unknowable. The state-of-the-art in verifying compliance with these procedures is currently self-reported surveys (CSI/FBI, SANS, IDC, etc.) against “best practices” and the occasional vulnerability scanning and network penetration testing. These methods provide only anecdotal, static, or circuitous indicators of actual practices and compliance. Basing industry operations that proceed on perceived risk and response — insurance actuarials and premiums; business-enterprise decision making; and legal liability claims — on the accuracy of such security assurances threatens to institutionalize poor practices.

The approach to cyber risk evaluation outlined herein aims to converge IT capabilities and operations by applying forensic techniques to digitally excavate systems management. As human operators configure and use computers, their activities leave shadows and artifacts in the systems’ files, logs, and networks. Some of these artifacts are plainly visible (e.g., Web browser cookies or log entries), but some are subtle (e.g., the last time of access of a set of header files). Techniques and tools such as The Coroners Toolkit have demonstrated that it is possible to effectively reconstruct and summarize the activities of intruders as they reconnoiter, invade, subvert, and otherwise attack systems. The same concepts and practical methods can be applied to study the activ-
DEFINING THE PROBLEM: WHAT TYPICAL AUDITS DO NOT REVEAL
Are Human Managers Operating Systems Securely?

It is difficult to ascertain whether human operators are complying with documented procedures because the tools to audit compliance and the effectiveness of compliance are not available. A typical audit of an organization’s security practices compares its stated policy and procedures against a set of best practices. The comparison is highly subjective and fundamentally suspect because implementing more procedures from a best practices menu does not necessarily correlate with better system management.

More damning is the fact that the comparison ignores the actual systems, which are fundamentally the subject of the security practices. While some more knowledgeable auditors may randomly sample several systems for signs of “poor” system management such as lack of access controls or signs of previous intrusions, systems are not generally examined for signs of safe or unsafe operational procedures. However, the systems themselves (i.e., machines, disks, databases) are the best, most reliable source of information about the effectiveness of operational practices. For example, if an operator misconfigured a control or failed to execute a program, the evidence of that is discernible within the system, and forensic examination can help reconstruct the misstep. By comparison, current methodologies necessarily leave open the question of effectiveness due to their reliance on self-reporting and lack of transparent and repeatable system autopsy methods.

The current paucity of hands-on examinations of operational practices has implications for system management quality. Firewall deployment in the 1990s offers a historical analogy. Previously, when network managers were evaluated on how well their networks were connected, there was an incentive to remove barriers between networks. Later, it became an accepted practice to defend critical systems with firewalls, but it was a challenge for network managers to know (1) if all the systems were behind the firewall and (2) whether or not firewall administrators were maintaining effective shields during a period of intense network debugging. The advent of robust network vulnerability scanning and penetration testing tools such as nmap quantified and exposed the risks of unadulterated connectivity and permitted managers to institute corrections. These assessment methodologies increased the incentive to comply with as-needed connectivity requirements, to install and properly manage firewalls, and to limit routing broadcasts — thus improving the security of networks.

It is proposed that similar incentives to enhance security operations can be devised by developing forensic tools in conjunction with an assessment methodology. By forensically gauging how system managers...
operate security products and mechanisms, one can measure the “foreseeable misuse” of products. This metric is useful in assessing security risk (for insurance and business decision making) as well as shaping product design and defining relevant levels of reasonable conduct (for legal liability judgments).

ARE OTHER APPROACHES ADDRESSING THE CAUSE — OR JUST THE SYMPTOMS?
The methodology proposed here is markedly different from vulnerability scanning (e.g., SATAN, COPS, ISS). While vulnerability scanning tools can measure the exposure of the system at an instant in time, they do not examine the processes by which hosts are secured or by which previously secure hosts become exposed to attack. In other words, such scans convey the what (that a potentially vulnerable service is exposed) but not the why (how the vulnerability came into being). Did system operators attempt to patch the vulnerable software on several systems but fail due to the peculiarity of a single host? Was a patch applied but then invalidated by reinstallation of a component after a failure? Did an overextended administrator disable one safety mechanism in order to apply another? Knowledge of results without understanding underlying causes encourages a reactive posture and discourages safe and secure practices.

Anecdotally, 80 percent of system compromises are the result of missing patches, “vanilla” out-of-box installs, or system access-control misconfigurations. It may be true that some system platforms expose more vulnerabilities than others, but their real vulnerability may arise because of a high turnover rate for system administrators or because frequent system re-architecting makes it a challenge for system administrators to gain adequate experience. Vulnerability scans and penetration tests yield no data useful for making such determinations.

This approach addresses these problems by building tools and associated methodologies that make use of system-contained evidence to recognize, recover, reconstruct, and ultimately quantify how human operators are configuring and maintaining their systems. The resulting data and metrics can be applied to industries that rely on the accuracy of information and security assurance:

1. Insurance actuarial and premium determinations that rely on accurate assurances of risk factors
2. Legal principles governing the assignment of duty and liability for negligent security (foreseeable misuse, third-party liability claims, shareholder suits)
3. Business-enterprise decision making regarding due diligence inquiries, risk assessment, and cost expenditure (ROI) for information security products and personnel
OBJECTIVE: OPERATIONAL METRICS
Measuring System Management and Operations Effectiveness

With each critical system event (e.g., an employee separation or a discovered intrusion), there exists an opportunity to study how operators react and address the incident. Forensic techniques can be employed to collect the digital residues of activity embedded in the system and reconstruct steps that human operators take in managing their systems (see Exhibit 1). Using these electronic “fingerprints,” one can compare the system administrator’s behavior to a super-set of best practices and characterize their effectiveness.

Our tools extract patterns from system artifacts, system backup images, tripwire databases, log files, etc. into a repository (implemented in a database). Queries related to specific management practices can then generate statistical profiles of best practices. Weighted results of queries yield a score of their system management effectiveness. Driven by a model of best practices, one can analyze actual, historical evidence of safe and unsafe management practices and rate operator effectiveness.

By correlating with other data from studies of system vulnerabilities, such as time-till-break-in [SDSC01], one can quantitatively determine foreseeable misuse and even begin to construct actuarial tables describing probability of damage.
A Concrete Example
In July 2001, shortly before the writing of this article, the “CodeRed” worm infected more than 250,000 Windows NT 4 and 2000 Web server systems in just nine hours (CERT/CC CA-2001-23). The worm exploited an unchecked buffer in the IIS index server extension to display a hacked Web page, infect other systems, and target a denial-of-service attack on the White House Web site.

A month before CodeRed infections became rampant, Microsoft had acknowledged the vulnerability, offered a patch, and strongly urged immediate action by all Web server system administrators. That the administrators of over a quarter-million Web servers had not received, or had failed to effectively act on, bulletins from a variety of sources speaks volumes about common operational practices on these systems.

By the time this article reaches print, checks for the IIS Indexing Service vulnerability will have been incorporated in vulnerability scanners, penetration test rubrics, and system audits. These checks will reveal whether or not systems are still vulnerable — but they will tell very little about the system administrative practices that may have left the systems vulnerable to complete compromise for a month before the exploit was widely publicized.

Common forensic techniques reveal a great deal more. In addition to checking that the affected library file has the same last-modified date/time as the patched file, one can examine the hidden (in the sense that it is not revealed by commonly used system utilities) Master File Table (MFT) record modification date/time to determine when it was installed on the system. The system log can be examined to determine if the system was restarted after installation of the update file (the update would not take effect until restart). And, if an update utility, such as those distributed by Microsoft, was used to apply the patches, even more information will be available because these hotfix utilities leave traces spread through the system registry and log files.

In a multiple-server environment, these forensic traces can be assembled into a collection that shows a pattern of application to administered hosts. One’s understanding of systems operational practices becomes richer when one compares the patterns of multiple updates. One can determine not only the mean windows of vulnerability, but also find whether or not vital administrative practices are occurring in planned, repeatable patterns.

IMPACT OF OPERATIONS METRICS
Insurance Applications of Operations Metrics
Insurance is a tool for risk management in information security. Computer security insurance, like all forms of insurance, represents a systemic approach to mitigating loss by subsidizing costs associated with securing systems and mitigating expected losses.
As previously stated, current measurements related to the cost of security failures and the potential risk to private, commercial, and civil infrastructures are inaccurate. Not only are damages and risk factors difficult to measure and often exaggerated, but anecdotal data from surveys yields numbers as indecipherable as pronouncements by a Delphic oracle. Realistic actuarial data will increase the accuracy of probability of loss estimates by countering “adverse selection,” which occurs when insured parties hide information that would be pertinent to the insurer in calculating risk of loss.

Furthermore, metrics regarding operator management can help close the window of exposure. The emergence of new vulnerabilities, new attack tools, and new legitimate services present windows of exposure (measured by time) within which systems are at risk for intrusion. Prominent factors that define the window of exposure include breadth of knowledge about the vulnerability, speed with which vendors post patches for vulnerabilities, and installation of patches by system administrators. If the goal of system security is to narrow the window of exposure, efforts should be directed at enhancing all of these factors. By quantifying the response time of system operators to implement security measures (e.g., patches), one can more accurately infer the probability of loss variable that is vital to risk valuation.

LEGAL APPLICATIONS OF OPERATIONS METRICS

The law is a vehicle to shape systems security by defining duty and ascribing liability for damages wrought by security failures. In this way, negligence law offers an incentive to implement security or to risk paying tort judgments for breaching a duty to act reasonably in securing systems. Furthermore, negligence law in valuable to defining duty and assigning responsibility between the various human factors that affect computer security, such as manufacturers, businesses, and service providers.

As applied to legal standards that ascribe liability for negligence or product liability, misconfigurations (e.g., insecure permission setting, failure to install patches or upgrades) that are deemed “foreseeable” may alter the assignment of responsibility for damages resulting from security breaches. By analogy, this would be tantamount to Chrysler shipping a car with configurable airbag options accompanied by accurate and explicit directions to enable/disable the device. If, for whatever reasons, operators consistently sustained injury due to misconfigurations, courts would find it easier to adjudge Chrysler liable for airbag design defects. The common thread between this scenario and the computer security situation is that faulty permission settings could be deemed a foreseeable misuse, thereby resulting in foreseeable harm to operators and users.

Quantitative data supporting “foreseeable misuse,” as a rule rather than an exception, in security operations can lead to more accurately determining which party can best bear the costs associated with risk of insecure
systems. Ultimately, this information is useful in resolving the “blame game” discrepancies between vendors and end users, wherein the former credit system exploits on misconfigurations by the end operator, and the latter argue that vulnerability-riddled software is the culprit. In the absence of data upon which to draw assessments of reasonable conduct, assignment of liability and security assurances will be ad hoc, at best.

Finally, the increasing awareness and publicity paid to computer security incidents (i.e., DDoS in February 1999, Love Bug virus in 2000, CodeRed worm in July 2001, CD Universe credit card extortion case) is catalyzing efforts to seek recompense for damage. Despite the fact that a digital miscreant may be unidentifiable or insolvent, third-party or shareholder liability suits offer a means to seek damages owing to security vulnerabilities. Given that the identifiable and solvent parties will likely be companies that facilitate the cyber damage, metrics and methodology are valuable in resolving disputes more efficiently and consistently.

**BUSINESS-ENTERPRISE APPLICATIONS OF OPERATIONS METRICS**

These operations metrics can be applied to business-enterprise decision making regarding due diligence, expenditure for information security products and personnel, and accurate assessment of remaining risk.

Reliable metrics will help corporate decision makers grasp how operators currently may be tasked with responsibilities without the necessary resources. System security at present is viewed as a drain on the bottom line, rather than an enabler. A quantifiable, repeatable, and transparent methodology for measuring the risk factors associated with system mismanagement enables accurate cost-benefit analysis for system management expenditures.

**SUMMARY**

The tools and methodologies discussed in this article forensically analyze the traces left by system management, compare the observed patterns with “best practices” templates, and develop statistical profiles of system management practices.

By collecting profiles on an aggregate level, correlating these patterns with incident data, and drawing comparisons with the data for a particular enterprise, one can dramatically improve risk assessment capabilities. As with other actuarial information, the profiles for an individual enterprise will not predict particular outcomes, or prove what did in fact happen in a particular case. Instead, the authors hope it will provide a reliable statistical association with outcomes. Comparison of individual profiles with best practices models will also offer opportunity for in-depth analysis of practices.

Data gathered via these new methodologies will allow more accurate risk identification. Comparison of resulting data by industry, platform,
operations techniques, and experience will provide the vital assessments necessary for the maturation of risk management decisions. Comparisons of particular cases with corresponding industry data will enable quantifiable measures of information assurance and computer security.

The authors’ approach addresses an often-overlooked component of business security risk — the actual practices of system managers. It will answer basic and pressing questions about how humans can operate large, complex, heterogeneous systems effectively and responsibly.

Notes
1. Within this article, forensic techniques are defined as detailed, repeatable examinations of the system state using tools that capture data from the original sources, such as files, disk, memory, or log files, rather than relying on application reports.

Christopher Wee is the founder and president of Anzuru Technologies, Inc., a research and information security consulting company. Dr. Wee has created a successful and profitable business in security design and forensic services, and recognizes the urgent need for additional research in computer forensics. Prior to Anzuru, he was a co-founder of Infidel, Inc. (with Rebecca Bace and Terri Gilbert), specializing in advice and due-diligence services for venture capitalists. As a postdoctoral researcher at the University of California, Davis, Computer Security Lab, he conducted research in intrusion detection systems (DIDS, GrIDS, LAFS), policy-directed host auditing (AWB), and network security. He has relevant system management experience as senior information security analyst for Intel Online Services, a managed-Internet-hosting provider. Dr. Wee holds a Ph.D. and other degrees from the University of California, Davis.

Erin Kenneally is a forensic analyst with the Pacific Institute for Computer Security (PICS), San Diego Supercomputer Center. Ms. Kenneally is a licensed attorney who holds Juris Doctorate and Master of Forensic Sciences degrees. She conducts substantive research, publication, and consultation relating to prevailing issues in computer forensics law and computer security in both the criminal and civil arenas. This includes evidentiary, procedural, and policy implications of new technologies and digital evidence. She has lectured and helped coordinate training conferences for state and local judges, law enforcement, and industry professionals concerned with high-technology crime.

In addition, she liaises with a broad range of constituents and professional organizations, including the Southern California Hi-Tech Task Force and RCFQ (Regional Computer Forensics Laboratory). She holds leadership positions within the advisory board and board of directors for the Computer and Technology Computer High Tech Task Force (CATCH) and the High Technology Computer Investigation Association (HTCIA).

Steve McMahon is a security researcher for Anzuru Technologies, where he specializes in Web server and Windows operating system security and administration issues. He has nearly 20 years of experience in applications and systems programming.