Concepts and Tools for Protecting Sensitive Data in the IT Industry: A Review of Trends, Challenges and Mechanisms for Data-Protection

Omar Tayan
Dept. of Computer Engineering, College of Computer Science and Engineering (CCSE), IT Research Center for the Holy Quran and Its Sciences (NOOR), Taibah University, Saudi Arabia

Abstract—Advancements in storage, dissemination and access of multimedia data content on the Internet continues to grow at exponential rates, while individuals, organizations and governments spend huge efforts to exert their fingerprint in this information age through the use of online multimedia resources to propagate thoughts, services, policies, e-commerce and other types of information. Furthermore, information at different levels may be classified into confidential, sensitive and critical data types. Such data has been subject to numerous tools and techniques for providing automated information processing, information management and storage mechanisms. Consequently, numerous security tools and techniques have also emerged for the protection of data at the various organizational levels and according to different requirements. This paper discusses three important types of information security aspects that includes; data-storage, in-transit data and data access-prevention for unauthorized users. In particular, the paper reviews and presents the latest trends and most common challenges in information security with regards to data-breaches and vulnerabilities found in industry today using simple brief summaries for the benefit of IT practitioners and academics. Thereafter, state-of-the-art techniques used to secure information content commonly required in applications-software, in-house operations software or websites are given. Mechanisms for enhancing data-protection under the given set of challenges and vulnerabilities are also discussed. Finally, the importance of using information security policies and standards for protecting organizational data content is discussed along with foreseeable open issues for future work.

Keywords—sensitive-data; data-breaches; data-protection; trend analysis; classification

I. INTRODUCTION

The Digital era has witnessed an ever increasing dependence on the Internet and world-wide web (WWW) in our lives and daily activities. Moreover, the continuing growth of such information and communication technologies has played a crucial role in establishing the Internet and WWW as the dominant IT platform for digital content distribution, communication, and other general information sharing activities. Hence, millions of worldwide users have benefited from the advantages of fast and simple mechanisms for digital information exchange. On the contrary, such benefits are also vulnerable to the problems and threats associated with securing the digital content. The literature of digital multimedia content has identified a number of security issues to be addressed that includes: digital copyright protection, counterfeit prevention and data-authentication. Such requirements are more predominant in the case of specialized and sensitive data. Generally, all digital multimedia content on the Internet can be classified into text, images, audio and video content with the challenge being to provide secure, robust and reliable storage and dissemination for each media type. On the other hand, many electronic-transactions impose an additional requirement to ensure data-confidentiality, particularly for the case of sensitive customer and client information. This paper explains the important and timely role of information assurance and related security techniques concerned with the storage, propagation, reproduction and communication of sensitive online data-content.

Background Concepts in Information Security

Some of the key objectives of digital multimedia security can be classified into; requirements for assuring authenticity and integrity of content, usage-control, binding of identification data with the cover-content, and ensuring secrecy and non-repudiation in the transmitted content [1 - 3]. The state-of-the-art techniques in information security can be used to achieve the necessary security requirements according to the target application and content-type in most cases. The protection of sensitive digital multimedia content can be achieved using authenticity and integrity based techniques to ensure that 100% accurate content is transmitted and stored, whereas secrecy of the data can be achieved using cryptographic approaches prior to transmission. Integrity is concerned with ensuring that the transmitted data is not altered or tampered with, and is exactly similar to the version sent. Integrity can be achieved using numerous techniques such as; encryption, hashing, watermarking etc. Authentication, on the other hand, is associated with establishing trust between communicating parties, such as assurance by verifying that the data-content had originated from a trusted source/publisher. Authentication can be achieved using digital signatures/certificates and digital-watermarking. In contrast, confidentiality and non-repudiation requirements are typically used with e-transactions that are concerned with data-secrecy during the communication and are achieved using encryption schemes.
II. VULNERABILITY TRENDS IN THE IT SECTOR AND A CLASSIFICATION OF CHALLENGES FOR DATA-PROTECTION

Organizational employees with access to networked-devices have a key role in protecting the organization’s information assets since those devices can provide a gateway to information stored elsewhere on the same network and can be exploited as vulnerable access-points for internal or external intruders. In fact, an organization faces a number of emerging penetrations due to security activities and cloud-based applications [4-9].

- Fraudulent websites that can imitate other sites
- Data-theft
- Fake purchases
- Intruder attacks
- Damage to an organization’s reputation

Moreover, this information-era has witnessed many ways in which data and security-breaches have penetrated our normal business operations and daily-life activities. Such security-breaches can now be found in most/all IT systems covering new and known application-domains and functions, including: e-Banking and e-Commerce applications [4], e-Healthcare systems [5], wireless and mobile devices [6,7], cloud-assisted applications, wireless sensor-networks (WSN) and Internet-of-Things (IoT) [8] and Big-Data processing activities [9]. Figure 1 illustrates those recent domains with emerging penetrations due to security-breaches [4-9].

Table 1 classifies the actual extent of damage faced by organizations and individuals due to information security vulnerabilities and cyber-attacks and includes threats/vulnerability statistics industrial-sector:

**TABLE 1. CLASSIFICATION OF CYBER-ATTACKS AND VICTIM COUNTS IN 2015 [10]**

<table>
<thead>
<tr>
<th>Category</th>
<th>Attack/Vulnerability Classification</th>
<th>Quantitative Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal Digital Identities</td>
<td>Personal Identities Lost or Stolen</td>
<td>Over Half a Billion</td>
</tr>
<tr>
<td>Data-Breaches</td>
<td>Total Breaches</td>
<td>318</td>
</tr>
<tr>
<td></td>
<td>Average Num of Identities Exposed/Breach</td>
<td>1.3 Million</td>
</tr>
<tr>
<td></td>
<td>Total Identities Exposed</td>
<td>429 Million</td>
</tr>
<tr>
<td></td>
<td>Top Causes of Data Breaches by Identities</td>
<td>Attackers, Accidental-loss, Theft</td>
</tr>
<tr>
<td></td>
<td>Top Ranked Industry Sectors Targeted for Data-Breaches</td>
<td>Services (Inc., Healthcare), Finance, Public Admin., Wholesale Trade, Retail Trade</td>
</tr>
<tr>
<td>Email Threats, Malware and Bots</td>
<td>Overall Email/Spam Ratio</td>
<td>53% of emails</td>
</tr>
<tr>
<td></td>
<td>Email Phishing Ratio</td>
<td>1 in 1846 emails</td>
</tr>
<tr>
<td></td>
<td>Email Malware Ratio</td>
<td>1 in 230 emails</td>
</tr>
<tr>
<td></td>
<td>Number of Bots</td>
<td>1.1 Million</td>
</tr>
<tr>
<td></td>
<td>Number of New Malware Variants</td>
<td>431 Million</td>
</tr>
<tr>
<td></td>
<td>Top-Ranked Industries Targeted by Spam Emails</td>
<td>Mining, Manufacturing, Construction, Services, Agriculture</td>
</tr>
<tr>
<td>Mobile Devices</td>
<td>New Vulnerabilities Detected</td>
<td>528</td>
</tr>
<tr>
<td></td>
<td>New Android Malware Variants</td>
<td>3944</td>
</tr>
<tr>
<td></td>
<td>Ratio of Apps Analyzed and Classified as Malware</td>
<td>Over 30%</td>
</tr>
<tr>
<td>Vulnerabilities</td>
<td>New Vulnerabilities</td>
<td>5585</td>
</tr>
<tr>
<td></td>
<td>Most Frequently Targeted App for Zero-Day Exploit Vulnerabilities</td>
<td>Adobe Flash Player</td>
</tr>
<tr>
<td>Web-Attacks</td>
<td>Scanned Websites with Vulnerabilities</td>
<td>78% (15% of which were critical)</td>
</tr>
<tr>
<td></td>
<td>Websites Detected with Malware</td>
<td>1 in 3172</td>
</tr>
<tr>
<td></td>
<td>Top-Ranked Most Frequently Exploited Sites</td>
<td>Technological sites, Business sites, Searching sites</td>
</tr>
</tbody>
</table>

Other interesting facts related to attack-frequencies, types and detection-rates with a classification of countries and regions are now summarized in Table 2.

---

![Fig. 1. Emerging Information Security Breaches by Domain](image-url)
TABLE II. STATISTICAL SUMMARY OF VARIOUS VULNERABILITIES, THREATS AND ATTACKS

<table>
<thead>
<tr>
<th>Category</th>
<th>Reference / Citation</th>
<th>Summary of Statistical Trends</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infected Smartphone Apps</td>
<td></td>
<td>150M apps scanned, with 9M malwares, 9M suspicious apps and 3M affected devices detected</td>
</tr>
<tr>
<td>Top 10 Countries Ranked by Infections for Smartphone Devices in 4th-Quarter, 2015</td>
<td>Mobile Threats Report [11]</td>
<td>India (86k infections), USA (82k infections), Brazil (68K infections)</td>
</tr>
<tr>
<td>Largest Rate of Increase in Mobile Malware Threats</td>
<td></td>
<td>From Third to Fourth Quarter, 2015 (72% increase)</td>
</tr>
<tr>
<td>Total Worldwide Malwares Detected</td>
<td>McAfee Labs Threats Report [12]</td>
<td>Total of nearly 500M, with the largest growth in the fourth-quarter in 2015 due to newly emerging mobile threats (representing 12M in the fourth-quarter of 2015). Highest regional malware infection rates were reported Africa, followed by Asia and South America.</td>
</tr>
<tr>
<td>Total Worldwide Web- Threats Detected</td>
<td></td>
<td>New suspect and phishing URLs reaching 17M and 1.4M, respectively by the end of 2015, with global spam emails reaching over 4.5 trillion messages.</td>
</tr>
<tr>
<td>Top Network Attacks Detected</td>
<td>Browsers (36%), Brute-Force attacks (19%), DoS attacks (16%), SSL attacks (11%),</td>
<td></td>
</tr>
</tbody>
</table>

In [13-18], cyber-attacks/e-crimes were classified into three categories. The first category relates to external-attackers, which involves attacks using viruses, worms, Trojan-horses, and Denial-of-Service (DoS). Next, internal-crimes were classified as those that include unauthorized access, theft of IP-rights/theft of knowledge by employees and breach of conduct by business partners. Finally, the social-engineering category of attacks had included; phishing and spoofing. Moreover, the work in [14] provides a report on the common technical vulnerabilities in web-applications and websites, which can be summarized into the following aspects:

- Cross-site Scripting (XSS)
- SQL Injection
- Malicious File-Inclusion
- Insecure Direct Object Reference
- Information Leakage/Improper Error Handling
- Insecure Cryptographic Storage
- Insecure Communications
- Broken Authentication/Session Management
- Failure to Restrict URL Access

Numerous examples exist relating to the extent of such security breaches within the various IT-based industrial-sectors, and particularly in the case of many highly-reputable and financially-strong organizations as shown in Table 3.

TABLE III. WORLDWIDE IMPACT OF SECURITY BREACHES ON VARIOUS IT-BASED INDUSTRIES [15]

<table>
<thead>
<tr>
<th>Organization/Sector</th>
<th>Victim of Security Breach</th>
</tr>
</thead>
</table>

Notably, a number of data security breaches can also be recalled that relate to some recent and famous incidents with impact on most online users today. Some of those recent events include:

- LinkedIn Accounts – 6.5 million accounts were hacked on 5th June’12 and passwords publicly posted on 6th June’2012.
- ARAMCO attack – 15th August 2012 – virus Shamoon attacks 30,000 PCs at company, taking Aramco two-weeks to recover.
- Facebook – most popular social networking site had around 600,000 “compromised” accounts/day.

Table 4 classifies the top fifteen countries involved in the generation of those attacks that had resulted with consequent data security breaches according to another study [13].
III. CLASSIFICATIONS OF TECHNICAL AND ORGANIZATIONAL-LEVEL TECHNIQUES FOR PREVENTING DATA-BREACHES AND ENHANCING DATA-PROTECTION

The discussion presented in this section comprises of technical approaches, organizational approaches and strategies for managing information-security requirements, as follows:

A. Technical Approaches

Some of the main technical requirements concerned with the protection of sensitive content are summarized in Table 5.

B. 3-Tier Organizational Approach

A summary of the procedures and guidelines that forms part of an organizational action-plan for protecting digital information can be further classified into three levels (management level, implementation level and systems level) as follows:

Management Level Protection (General advice):
- Assign a Chief Security Officer (CSO).
- Develop an organizational security-policy
- Seek third-party accreditation that ensures high-security standards are achieved, e.g. ISO 27001, ISO 9001 for improving quality-standards and overall reputation.
- Perform regular risk assessments and revise management solutions currently in-place.

Implementation Level Protection (summarized from [16]):
- Educate employees of the organization’s security policies.
- Raise awareness of the network-administrator/IT helpdesk role and contact details.
- Be mindful of how to share sensitive data across the network.
- Do not open unexpected email attachments or downloads.
- Perform regular backups, password-updates, encryption, biometric control.

Table 6 highlights common state-of-the-art security-techniques that have emerged as a consequence, together with their goals the corresponding application-domains.
• Caution should be taken not to email content that you would not want to be distributed to unauthorized parties.
• Ensure data-sharing features on the PC are off or set to allow access to authorized persons only.
• Keep the system and security updates active and patched on PCs.
• Do not store sensitive data in an unsecure location online.
• Remote access to an organization’s PCs should be done via secure methods (e.g. SSH/VPN).

Systems Level Protection (summarized from [17]):
• Select a secure e-commerce hosting platform.
• Use a secure connection for online transactions that is PCI compliant (e.g. SSL certificates).
• Do not store sensitive customer details (e.g. card numbers).
• Use address and card verification systems.
• Request customers to use strong passwords.
• Setup alert systems for suspicious activity (e.g. same IP/person may be using many card numbers).
• Use Layered Security (e.g. Perimeter, Network, Host, Application and Data layers) such as firewalls, contact-forms, and login boxes.
• Provide Security training to employees.
• Use tracking-numbers for all e-transactions or orders.
• Monitor your site regularly (e.g. use RT-analytics tools to view interaction) and ensure that the hosting platform continuously monitors their own servers (e.g. against malware, viruses, updates needed).
• Perform regular PCI scans (e.g. using Trustwave, PrestaShop).
• Patch/Update systems and third-party code (including perl, java, php, joomla, wordpress).
• Use DDoS protection service and mitigation service (e.g. Cloud DDoS protection and DNS service).
• Consider a fraud-management service from a card-company.
• Ensure the platform host regularly backs up the site and has a disaster-recovery plan.
• Encrypt stored, transmitted and processed data.

Table 7 identifies a number of quick-tip solutions for several very common web-based attacks at the system-level and implementation-level.

TABLE VII. QUICK TIPS PROVIDING SOLUTIONS TO MOST COMMON ATTACKS [18]

<table>
<thead>
<tr>
<th>Website Attack (Type)</th>
<th>Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>SQL Injection (DB Access)</td>
<td>- In PHP use: <code>mysql_real_escape_string</code> function for any variable in SQL queries.</td>
</tr>
<tr>
<td></td>
<td>- Set DB access permissions</td>
</tr>
<tr>
<td>Secure private/personal data (e.g. Transactions)</td>
<td>Use encryption, e.g. SSL when passing data between website &amp; webservers/databases.</td>
</tr>
<tr>
<td>Computer Security</td>
<td>Anti-spyware, anti-virus, scanners, firewall, software updates.</td>
</tr>
</tbody>
</table>

C. Strategies for Managing Information Security

When an organization evaluates the need and extent for information security techniques against the deployment costs, a number of considerations must be made as part of a complete strategy that includes [19]:

• A chief-security officer (CSO) must balance the trade-off between risks and costs for securing the organization’s assets.
• The security-management approach should consider:
  – Determining the information assets and their value
  – Determining the maximum time which the organization can function without a given asset.
  – Implementing security-procedures to protect each asset.
  – **Loss Calculations** should be used to justify costs for purchasing security techniques: *Annual Expected Loss* = *Single Loss Expectation* × *Annual Occurrence Rate* [19].
• Security Cost-Benefit Analysis: develop a quantitative analysis to calculate the potential business benefit and costs involved with addressing security risks.
• Net Benefit Calculation provides an efficient tradeoff measure: *Return Benefit* = *Annual Expected Loss* − *Annual Cost of Action* [19].
• A business continuity plan (BCP) is needed for each organization.
• Determine the category for tolerable downtimes for the organization’s services: e.g. <12 hours, 24 hours, 72 hours, 7 days, 30 days.
• Develop an Information-Security Policy: a policy document is needed that describes what is and what is not permissible use of information in the organization and the consequences for violating the policy.
• The Policy document includes: access-control, external-access, user and physical policies.
The Policy should be developed by a policy-committee with members from user-groups and stakeholders.

The policy-committee should meet regularly and should be updated with the organization’s needs and current laws.

Good training and communication of a new policy is needed for awareness.

Further reading with best practices using summarized guidelines for organizations can be found in [10].

IV. ESSENCE OF INFORMATION SECURITY STANDARDS AND INFORMATION SECURITY POLICIES

The necessity for developing and conforming to IT and information security standards at the business or institutional level cannot be understated or emphasized enough since it provides a multi-layer protective shield to many of the security deficiencies and consequent vulnerabilities described in this paper. One example of a set of standards considered as highly relevant to the domain of IT and information processing is that of the WWW Consortium (W3C) Web standards, which are developed with the aim of attaining two key agendas, namely: (i) design principles that includes; Web-for-All (human communication, commerce and knowledge-sharing available to all people, hardware types, software, network infrastructures, native languages, geographic locations, and physical/mental abilities) and Web-on-Everything (all types of web-access devices) , and (ii) a Vision for W3C standards that includes; Web-for-Rich-Interaction, Web-of-Data-and-Services, and a Web-of-Trust [20]. Effectively, such ‘web-standards’ have established technologies for creating and interpreting web-based content designed to benefit users while remaining compatible with future Web-developments [21].

Other standards particularly relevant to the information-security domain include the ISO 27001 and ISO 27002 standards which establish protocols and guidelines for different levels of security policies within an organization. ISO 27001 formally specifies an Information Security Management System (ISMS) that includes a suite of activities for the management of information security risks and covers all sizes and types of organizations (commercial enterprises, government agencies and non-profit) and industries/markets (retail, healthcare, defense, banking, government and education) [22]. Additionally, the ISO 27001 can be used as the basis for formal compliance assessment by accredited certification bodies in order to certify an organization.

Similarly, the ISO 27002 standard is also relevant to all types of organizations that handles and depends on information processing. This standard explicitly refers to the security of all forms of information, and is not only limited to IT-systems security (e.g. cyber-security). However, whilst the ISO 27001 specifies a mandatory for implementing an ISMS, the ISO 27002 standard specifies suitable controls within the ISMS and is presented as a Code-of-Practice complementary to the ISO 27001 standard [23]. Furthermore, organizations cannot obtain certification by an accredited body through adherence to the ISO 27002. Hence, ISO 27002 is a standard which is normally used more flexibly in accordance to an organization’s context [23]. In short, every organization should develop its own information-security policy based on a standard (e.g. such as ISO 27001 with/without ISO 27002). An example document-structure for an organizational policy is provided in [24]. Once a policy-document has been developed, some training for IT staff and employees is required to ensure all are clear of what is required at all levels of responsibility.

V. CONCLUSIONS AND OPEN RESEARCH ISSUES

The rapid growth of the Internet and the World Wide Web (WWW) suggests that more attention is required for the security and protection of online sensitive data at various levels. There is an essence and need for Information Assurance in the digital community that encompasses the protection of information in the public and private sectors, academia, or other purposes. Those various sectors are required to take the necessary technical and administrative measures to protect its information assets. In this paper, a number of remarkable data-breach cases and their trends and statistics in the IT sector were shown, along with the technical and organizational-techniques for mitigating such attacks.

Emerging challenges and open research issues which persist in the domain of information security includes: mobile-security, scripting-languages and web-security, and cloud-based security. A notable trend for the development of a more complete information security approach was observed in the literature and related products in the marketplace, which includes: the encryption of something you have or wear (e.g. personal smart-phones) and the encryption of what you are (e.g. using biometric-data). Finally, Figure 2 summarizes and classifies the future research directions and open-issues as a result of our analysis and research findings from a number of recent works.

REFERENCES


http://www.gogulf.com/blog/cyber-crime/

http://www.infosec.gov.hk/english/business/other_sywa_1.html

http://www.informationisbeautiful.net/visualizations/worlds-biggest-data-breaches-hacks/

Fig. 2. Summary and Classification of Future Directions and Open Issues in Information Security

- New authentication schemes (using context or biometrics based one-touch protocols)
- Cloud-Based Authentication (for higher scalability and less complexity)
- Migration from user-authentication to transaction-authentication schemes

- Network segregation
- Network traffic-analysis
- Regular vulnerability tests
- Continuous updating/patching
- Regular malicious-activity detecting tools
- Proxy solutions

- Joint optimization of security, reliability and throughput
- Cross-layer wireless security design and analysis
- Physical-layer security investigations & field-tests for emerging 5G systems

- Verifying of Trust and Privacy
- Verifying data-integration
- Modeling and Verification of Big-Data in Healthcare
- Cloud-secure provenance with tracking and monitoring actions

- Personalized Privacy Preservation
- Data Customization
- Provenance for Data-Mining

- How to create a good information-security policy