Information Disclosure Incidents and Computing Education

Stefan A. Robila
Department of Computer Science, Montclair State University

Abstract

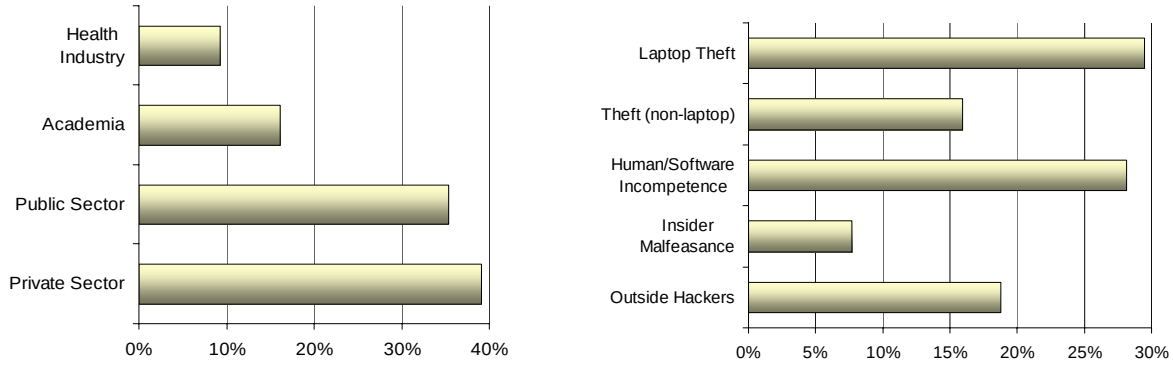
We present an introduction to security incident encountered by academic institutions and follow up with our approach to user education by infusing information disclosure incidents in two courses laying at the extremes of the computer science curriculum: a General Education Introduction to Computing and an Advanced Topics Information Security course. The choice of the two courses is such that, while in the Intro to Computing course the students view the incidents from the user's point of view (and are either victims of larger incidents or the disclosers of their own information), in the Information Security course the students view it mainly as computing professionals asked to prepare against and handle such events. In each course, theoretical and hands-on activities were developed to increase both the students' awareness and the acquiring of skills enabling to recognize and defend against such incidents. Special activities included the development of a phishing education module, invited guests discussing information disclosure incidents, analysis of term of use contracts, and practical penetration testing tasks. In both courses, we note that the students' understanding of the topics increased, allowing them to prevent information disclosure incidents from occurring or to better handle the recovery aspect when they occur.

1. Introduction

Information disclosure incidents, also known as data breaches, refer to situations when private information is improperly disclosed. Often, confidential data handled by various entities are compromised by loss, theft or pure negligence. Moreover, individual users fall prey to phishing attacks unwillingly providing their own personal data. Within the financial and health care sectors, such incidents have led to federal legislation mandating strict policies on how the confidential information is handled and severe legal implications when these policies are not followed. Consequently, most of the individual states have now enacted regulations mandating that security incidents be disclosed and individual parties be informed. An in depth analysis of the current legislation is provided in [1].

A significant increase in the disclosure and tracking of the security incidents has been noticed especially starting with 2005 following the "ChoicePoint incident". ChoicePoint, based Alpharetta, GA is a billion dollar company specialized in data aggregation for use in the banking and insurance industries as well as by law enforcement and state agencies. According to current reports, the company handles personal data (such as names, addresses, data of birth, social security information, credit reports, etc.) for over 220 million people and has commercial contracts with over 1000 economic agents and 7500 other various agencies [2]. While the company claims to have a system in place for screening the potential customers (in order to preserve the confidentiality of the data), in early 2005 it was revealed that confidential data associated to over 140,000 people have been sold to illegal organizations that further used it for identity theft [3]. Since then and up to date, various non-official statistics have counted over 1000 other data breaches totaling over 104 million potentially affected people [4].

A recent analysis of the data for 2006 alone reveals that the approximately 328 reported breaches have affected over 100 million users with entities from all the sectors being affected [5]. We note that the report indicates that in 90 of the incidents (or more than a quarter of all cases) the number of affected people was either not made public or was impossible to compute, leading us to believe that 100 million may be a seriously underestimating the problem. Figure 1 displays a percentage breakout of the incidents both on the type of the organization as well as on the type of incident.



(a) (b)

Fig 1. Data breaches recorded in 2006. a) percentage breakout based on the type of the organization, b) percentage breakout based on the type of the incident.

In this perspective, it is useful to remark that universities and colleges across the country are often found on security breaches lists making up for 16% of all the cases reported last year. An increase in information security events within the academia may lead to federal and state legislative efforts that would ultimately add upon the requirements in handling private information within the academic world. Unlike other entities, academic institutions however have to strike a delicate balance between the free circulation of ideas and the need for protection of personal information. With that in mind we have pursued to investigate two different aspects of the computers security within the academic realm. In the first phase, in the next section we are providing a statistical overview of the academic world security incidents and reaction to them. In the second phase, detailed in the third section we are presenting various educational modules and activities that were implemented in the curriculum, in an effort to increase the student population's understanding of the security issues in today's campus.

2. Data Breaches and Other Security Incidents in Academia

Unfortunately, up to this date, no official repository exists detailing the data breaches in colleges and universities. While in general, state legislation mandates that the affected parties be informed of the breach [1], no uniform requirement exists in detailing the number of cases and the exact nature of the incidents. Instead, researchers need to rely on a combination of news reports, press releases and private communications available in various forms and at various levels of detail.

However, more or less complete directories exist and maintained by either individuals or privacy rights advocates. Based on data retrieved from [5], [6], [7], and [8], we have identified over 119 incidents that are directly associated to academic units that affected private data associated to almost four million records. Figure 2 provides general statistics on the number of affected cases broken down by the year the incidents have occurred. We note that while the number of affected records has been relatively the same for the last two years, a decrease in the severity of the breaches may occur in 2007. On average, while the previous two years are marked by over 36,000 records / incident, 2007 so far has reduced to only a third (11,000 records / incident). As such, at first, one may conclude that academic institutions have increased their efficiency in protecting the private date. Unfortunately, a closer analysis reveals some disturbing trends. When taking in consideration the number of incidents reported (Figure 2c) we note that the monthly average has increased every year, meaning that today more institutions are affected by data breaches than ever before.

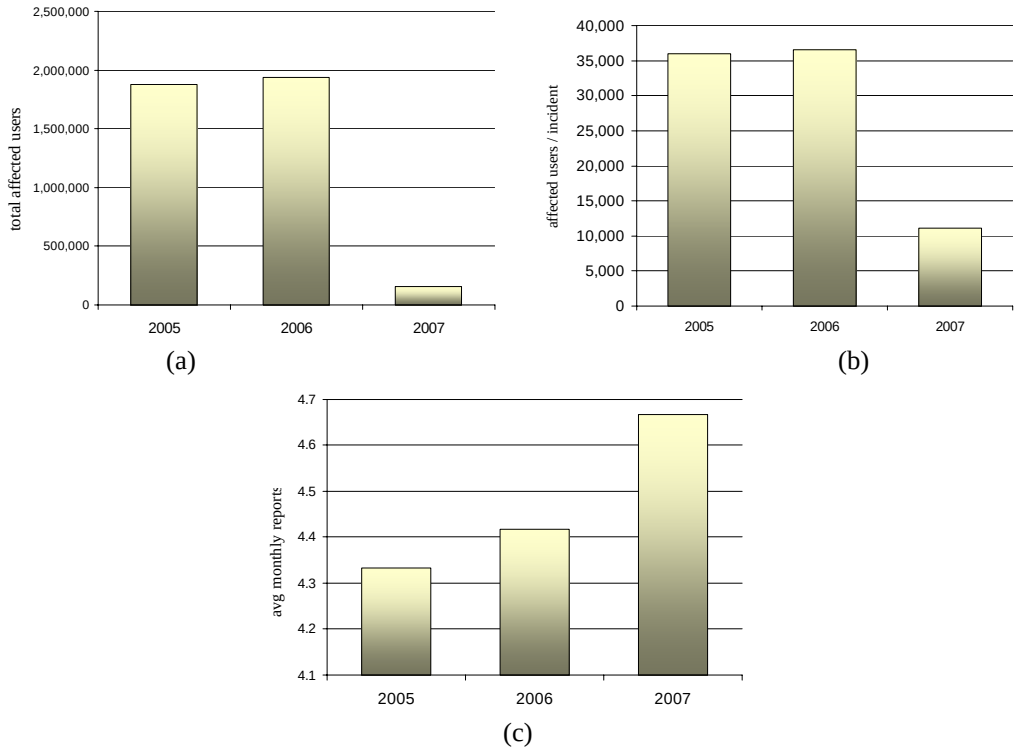


Fig 2. a) Total number of affected private records for academic institutions versus year of reporting, b) Average number of affected records per incident, and c) average number of incidents per month.

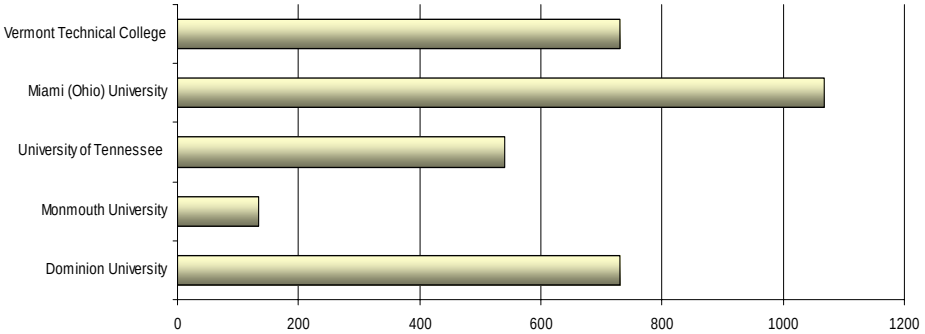


Fig 3. Sample of inadvertent disclosure incidents in which the data was posted on the web. The graph plots the number of days the data were publicly available.

Furthermore, both 2005 and 2006 were marked by spectacular incidents such as hacking attacks on University of Southern California (2005) and University of California Los Angeles (2006) that affected over one million records combined [9,10]. A similar incident this year would significantly increase the number of affected individuals. While most of the attacks can be attributable to computer hacking or theft, we also note that in 2006 a fifth of the incidents were described as 'human/software incompetence'. This is usually taking the form of either mailing or emailing confidential data to unauthorized sources or, more often, posting information in publicly accessible locations (such as websites). The severity of such incidents is indicated by the number of days, the data have stayed publicly available (see Figure 3). We note that in some cases the length has been for one or more years [5, 7]. Such incidents are particularly disturbing for the academic population since they are perceived as lapses in the university management [11, 12]. Finally, we note that some of the breaches were discovered by the staff or students themselves. Figure 4 shows the result on Google.com when searching for a student's Social Security Number. An excel spreadsheet was publicly available through the search engine for several months. Both the file and the link to the file are now deleted.

Fig 4. Sample search engine result when a student has used her (his) social security number.

In the light of such a significant number of vulnerabilities, one must consider ways that will support the students' understanding of the issues [13, 14]. While many aspects of the security incidents need to be addressed by the academic institutions as organizations (such as better security of the equipment and data, better staff training, etc.) it is also important to focus on the student body by supporting efforts to inform them better and educate them on privacy protection issues [15, 16].

3. Educational Initiatives on Security Incidents

In the following, we provide several educational activities that infuse information disclosure incidents in two courses laying at the extremes of the computer science curriculum: a General Education Introduction to Computing and an Advanced Topics Information Security course. The choice of the two courses is such that, while in the Intro to Computing course the students view the incidents from the user's point of view (and are either victims of larger incidents or the disclosers of their own information), in the Information Security course the students view it mainly as computing professionals asked to prepare against and handle such events.

3.1. Computer Security Course

According to the university catalogue, the course is a survey of topics related to internet and intranet security. It introduces the undergraduate students to many contemporary topics ranging from data encryption, computer authentication, network security, to cyber-warfare and security ethics. The course was developed based on Pfleeger & Pfleeger's textbook [17] with some materials from [18], following NSA recommendations on terminology and content. An important component of the course constitutes the practical assignment and the final projects. Each includes significant work both in application development as well as in writing. The course was initially offered as elective for upper level CS majors and graduate students, and constitutes the major security component of the program although other components are integrated throughout our curriculum. The current offering is required for undergraduate IT majors. The course size is usually approximately 18-25 students most in their senior year.

As initial assignment in the course, the students were asked to provide a survey *of the security issues* in campus. Suggested issues included the electronic mail system, the course delivery system (Blackboard), the student information system, access to labs and lab equipment and offices, library systems, etc. The students also had to analyze the terms of use agreement that the university requires the constituent members to abide by.

A review of the submitted work reveals that students perceive security depending on their immediate needs and interactions with the system. While most answers have included the student information system as main point of interaction, other aspects such as admission on the field as player in varsity games or falsification of parking permits have also surfaced. Students have accessed official university sources but have also relied on security discussion boards and blogs to extract additional information on the vulnerabilities of known applications (such as various operating systems or Blackboard). The strength of password protection was also addressed. The goal of this activity was to allow the students to understand the complexity of the academic information system and also to investigate the interdependency of various modules in sharing and protecting data.

An additional assignment is asking the students to expand their horizon and see *how security is handled throughout the academic world*. Here, the students had to search and analyze a number of recent data breaches. An important point was to analyze the reaction from the academic institution. The data are put back together in a compact format and combined with previous results obtained by students working on independent study projects (see Figure 5). The results for the students are significant. Based on 54 data breaches reported until May 2006, most academic institutions have not provided any significant support for the parties affected by the incidents. Only in four cases has the institution offered to pay for a credit monitoring service (even this limited to one year) and in some cases, the universities did not even notify the students directly but instead published general public releases.

3.1. Fluency Course

Approved to satisfy the General Education requirements for computer fluency in a liberal arts college, the Introduction to Computer Applications, Fluency with Information Technology course is shaped as a combination of lectures and hands on labs, with the lectures dedicated to various computing issues such as computer organization, structured programming, networks, privacy and ethics. The labs are focused on acquiring skills on various productivity packages as well as learning basic concepts in programming, web development and netiquette. The course is attended by large numbers of non-science majors most in their freshman year. The regular section size is 26.

Apart from regular modules that address computer privacy and E-commerce, we also developed a new strategy for educating users by combining phishing IQ tests and class discussions. Phishing is associated to electronic mail messages claiming to be from legitimate business and trying to attract the recipients to provide confidential information [19]. The technique that we used is based on a 12 item phishing IQ test that is asking the students to analyze a set of possible email messages and indicate whether they are phishing attempts or regular messages [20]. Following various class discussions and examples, a second test is administered. A full description of the technique is described in [21], with the results suggesting an increased level of awareness and better recognition of attacks. To increase our understanding on how efficient the test is, we have recently applied the same phishing survey to the computer security class. Figure 6 displays the correct identification rate for both classes with the checkered bars corresponding to fluency students and the solid bars corresponding to computer security students. The questions are divided in phishing attempts (Figure 6a) and legitimate (Figure 6b). We note that the computer security students have a better understanding of the phishing concept and recognize attacks at significantly higher rates compared to regular undergraduate students. However, their cautious approach is causing them to mislabel most of the legitimate messages as phishing.

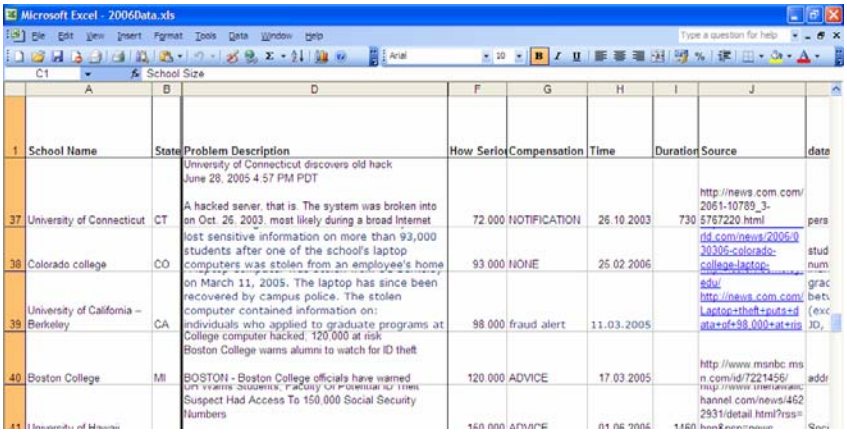
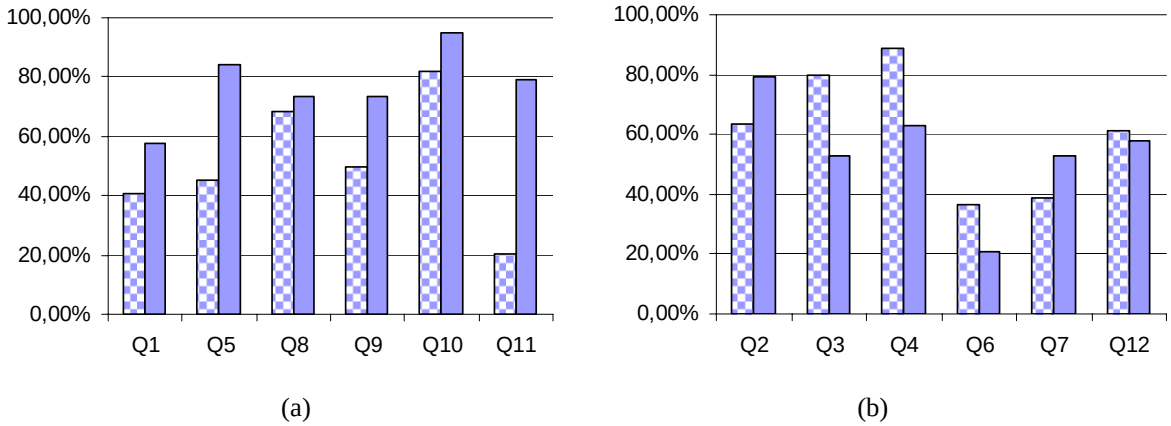


Fig 5. Sample database created by students to monitor security incidents within the academic world.



(a) (b)

Fig 6. Correct identification rates for fluency students vs. computer security students a) phishing attempts, b) legitimate messages.

4. Conclusions

We have presented several educational modules that are varied in approach and fit to many of the IT or CS courses currently offered. Through them, the students' understanding of the topics increased, allowing them to prevent information disclosure incidents from occurring or to better handle the recovery aspect when they occur.

The following motto, visible on the top of the ESI's list is an excellent characterization of the academic world: "Sometimes the free flow of information is unintentional" [8]. In recent years, the academic communities are facing an increasing number of security incidents. Over and over, when one analyzes the size of the phenomenon the results are shocking. Data are disclosed, lost, stolen, destroyed, and millions of students, alumni and staff have their confidential information put at risk. Our analysis is not different. Academic institutions are poorly equipped to handle confidential information that resides or is accessible from almost any computer system in campus. Most of the time, universities react to incidents by notifying either the public or the affected parties and do not provide any systematic paid support to assist the potential victims. In

this case, it becomes the responsibility of educators to provide to the students (i.e. the possible future victims) the necessary knowledge on how to protect, detect and react to data breaches.

References

[1] Consumers Union, Security Breach Notice Laws, www.consumersunion.org/campaigns/Breach_laws_May05.pdf, (accessed March 15, 2007).

[2] Wikipedia, ChoicePoint Entry, http://en.wikipedia.org/wiki/ChoicePoint, (accessed October 10, 2005).

[3] Scatlet S. D., 2005, The five most shocking things about the ChoicePoint debacle, CSO Magazine, May 2005, http://www.csoonline.com/read/050105/choicepoint.html, (accessed January 5. 2007).

[4] Privacy Rights Clearinghouse, A Chronology of Data Breaches Since the ChoicePoint Incident, http://www.privacyrights.org/ar/ChronDataBreaches.htm, (accessed March 10. 2007).

[5] Privacy Rights Clearinghouse, Chronology of Data Breaches 2006: Analysis, http://www.privacyrights.org/ar/DataBreaches2006-Analysis.htm, (accessed March 10. 2007).

[6] Attrition.org Data Loss Archive and Database, http://attrition.org/dataloss/, (accessed January 5. 2007).

[7] Hasan, R. and Yurcik, W. 2006. A statistical analysis of disclosed storage security breaches. In Proceedings of the Second ACM Workshop on Storage Security and Survivability StorageSS '06. ACM Press, New York, NY, 1-8.

[8] Adam Dodge, Educational Security Incidents (ESI), http://www.adamdodge.com/esi/, (accessed March 20, 2007)

[9] North County Times, 2006, San Diego man charged with accessing university applicants' information, http://www.nctimes.com/articles/2006/04/21/news/sandiego/17_01_084_20_06.txt, (accessed October 10, 2006).

[10] CBS News, 2006, UCLA data breach leaves 800K at risk, http://www.cbsnews.com/stories/2006/ 12/12/tech/main2249716.shtml, (accessed February 14, 2007).

[11] Eyeopener Online, 2007, Student privacy violated - marks and names put online, http://www.theeyeopener.com/article/3277, (accessed March 14, 2007).

[12] WRAL, 2007, ECU mistakenly posts personal info online, http://www.wral.com/news/local/story/ 1198897/, (accessed March 1 2007).

[13] Rollason-Reese, R. L. 2003. Incident handling: an orderly response to unexpected events. In Proceedings of the 31st Annual ACM SIGUCCS Conference on User Services SIGUCCS '03. ACM Press, New York, NY, 97-102.

[14] Masuya, M., Yamanoue, T., and Kubota, S. 2006. An experience of monitoring university network security using a commercial service and DIY monitoring. In Proceedings of the 34th Annual ACM SIGUCCS Conference on User Services SIGUCCS '06. ACM Press, New York, NY, 225-230.

[15] DeWitt, J. and Cicalese, C. 2006. Contextual integration: a framework for presenting social, legal, and ethical content across the computer security and information assurance curriculum. In Proceedings of the 3rd Annual Conference on information Security Curriculum Development InfoSecCD '06. ACM Press, New York, NY, 30-40.

[16] Pothamsetty, V. 2005. Where security education is lacking. In Proceedings of the 2nd Annual Conference on information Security Curriculum Development InfoSecCD '05. ACM Press, New York, NY, 54-58.

[17] Pfleeger, C.,S. 2003, Pfleeger, Security in Computing, 3rd Prentice Hall.

[18] Stallings, W., 2003, Network Security Essentials, 2nd Ed., Prentice Hall.

[19] Dhamija, R., Tygar, J. D., and Hearst, M. 2006. Why phishing works. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. R. Grinter, T. Rodden, P. Aoki, E. Cutrell, R. Jeffries, and G. Olson, Eds. CHI '06. ACM Press, New York, NY, 581-590.

[20] Robila, S. A. and Ragucci, J. W., Montclair Phish, http://csam.montclair.edu/~robila/RSL/Phish/, (accessed June 5, 2006).

[21] Robila, S. A. and Ragucci, J. W. 2006. Don't be a phish: steps in user education. In Proceedings of the 11th Annual SIGCSE Conference on innovation and Technology in Computer Science Education, ITICSE '06. ACM Press, New York, NY, 237-241.