A Data Centric Security Cycle Model for Data Loss Prevention of Custodial Data and Company Intellectual Property

Mathew Nicho  
College of Information Technology  
University of Dubai  
mnicho@ud.ac.ae

Avinash Advani  
Boole Server  
Milano, Italy  
a.advani@booleserver.com

Abstract: Review of data breach trends in the last five years reveal that data at rest, in use, and in motion, inside and over the extended network, is being increasingly affected. While organizations primarily focus on protecting sensitive customer financial information, the protection of custodial data and company secrets has been a back burner issue. Moreover, errors, mistakes and accidents on the part of the employees working/travelling/residing onsite and off-site with company media/data, have worsened the situation such that current technical and socio-technical controls are not adequate in preventing theft of media or the accidental or intentional misuse/loss of portable data. To overcome this issue, the security action cycle model of Straub and Welke (based on the general deterrence theory) is used as a theoretical lens to build a data centric security cycle model to safeguard the data that are “at rest, in motion and in use”. Finally, the paper discusses how the model can be further empirically validated using the updated IS success model of DeLone and McLean.

Keywords-IS Security; data breaches; data centric security.

I. INTRODUCTION

Security and privacy has remained one of the top ten key issues for Information Systems [IS] executives since 2003 [1] and crucial to the continuous wellbeing of modern organizations [2] with the result that organizations need to protect information assets against cyber crime, denial-of-service attacks, web hackers, data breaches, identity and credit card theft, and fraud [3]. A firm’s information related assets are now among its most valuable assets [4]. Thus, the protection of this asset through the process of information security, is of equal importance [5]. The application of existing technical IS security frameworks and IS controls have been effective in preventing attacks from external entities into the organizational networks, but the mobility of the organizational staff and information technology (IT) assets along the extended network have posed serious risks to organizational data.

Inside organizations, valuable and critical data may exist in laptops, portable storage media, storage media at distant locations, and in emails in online and offline mode. In this instance, sensitive data is not only out of the organizational security defenses but that the existing organizational IS controls and assurance cannot be monitored and thus becomes voluntary on the part of the employees to protect it from theft, accidental loss and misuse. While information security controls and models exist for securing the organizational network, data loss prevention is an area of scant research.

The object of this paper is data, which primarily refers to custodial data (protection required by regulation/law) and company secrets (high value intellectual property and assets). The objective of the paper is to propose a model to protect and control this data. Protection, focus on the data in its raw form as it rests in the file system, flows through the network and while it is being used. The controlling aspect considers how the data or information is used once authorized users have gained access to it and even revokes the access to it. The output of this paper is a data centric security cycle model to protect the data (hereinafter refer to data that are “at rest, in motion and in use”) in real time. The theoretical basis of the model rest on the security action cycle model of Straub and Welke [6], which in turn is derived from the general deterrence theory.

The paper is divided into five sections. The current state of IS security breaches is analyzed from statistics to highlight the relevance of data loss prevention (Section 2). This is followed by analyzing the high profile data breaches occurred during the last five years from 2007 to 2011 to find out the gaps in the socio-technical security structure (Section 3). Then, the current literature on IS security is evaluated to ascertain the existence and appropriateness of relevant IS security models for data protection and to select an appropriate theoretical basis for the model (Section 4). Finally, the selected theory is employed to build up the proposed Data Centric Security Cycle (DCSC) model, as well as provide the rationale for using the updated IS success model of DeLone and McLean [7] for evaluating the success of the model (Section 5).
II. DATA BREACH ANALYSIS

The central objective of any security system is the ability to prevent undesired access, while still allowing authorized access to information [8] but with cyber incidents growing in intensity and severity [9] the risks related to information security have become a major challenge and a top management priority for many organizations [10]. Thus, despite the critical role and relevance of information and information security in an organization, unauthorized breaches into organizational internal and the extended networks occur with greater frequency and severity [6] [9] [11] [12]. IS security have remained within the top ten key issues in information systems since the advent of the Internet in the early 1990s [13], and maintained this position in subsequent studies [1] [14] [15] [16] [17]. This section analyses data breaches from a statistical as well as from a methodological perspective to ascertain the severity and cause of breaches.

A. Data Breaches – a Statistical Perspective

As organisations rely heavily on information to conduct their daily activities [5], any disruption or intrusion poses grave threat to the organisation and its extended enterprise. Recent statistics taken from different sources reveal the seriousness of the issue. In the annual 2010/2011 CSI computer crime and security survey (285 respondents), 41.1% of those surveyed reported a security incident in their company [18]. Similarly, a 2011 sector wise study by Deloitte on 138 organizations revealed that 75% experienced an IS security breach, which was an increase over the previous year by 62% [19]. Also the average total cost of a data breach rose to $6.75 million in 2009, with major increase in number of records lost per incident [20] where stolen laptops were cited as the number one cause of a data breach in 2009.

From a numerical perspective, in 2011, the US Identity Theft Resource Centre (ITRC) reported a total of 419 breaches resulting in 22,918,441 records being compromised [21]. Compared with ITRC, Datalossdb [22] reported 890 breaches in 2011 with insiders (accident and malicious) accounting for 39% of the breaches. Among the three types of breaches namely reported abuse, discovered abuse and undiscovered abuse [23], the above figures mainly represent only the reported abuse all of which involve substantial loss to the stakeholders. The breaches identified by the various sources are not uniform as they use newspaper articles, copies of letters reporting a breach to consumers, notification lists of state agencies, direct entry of incidents by the public, and other web sites as sources for the breaches [24].

ITRC statistics that are based on publicly available breaches in United States provides categorized statistical data on breaches (see Figures 1 and 2). Considering accidental exposure and data on the move, statistics from 2007 till 2011 does not indicate any drastic reduction in either the number of breaches or records breached.

![Figure 1. Total number of reported/known breaches in US from 2002 to 2011 (Source: ITRC, 2007 - 2011)](image1)

While payment card information and authentication credential are still the most sought after data, the largest breaches (in terms of the number of records) reveal that the personal information along with company secrets are a potential target (Figure 3). The above statistical figures reveal the increase in intensity and frequency of threat to all financial and organizational data from within and external to the organization.

![Figure 2. Total number of records breached from 2007 to 2011 from reported/known sources in US I (Source: ITRC, 2007, 2008, 2009, 2010, 2011)](image2)
Information can be categorized into custodial data and secrets [26]. It has been noticed from the above statistics that hackers have shifted from financial information to the wider custodial and company secrets (see TABLE I). Irrespective of the type of date, the cost incurred by the organisation comes at a price.

B. Cost of Data Breach

While IS security breaches cost dearly for the organizations, the question is not whether organizations need more security, but to look at cost-benefit methods to evaluate IT security so as to ‘optimize’ security countermeasure investments and reduce spending without sacrificing protection (Arora et al., 2004). The financial loss suffered by US companies average $ 5.1 million for a single data breach and the cost incurred for one compromised record comes to $214 [27]. Taking the 2011 ITRC reported breaches into consideration, this would amount to $ 2136 million (at the rate of US $ 5.1 million x 419 breaches). Calculating the loss from a records compromised, the loss amounts to $ 4904 million (22.9 million records based on ITRC x $ 214). Verizon (2012) reported 855 incidents with 174 million compromised records in 2011 from the small global sample from the 33 participating countries. If this statistic is taken to calculate the loss, it comes to an annual loss of $ 37,236,000,000 from this small sample.

C. Human Factor and Mobility in Data Breaches

The human being remains the weakest link in the control and security of systems and networks [28] and “frequently security violations involve those who are authorized or have access to the sensitive data of concern” [29] (p. 26). Moreover, it is estimated that at least half of the breaches to IS security are unauthorized system access made by internal personnel [30]. Thus, the involvement of humans in information security is equally important and many examples exist where human activity can be linked to security issues [2]. As noted by Schultz [31], information security is primarily a people problem and technology is designed and managed by people, leaving opportunities for human error. It has been observed that organisational security target the prevention of external threats, such as hackers and viruses, leaving organizations open to breaches from the inside [30]. Hence, the occurrence of IS security breaches by internal personnel may be reduced if greater emphasis were placed on internal threats to IS security that can occur when employees handle information in their day-to-day jobs (ibid).

The environment of today’s worker is evolving from centralization and control to mobility and performance [32]. With the rapid adoption of mobile office, modern organizations are exploiting mobile media and their infrastructure in a more strategic manner, thus developing work styles and office designs that are evolving around new technologies like mobile phone, laptop and email address [33]. While these technological advances provide advantages to organizations [34], all of these pose significant threat to information. This mobility highlights the relevance of protection to the data, focusing on the data itself and evaluate how sensitive data and information can be securely delivered and shared/transferred beyond the organisational network. This calls for a data focused security model that can incorporate the policies and controls into the technical security architecture to protect data from unauthorised access, wherever it resides (fixed infrastructure, end points and mobile devices).

III. CAUSES OF DATA BREACH

The year 2011 witnessed a spate of high profile cyber-attacks, and the top ten reported cases (number of records breached) in the US and a few organizations in Europe is taken to analyze the gaps in IS security. Since the ITRC and the Privacy Rights Clearinghouse (PRC) [35] documents publicly available breaches in US, these sources were combined and analyzed to come up with a list of top ten data breaches in 2011.

![Figure 3. Varieties of data compromised by number of breaches and records [25]](image-url)
Table I reveals theft/accidental loss of media as a common cause of breach rather than hacking into the company network. In the above ten cases, the data stolen mostly contained personal and company information, rather than credit cards or financial data which can be sold to third parties or used for further social engineering attacks. Secondly, in all of these top ten breaches, only cases 1, 2 and 5 attacks came from external parties into the organizational network for which the IS security manager have control of. In these three cases, a simple analysis of the IS security defenses and improved IS security can prevent further attacks. Regarding the rest seven cases (3, 4, 6, 7, 8, and 9), data loss/theft occurred with ease, and since the data was in media out of the organizational perimeter defense it would have been impossible to track with the normal internal controls. These cases prove that the IS security manager need to have more control over the data that are at rest, in motion and in use. Thus, IS security should not only be built like a staircase of combined measures in order for information security measures to become effective [36] but mutually dependent on each other [37] Berghel, 2005 cited in [36] Hagen, Albrechtsen and Howden.

The cases analyzed in this section highlights the relevance of, control and changing access to information in real time even after it has been sent out to the extent of even restricting functions on secured documents, e.g. print, copy, save as, print screen. If the data is outside the authorized perimeter then the manager should have the right to revoke access instantly, even after delivery has been taken, no matter where it has ended up (except in the case of back up tapes). This can be equated to digital rights management.

**TABLE I. EVALUATION OF THE TOP TEN BREACHES (Source: ITRC, 2012; PRC, 2012)**

<table>
<thead>
<tr>
<th>Organisation</th>
<th>Records</th>
<th>Nature of breach</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sony Playstation</td>
<td>70,000,000</td>
<td>Stolen information include - name, address, country, email address, birthdate, PlayStation Network/Qriocity password and login.</td>
<td>External attack - the hacker used spear phishing rather than highly sophisticated hacking to break into the network.</td>
</tr>
<tr>
<td>Zappos (owned by Amazon.com)</td>
<td>24,000,000</td>
<td>Customer information - names, email, billing and shipping addresses, phone numbers, the last four digits of credit card numbers, and cryptographically scrambled passwords were stolen</td>
<td>External attack – where the hacker gained entry into the company server (Methodology is not known as the probe is still ongoing).</td>
</tr>
<tr>
<td>TRICARE-SAIC</td>
<td>1,900,000</td>
<td>Backup tapes containing SAIC SAIC data stolen from the car of a Tricare employee.</td>
<td>Non-technical – employee error; procedure not followed</td>
</tr>
<tr>
<td>Texas Comptroller</td>
<td>1,500,000</td>
<td>Unencrypted information transferred and kept in a server accessible to the public</td>
<td>Correct procedure not followed when transferring information across servers</td>
</tr>
<tr>
<td>Betfair</td>
<td>2,300,000</td>
<td>SQL injection attack on a code vulnerability</td>
<td>Technical breach, procedure not followed especially network segregation and file integrity monitoring.</td>
</tr>
<tr>
<td>Health Net IBM</td>
<td>2,000,000</td>
<td>Nine server drives went missing from the data centre of the California office of Health Net</td>
<td>Appropriate policies and procedures have not been followed by those responsible for both the physical and logical protection of critical data.</td>
</tr>
<tr>
<td>Jacobi Medical Centre</td>
<td>1,700,000</td>
<td>The files (cassette tapes in a box) was stolen from a van operated by GRM Information Management Services, when the driver left the van unattended and unlocked.</td>
<td>Correct procedure not followed prior to and when transporting storage media</td>
</tr>
<tr>
<td>Nemours Children Health System</td>
<td>1,600,000</td>
<td>Unencrypted computer backup tapes containing patient billing and employee payroll data stolen from a Nemours facility in Wilmington, Delaware</td>
<td>Correct procedure not followed when storing/storage media. As per the control, the tapes were supposed to be safely locked</td>
</tr>
<tr>
<td>Oregon Department of Motor Vehicles</td>
<td>1,000,000</td>
<td>USB or CD containing personal information lost from the department. Thief caught.</td>
<td>Correct procedure not followed when storing/disposing redundant data</td>
</tr>
<tr>
<td>Eisenhower Medical Centre</td>
<td>514,330</td>
<td>The computer used to check-in patients at the Center in Rancho Mirage was stolen from the open lobby area</td>
<td>Correct procedure not followed. The computer was not protected with drive encryption nor physically locked.</td>
</tr>
</tbody>
</table>
(DRM), which broadly refers to the set of policies, techniques and tools used to manage the use of digital content. While there are solutions in digital rights management, this are limited to a few media and does not have the time bound or rights revoke methods.

IV. INFORMATION SECURITY MODELS

The prime goals of information security is to provide confidentiality, integrity, authentication, non-repudiation (to data) and the key factor in getting value from security is to insure that technology investments protect the right things [38]. Based on this objective, numerous models and frameworks have been proposed for securing information systems in an organisation, apart from a few studies to secure information. But, very few researches have been done for securing data “at rest, in motion and in use”. While it is impossible to totally secure any information systems, much is known today about how "systems risks" can be substantially reduced through effective management practices [6].

Lehman [29] presented a detection model using audit trails in tracking potential security violations. Deterrence theory was used for modeling sanctions by Siponen et al, and Starub [39, 40]. The use of marketing campaigns in security breach prevention has been proposed by McLean [41] along with training and education model by [42]. Straub and Welke, [6] proposed the security planning model that addresses deterrence, prevention, detection and remedies using the general deterrence theory. A five level information security management model by Solms, et al. [43] encompass deterrence, prevention, detection and discipline. Straub’s [40] model for detection and discipline of computer abuse was focused more on evaluating investment in IA security, while Trecz’s [44] layered multi-planes model for information systems security focus on e-business systems security. Ganame et al [45] proposed a distributed SOC (Security Operation Center) which is able to detect attacks occurring simultaneously on several sites in a network, and a six view perspective of system security was presented by Yadav [12]. The model of Siponen and Vance [46] uses neutralization technique and deterrence theory to develop and implement security policies. Analysis of these models reveal that they target organisational and security systems rather than data. In the realm of data centric security, the first and only attempt to provide security at data level resulted in the data centric security model proposed by IBM [47]. The purpose of the IBM model was to directly align business strategy and IT security through the common thread of data. This model focuses only on authentication, authorization and disclosure control, thus only partly complying with the security action cycle (deterrence, prevention and detection, but not remedial actions).

A. The Security Action Cycle Model

An analysis of the security models reveals that either one or more of the four components of the security action cycle (Figure 4) namely, deterrence, prevention, detection and the discipline (remedies) element is used to provide a comprehensive protection to the information systems as a whole, rather than focus on the data itself. Since hackers and unauthorized personnel target data, the four components of the security action cycle is used to built a model to focus and protect the data. The general deterrence theory have been used in IS security models [40, 46]. The security action cycle model [6] which is based on the general deterrence theory outlines four components for preventing ‘systems risk’ (information systems damage or loss) such as deterrence, prevention, detection and remedies. Two classes of countermeasures—deterrents and preventives—have been found to be effective [23]. Deterrents are passive, administrative controls that take no active role in restricting the use of system resources. Examples include computer security awareness training sessions and distributed policy statements that specify conditions for proper use of the system. Preventives, on the other hand, screen access to the system to admit authorized users only and include physical restraints such as locks on computer equipment room doors and programmed restraints such as software locks on accounts, files, transactions, and data items [40]. Detection is the process of monitoring the events in a network. Remedies have been described by Starub and Welke as punishment and recovery procedures. The application of these four elements involves the use technical as well as non-technical information systems controls which is referred to as audit.

B. Role of Controls in IS Security

Internal controls are policies, procedures, practices, and organizational structures put in place to reduce risks [48]. Appropriate controls are necessary to protect organizations from suits against negligent duty and compliance to computer misuse and data protection legislation [49]. While a “control framework is a recognized system of control categories that covers all internal controls expected in an organisation” [IIARF 2002, cited in [50], an internal control provides reasonable assurance regarding the achievement of objectives in the area of effectiveness and efficiency of operations, reliability of financial reporting and compliance with regulations [51].
The analysis derived from the three sections direct the researcher to the need for audit, protect data, transfer data across networks, control the access of data and revoke access even after delivery. These analysis viewed through the security action cycle is given in Figure 4.

Figure 4. Data centric components embedded into the security action cycle

V. DATA CENTRIC SECURITY CYCLE MODEL

Data centric security is a relatively new field and despite the fact that there exist a few security ontologies in the literature, none has yet touched on the topic of data-centric security [52]. Data-centric security starts with a hard look at what data the business must protect and its classification where the focus is on the data instead of the network [53]. Data sharing among the extended stakeholders requires complex procedures in the form of data sharing agreements, contracts, access privileges, usage and routing control infrastructure in the form of security policies [52], but despite these controls, the extent of data breaches, data loss and data pilferages could not be contained.

The data centric security cycle (DCSC) model revolves around five dimensions namely protection of data, securely deliver and share sensitive information, audit every single activity done by users to information, control and change access on a continual basis, and if necessary revoke access even after the delivery of information has been taken by the user. Security of information here refer to where, when, who and how security can be used as an enabler of business, not as a restriction. This is done through storing information in encrypted format wherever needed, protecting through strong encryption, digital rights management and business-focused data leakage prevention. This ensures sharing data solely amongst authorised users, both internal and external to the organization and controlling it in real time of all user, administrator and file activity. The ISO 27001 IS security standard reiterates the cyclical nature and thus incorporates the Plan – Do – Check - Act (PDCA) cycle of Edward Deming.

Also, taking into account the cyclical nature of the security action cycle, the model presented in Figure 4 can be refined further to the data centric security cyclical model (Figure 5). A data centric security solution provides persistent protection (persistent -at rest, in transit and in use); Strong encryption that is dynamic and granular data rights management where the rights can be managed at any time and in real time and it can be assigned in a granular way (administrators, users and data). The concept puts data control back in hands of data owners by separation between administrators and owners/users (Independent management) thus making it flexible for both internal and external users, for any type of file, at any storage location, for any size and type of solution, customization and integration. This data centric security concept is easy to use, simple, easy to administer, implement, maintain and is complete in terms of securing data. Since the control of data use is in the hands of the administrator, the human factor in data breaches due to accident and negligence can be substantially reduced, but not fully contained.

Next, the application of the model is discussed. When converted into a tangible solution, the model consists of three components namely the server, the web client and the agent. The server component manages all operations namely data protection, encryption, all information related to user profiles and their rights to access and use of shared files. The web client users are able to access all the functionality of the server via any Internet browser, for both Windows and Mac. The agent provides added security functions like block screen capture, print screen, video streaming, revoke access any time, create encrypted local disks, synchronize a
local disk with the centralized resources through the web client, encrypt local resources making them accessible even without connection to the server and even provide time bound offline access. Once cases are selected this will be provided to them for implementation and thereafter for empirical evaluation, the researchers plan to use the updated IS success model of DeLone and McLean [54] using the variables namely information quality, systems quality and service quality from a IS security perspective. Once the users use the model, feedback (interview questions) will be framed based on these above variables.

VI CONCLUSION

Regulation and compliance are increasingly important where compliant doesn’t mean secure and secure doesn’t mean efficient. While organisations need to secure networks and financial information, the concept of focussing on the data that are in use, in motion and at rest has gained relevance as is evident from the data breaches.. With this objective, this paper propose a model that focus on the real time protection/control of data. The model that was based on the security action cycle model can audit, protect, control, secure and even revoke rights to data in online and offline mode by the data custodian. Further research can focus on protecting/controlling media and digital backup tapes using RFID technology.

REFERENCES


