SECURITY ACROSS THE CURRICULUM: IMPLEMENTATION IN A DATA ANALYTICS PROGRAM

Mark Ciampa, Western Kentucky University, mark.ciampa@wku.edu
Evelyn H. Thrasher, Western Kentucky University, evelyn.thrasher@wku.edu

ABSTRACT

Because the end user is widely recognized as the weakest link in computer security, educating users on cybersecurity is essential for fortifying systems and networks against attacks. Security training and instruction are important not only to meet the current demands of securing systems, but also to prepare students for employment in their respective fields. As employees, these college graduates will have access to critical enterprise data in order to perform their daily tasks; without proper instruction they could be unknowingly putting the data at risk. Yet for business data analytics graduates the level of cybersecurity training falls higher than that for the average user yet lower than that of the security technology specialist. To address the need for the proper level of security training for these soon-to-be data scientists, this study describes an initiative to achieve both goals in a unique approach, that of distributing the proper level of security training and instruction across all courses instead of in a single, high-level security course. This study details the efforts of an undergraduate Business Data Analytics program to distribute appropriate-level cybersecurity instruction and training across the required curriculum. The implementation steps are explained along with a discussion of the initial results.

Keywords: Security, Data Analytics, Higher Education, Information Technology (IT)

INTRODUCTION

Computer security is universally recognized as essential in today's technology environment, with billions of dollars and man-hours spent annually in attempts to ward off attacks. However, tangible results from these defensive actions are rare. One of the factors that is generally recognized as a significant weakness in cybersecurity defense is that of the end user. Despite the best technology appliances purchased and installed to prevent attacks, a user who unknowingly--or uncaringly--opens a malicious attachment or clicks on a link in a phishing email can instantly negate the very best cybersecurity defenses. Security training and instruction is considered essential against the growing tidal wave of attacks.

This situation is particularly important for college students currently enrolled in business data analytics who are training to be the new cadre of tomorrow's data scientists. These individuals handle critical data on a continual basis, even more so than the normal users in the enterprise. Students need security training to be secure end users--but they also need instruction on selected technical security topics to be able to effectively communicate with the enterprise security technical staff regarding the unique requirements of data analytics. Thus, the upcoming data scientist needs a solid understanding of both levels--basic end user and (limited) advanced security technician--to be able to practice good security and also interact with higher level security professionals.

Very few colleges address this need. A small number of schools are beginning to explore offering practical security training for all students, such as password management and proper email hygiene. At the other end of the spectrum most schools teach advanced security courses to those students who are seeking employment as security professionals. But what about business data analytics students who need both basic user training and limited advanced security instruction so as to effectively interact with higher level security professionals in the enterprise?

This study describes an initiative to achieve both goals in a unique approach, that of distributing the proper level of security training and instruction across all courses instead of a single, high-level security course. This study details the efforts of an undergraduate Business Data Analytics program to distribute appropriate-level cybersecurity instruction and training across the entire required curriculum.
BACKGROUND

The ubiquity of cybersecurity attacks has escalated to the point that most users are numbed by the never-ending reports of the latest attacks and required defenses. In just the first six months of 2015 it is estimated that 234,919,393 data records were lost or stolen as the result of data breaches, which, if spread over the entire period, would equate to 56,611 records stolen every hour (Gemalto, 2015). In February 2015 unknown attackers electronically broke into the Bangladesh Central Bank and tried to transfer almost $1 billion. Despite the fact that their fraudulent transactions were cancelled after a typographic error raised concerns about one of the transactions they nevertheless managed to transfer $81 million. A subsequent investigation revealed that the Bangladesh Central Bank had no firewall installed and was using a second-hand $10 network device (Bright, 2016). An aging and neglected network infrastructure continues to leave organizations increasingly vulnerable to compromise. Cisco analyzed 115,000 Cisco devices on the Internet and found that 92 percent of the devices in the sample were running software with known vulnerabilities. In addition, 31 percent of the Cisco devices were “end of sale” and 8 percent were “end of life” devices (Cisco, n.d.). And stolen payment card numbers are so numerous that they sell in attacker underground sites for as little as $5 (McFarland, Paget, & Samani, 2016).

Most attacks are not directed at uncovering unknown technical vulnerabilities, according to the National Security Agency’s director of Tailored Access Operations (Zetter, 2016). Instead, attackers are focusing on using social engineering, or relying on the weaknesses of individuals, as a means of gathering information or crafting an attack. One of the most common forms of social engineering is phishing. Phishing is sending an email or displaying a web announcement that falsely claims to be from a legitimate enterprise in an attempt to trick the user into surrendering private information. Users are asked to respond to an email or are directed to a website where they are requested to update personal information, such as passwords, credit card numbers, Social Security numbers, bank account numbers, or other information. However, the email or website is actually an imposter and is set up to steal what information the user enters (Ciampa, 2017).

The number of phishing attacks continues to escalate. In the last three months of 2015 over 14 million new samples of phishing malware were observed (Group, 2016). The overwhelming number of phishing attacks (77 percent) are directed at targets located in the United States, while other nations trailed significantly, such as China (5 percent), France (3 percent), Great Britain (3 percent), and Australia (2 percent). Yet over the past three years the percentage of phishing attacks targeting US companies has only grown 9 percent. However, this is not the result of increased defenses to ward off these phishing attacks. Rather, it is likely because U.S. companies have been so saturated by the phishing market that there is little room for additional growth (Phishlabs, 2016). The average 10,000-employee company spends $3.7 million annually dealing with phishing attacks (Korolov, 2015).

Because phishing attacks are based on social engineering, it is difficult to counteract phishing through a purely technical solution. One strategy that has been used to minimize successful phishing attacks is to warn users about the threat. Different warning technologies have been proposed. One of the first warnings created was a website authentication indicator that used a padlock icon on the toolbar. This was displayed on the web browser itself and signaled the presence or absence of a secure sockets layer (SSL) connection between the browser and the web site (Cranor, 2006). A closed padlock indicated an SSL connection, which usually implies that communications with the site are encrypted and that the site has authenticated itself. However, SSL does not ensure that the website is necessarily trustworthy. This is because certificate authorities issue domain-validated certificates to anyone who can demonstrate domain ownership by only receiving e-mails addressed to that domain name (Jackson, Simon, Tan, & Barth, 2007). It does not make any implications regarding the validity of the web site. In response to these weaknesses, the certificate authority industry developed extended validation SSL (EV SSL) certificates. EV SSL, in addition to displaying a padlock, also turns the web browser’s address bar green and displays the name of the extended validation certificate owner. This shows that the transaction is encrypted and the organization has been authenticated according to higher standards.

Another example of warning users of phishing is having the web browser feed information back to proactively warn users. A yellow button labeled ‘Suspicious Website’ in the web browser address bar indicates that the user may be viewing a suspected phishing site while a red status bar indicates the user is visiting a known phishing site (Tulloch,
Northrup, & Honeycutt, 2007). The address of the website the user is trying to visit is compared to a blacklist of known phishing sites maintained by the browser’s vendor. However, phishers check their websites against blacklists and then modify their techniques until they are no longer designated as a phishing site. And because a phishing web site is active online for only a few days, in most instances the phishing web site will have already ceased to function by the time it is placed on the blacklist. Wu, Miller and Garfinkel (2006) list five additional vendor-security toolbar indicators, while Dhamija and Tygar (2005) along with others have created their own web browser add-ons that can warn users.

However, these different phishing warning technologies have consistently shown that they are marginally effective. Usability studies testing the effectiveness of the SSL padlock by Dhamija, Tygar, and Hearst (2006), Downs et al. (2006), Schechter, Dhamija, Ozment, and Fischer (2007), Whalen and Inkpen (2005), Wu et al. (2006), and others have demonstrated that this standard security indicator is limited in its effect. In addition, usability studies by Jackson et al. (2007) illustrated the ineffectiveness of EV SSL. Dhamija et al. (2006) concludes that “standard security indicators are not effective” (p. 581) while Cranor (2006) states that “a growing body of literature has found the effectiveness of many of these indicators to be rather disappointing” (p. 45).

A more effective strategy to protect users from phishing is training users to resist phishing attacks. User education and training is the key to counteract phishing, according to Dhamija, Tygar and Hearst (2006), Downs, Holbrook, and Cranor (2006), Jackson, Simon, Tan and Barth (2007), and Kumaraguru et al. (2007).

The effectiveness of training users to resist phishing corresponds with the general advocacy that information security training and awareness are two of the most effective offsets to mitigate the human risk posed to information security (Parsons, McCormac, Butavicius, Pattinson, & Jerram, 2014). Training is emphasized by Long (1999), Mangus (2002), Tobin and Ware (2005), Werner (2005), Witson (2003), and Yang (2001) among others. This is because it is frequently maintained that the user is the weakest link in computer security. Observations ranged from a mild statement of “certain user practices contribute to information systems vulnerabilities” (Mangus, 2002, p. 5) to a sharp rebuke of “the average home user is clueless about security and should be required to obtain a license to log on to the internet” (Werner, 2005, p. 96). Training should not only apply to end users but even to federal government agencies (Macmanus, 2013) and the U.S. Department of Defense workforce of military, civilians, and contractors (McDaniel, 2013).

Valentine (2005) noted that the National Strategy to Secure Cyberspace (NSSC) document, created by the U. S. President’s National Infrastructure Advisory Council, calls for a comprehensive national awareness program to empower all Americans, including the general population, ‘to secure their own parts of cyberspace’. Specifically, the Department of Homeland Security, through the NSSC, calls upon home users to help the nation secure cyberspace ‘by securing their own connections to it’.

Educating general users on security also has additional benefits. First, it can provide future users with the critical thinking and basic skills to collaborate with vendors and IT professionals who provide security tools (Werner, 2005). A second benefit is that it may serve to deter attackers. Long (1999) stated that countermeasure strategies to reduce systems risk fall into four distinct and sequential activities: deterrence, prevention, detection, and recovery. General deterrence theory has been used in the study of criminals and other anti-social personalities and maintains that individuals with an instrumental intent to commit anti-social acts can be dissuaded by the administration of strong disincentives and sanctions relevant to these acts. General deterrence theory has also been applied successfully to IT by Straub (1990) and Straub, Carlson and Jones (1993). Educating users can be a form of deterrence by providing information about the risks of security and penal actions that can be taken against attackers.

Another benefit is that security training can change employee attitudes (Berry & Houston, 1993). In time, this becomes the way things are done and inculcates a positive information security culture (Da Veiga & Eloff, 2010). Effective training and awareness both result in behavioral change in organizations and are critical in embedding information security principles at the employee level (Da Veiga, 2015).
Although Long (1999) advocated that security instruction should begin as early as kindergarten, most researchers state that institutions of higher education should be responsible for providing security awareness instruction, including Crowley (2003), Mangus (2002), Null (2004), Tobin and Ware (2005), Valentine (2005), Werner (2005), and Yang (2001). This instruction and training is important not only to meet the current demands of securing systems but also to prepare students for employment in their respective fields. Werner (2005) said that as employees, new college graduates will have access to critical data to perform their jobs, yet they could be the weakest link in a secure computer system primarily because of inadequate education, negligence, and inexperience. Long maintained that the need for organizations to develop appropriate policies requires all decision makers to have a certain level of awareness of standards for security.

Support for making institutions of higher education the primary source for security awareness training comes from several different sources. The Action and Recommendation 3-4 of the NSSC calls upon colleges and universities to model user awareness programs and materials (Valentine, 2005). Frincke and Bishop (2004) summarized several of the major groups and efforts currently involved in computer security education with institutions of higher education. These include the Colloquium for Information Systems Security Education (CISSE), the International Federation of Information Processing Working Group 11.8 on Information Security Education (IFIP WISE), and the Workshop on Education in Computer Security (WECS). The National Security Agency (NSA) also had developed an effort aimed at creating a larger core of computer security trained professionals known as the National Centers of Academic Excellence in Information Assurance Education, which even provides large numbers of college scholarships under its ‘Cyber Corps’ program.

Different techniques have been used by institutions of higher education to provide security awareness instruction. Several schools hold annual cyber security training fairs. The purpose of these fairs are primarily to promote a secure culture within the campus, provide information-security education and training to all constituents, provide hands on peer-to-peer mentoring about security, teach users how to protect data through the deployment of common security practices, and to evaluate the cybersecurity awareness levels of the student population (Larson, 2015). Gaming has also been promoted as a technique to teach security awareness (Hendrix, Al-Sherbaz, & Bloom, 2016). Huang (2015) notes that an issue with existing cyber security training is that it relies mostly on lecture-style instructions without much hand-on experience. He advocates a training solution that provides a realistic, human-in-the-loop environment for exploration, collaboration, and interaction in order to promote effective learning and calls the approach Cyber Situation Awareness (CSA).

The location of security awareness instruction and training in a college curriculum should not be isolated in upper-level courses for IT majors, according to Tobin and Ware (2005) and Werner (2005). This instruction should be taught to all graduates as a ‘security awareness’ course (Valentine, 2005) along with integrating it across through the curriculum (Yang, 2001).

**IMPLEMENTATION**

Our Business Data Analytics program is designed for students who wish to be data scientists or business analysts. These future employees will be handling large volumes of data on a continual basis, even more so than normal employees. Yet, due to unique requirements and regulatory compliance issues, it is essential that this data be secured. In order to do so these data scientists must often interact with enterprise security professionals regarding data access and security. While traditional approaches tend to focus on a definitional level of security training for the end user and very technical security training for those in information technology, the data scientist needs a solid understanding of both.

The goals of our Business Data Analytics Program are 1) to equip our major students, and others who choose Business Data Analytics courses as business electives, with the knowledge and skills to practice good security techniques on the job and at home from the end-user perspective; and 2) to develop an awareness and appreciation for the complexity of information security from a managerial perspective to be able to communicate with the
security team in the enterprise. With these goals in mind, we want our students to develop a fundamental and practical, rather than highly technical, understanding of threats and attacks in order to build a suitable context for defenses.

To accomplish these objectives, we have chosen to include security instruction in the six required and one elective Business Data Analytics courses. Doing so allows us to break down security into smaller modules and to discuss security as it best relates to the objectives of each course. We believe that a more consistent, constant emphasis across the curriculum will enable our students to see the connection to all facets of data analytics, as opposed to learning about security within the silo of a single course. In addition, this ensures that even those students who take only one or two data analytics courses as electives for their respective majors will be exposed to at least some aspects of security.

Using the topics covered in a popular security textbook (Ciampa, 2017), we developed a detailed outline of the topics most appropriate for a data analytics curriculum. Then, the faculty member for each course chose the topic(s) that fit(s) best with his/her course content and objectives. The table below shows the mapping of security topics to Business Data Analytics courses; the appendix provides a detailed description of four of these courses. To ensure an appropriate level of security instruction for future data scientists, we chose to not only define the various topics, but also to discuss their defenses.

<table>
<thead>
<tr>
<th>Course Title &amp; Description</th>
<th>Security Topic(s)</th>
</tr>
</thead>
</table>
| **BI310 Business Data Analytics**  
An introduction to the application of data analytics methods to business issues. Topics include business case studies, data analytics, model building techniques, and communications of results. | Introduction to Security  
- Challenges of Securing Information  
- Defining Information Security  
- Defining Attackers  
- Attacks and Defenses |
| **BI320 Web Analytics**  
An introduction to measuring, collecting, analyzing and reporting on online digital Web data using digital analytics and business intelligence. | Internet Security  
- How the Internet Works  
- Internet Defenses |
| **BI330 Structured Data Analysis**  
An introduction to the practical analysis and interpretation of different forms of data, emphasizing how and when to use particular tools, techniques, and metrics to maximize decision-making. | Workplace Security  
- Restricting Physical Access  
- Restricting Data Access  
- Crisis Preparedness |
| **BI410 Decision Support Systems Analysis and Design**  
An exploration of the analysis and design processes used to develop and deploy decision support systems (DSS) in businesses, which are technology-based tools that support decision-making activities. | Mobile Security  
- Mobile Attacks  
- Mobile Defenses |
| **BI350 Data Management**  
An introduction to managing the data used in business data analytics. Topics include data sources, acquisition, conditioning, storage, and security. | Computer Security  
- Attacks Using Malware  
- Computer Defenses |
| **BI420 Data Mining**  
An introduction to data mining and the demonstration of extraction principles from data stored in large heterogeneous volumes and how organizations can analyze data from multiple perspectives. | Privacy  
- Introduction to Privacy  
- Privacy Protections |
| **BI430 Data Visualization and Digital Dashboards**  
An introduction to the accumulation, analysis, and visualization of complex data sets for businesses, including the analysis of complex data sets and developing digital dashboards and scorecards. | Personal Security  
- Personal Security Attacks  
- Personal Security Defenses |
In some cases, the textbook chosen for the class contained one or more chapters on security. In other cases, customized materials had to be acquired and added to the course. While each faculty member developed the topic coverage in the way that best fit with the instructional design of his/her course, most chose a lecture-based instructional design for the initial implementation. Students were assessed on the security topics as part of their midterm or final course exams.

**DISCUSSION**

As noted previously, information security training and awareness are two of the most effective means to reduce the security risks posed by humans (Parsons, et al., 2014). In addition, many suggest that training and awareness are necessary at all levels of the organization, from the end user to the information technology professional (Macmanus, 2013, McDaniel, 2013, Werner, 2005). Yet, higher education tends to focus its security instruction on the end user or the information technology professional, with little regard for those who fall somewhere in-between, such as data scientists and business analysts. To that end, we chose to implement a distributed security instruction plan across the Business Data Analytics curriculum. By doing so, we are able to cover more aspects of information security in more detail. While our students do not necessarily need the most technical instruction, they do indeed require a more in-depth knowledge of security than the typical end user.

While we have just completed our first semester of this approach, the feedback from the faculty who taught this semester has been positive. For example, BI330 Structured Data Analytics included a chapter on workplace security from the text used in the course. Instruction was provided through lecture and class discussion. An example assignment from BI330 can be found in Appendix B. Student learning was assessed using a multiple-choice assignment and with specific security-related questions on a midterm exam. In addition, students were required to complete a book report for a professional book centered on data analytics. Most of those books included at least one chapter on data security. So, students were further exposed to security training through the completion of those reports. BI310 Business Data Analytics covered introductory security topics as supplemental discussion topics throughout the semester. To reinforce these topics, students had to address data security and privacy in their class projects involving data acquired from nonprofit organizations. Because the course textbook did not include a chapter on security, other secondary sources of information were used to provide instruction and resources. Instructors reported favorable outcomes regarding the student assessments, and faculty expressed an interest in increasing the richness and rigor of the security instruction moving forward. All departmental faculty are in favor of continuing this distributed approach for information security instruction and training.

As an extension to this research, we will develop a survey instrument to further assess the success of a distributed approach to security instruction and training. Appropriate versions of the survey instrument can be used to gather feedback and information from current and former students, faculty, and industry professionals. In addition, the data collected can assist our faculty in the refinement of the Business Data Analytics curriculum through the identification of additional security topic needs and the identification of strengths and weaknesses regarding our approach to security instruction.

**CONCLUSION**

Educating users, data scientists, and security professionals is equally important and necessary for organizations to build proper defenses and ward off attacks. Yet, while security training for users and security professionals is becoming increasingly common at the college level, few programs are addressing the security training needs of the data scientist. Nonetheless, the data scientist’s in-depth work with the organization’s data places him/her in a unique and critical position with regard to security. He/she must practice good end-user security while also working effectively with security professionals. To that end, our Business Data Analytics program has launched an initiative
To address this need in a different way. We have distributed security instruction and training across the required curriculum in an effort to provide an appropriate level of security training and a continuous emphasis on its importance for future data scientists. The feedback from the first semester of implementation was very positive, and plans are underway to continue this approach. While this initiative is an important beginning for security training in a data analytics curriculum, additional work is necessary. Some next steps include: 1) an investigation of security-related certifications and higher-level training that may be appropriate and valuable for college students preparing for a career in data analytics, and 2) additional research to assess the effectiveness of this initiative from both the student and employer perspective.

REFERENCES


APPENDIX A
SELECTED BUSINESS INFORMATICS SECURITY TOPICS

1. **Introduction to Security (Business Data Analytics BI 310)**
   a. Challenges of Securing Information
      i. Today’s attacks
      ii. Difficulties in defending against attacks
   b. What Is Information Security?
      i. Understanding security
      ii. Defining information security
      iii. Information security terminology
      iv. Understanding the importance of information security
         1. Preventing data theft
         2. Thwarting identity theft
         3. Avoiding legal consequences
         4. Maintaining productivity
         5. Foiling cyberterrorism
   c. Who Are the Attackers?
      i. Cybercriminals
      ii. Script Kiddies
      iii. Brokers
      iv. Insiders
      v. Cyberterrorists
      vi. Hactivists
      vii. State-Sponsored Attackers
   d. Attacks and Defenses
      i. Steps of an attack
      ii. Defenses against attacks

2. **Personal Security (Data Visualization and Digital Dashboards BI 430)**
   a. Personal Security Attacks
      i. Password Attacks
         1. Password weaknesses
         2. Attacks on passwords
      ii. Attacks Using Social Engineering
         1. Phishing
         2. Typo squatting
         3. Pretexting
         4. Hoaxes
         5. Dumpster diving
         6. Shoulder surfing
      iii. Identity theft
      iv. Social-networking risks
   b. Personal Security Defenses
      i. Password defenses
         1. Using password management tools
         2. Creating strong passwords
      ii. Recognizing phishing attacks
      iii. Avoiding identity theft
      iv. Setting social networking defenses

3. **Computer Security (Data Management BI 350)**
   a. Attacks Using Malware
      i. Circulation/Infection
1. Viruses
2. Worms
3. Trojan
   ii. Concealment
   iii. Payload Capabilities
      1. Execute commands
      2. Collect data
         a. Spyware
         b. Adware
         c. Ransomware
      3. Delete data
      4. Modify system security
      5. Launch Attacks

b. Computer Defenses
   i. Managing patches
   ii. Examining firewalls
   iii. Installing antimalware software
   iv. Monitoring User Account Control (UAC)
   v. Creating data backups
   vi. Recovering from attacks

4. Internet Security (Web Analytics BI 320)
a. How the Internet Works
   i. World Wide Web
   ii. Email

b. Internet Security Risks
   i. Browser vulnerabilities
      1. Scripting code
      2. Extensions
      3. Plug-Ins
      4. Add-Ons
   ii. Malvertising
   iii. Drive-by downloads
   iv. Cookies
   v. E-mail risks
      1. Spam
      2. Malicious attachments
      3. Embedded hyperlinks

c. Internet Defenses

APPENDIX B

The information below describes a data security assignment from one of our Business Data Analytics courses. This assignment works well to tie together the objectives of the course with the selected security topic. This assignment is taken from Kroenke, D.M. and Auer, D.J. (2015) Database Concepts, 7th Edition, Pearson Education, Inc., Hoboken, NJ, pp. 386-387. BI330 Structured Data Analysis focuses on the use of database tools for the management and use of data. Primarily, we use Microsoft Access 2013 to demonstrate the concepts and skills discussed and demonstrated. Each chapter in the text includes a section called the “Access Workbench”, and the text includes a chapter specifically on database administration and security.

The Access Workbench assignment for this chapter asks students to do the following:
1. Create a database security plan for the database they have created this semester.
2. Create a My-Trusted-Location folder, store a copy of the Access database in the My-Trusted-Location folder,
and test that the database can be opened without displaying the security bar warning.

3. Encrypt a copy of the database with a password, and open the database to test the success of the password.

4. Create a Digital Certificate. Create an AACDE version of the database, and then create a signed package using the database and the digital certificate.

This assignment allows students to test the security features available inside Microsoft Access. It also encourages a discussion of the need for these features, how and when they might be implemented, and the pros and cons of each.