

In this issue:

- 4. Consumer Acceptance of Biometric Credit Cards as an Identify Proofing Mechanism**
Laura Poe, Longwood University

- 12. Teaching Public Key Cryptography: A Software Approach**
David Carlson, Saint Vincent College

- 21. Teaching Case**
Digital Forensics and Incident Response (DFIR): A Teaching Exercise
Jennifer L. Breese, Penn State Greater Allegheny
Maryam Roshanaei, Penn State Abington College
J. Andrew Landmesser, Penn State Brandywine
Brian Gardner, Penn State Schuylkill

- 35. Phishing: Gender Differences in Email Security Perceptions and Behaviors**
Jie Du, Grand Valley State University
Andrew Kalafut, Grand Valley State University
Gregory Schymik, Grand Valley State University

- 48. Feasibility of Creating a Non-Profit and Non-Governmental Organization Cybersecurity Incident Dataset Repository Using OSINT**
Stanley J. Mierzwa, Kean University
Iassen Christov, Kean University

The **Cybersecurity Pedagogy and Practice Journal (CPPJ)** is a double-blind peer-reviewed academic journal published by **ISCAP** (Information Systems and Computing Academic Professionals). Publishing frequency is two times per year. The first year of publication was 2022.

CPPJ is published online (<https://cppj.info>). Our sister publication, the proceedings of the ISCAP Conference (<https://proc.iscap.info>) features all papers, panels, workshops, and presentations from the conference.

The journal acceptance review process involves a minimum of three double-blind peer reviews, where both the reviewer is not aware of the identities of the authors and the authors are not aware of the identities of the reviewers. The initial reviews happen before the ISCAP conference. At that point, papers are divided into award papers (top 15%), and other accepted proceedings papers. The other accepted proceedings papers are subjected to a second round of blind peer review to establish whether they will be accepted to the journal or not. Those papers that are deemed of sufficient quality are accepted for publication in the CPPJ journal.

While the primary path to journal publication is through the ISCAP conference, CPPJ does accept direct submissions at <https://iscap.us/papers>. Direct submissions are subjected to a double-blind peer review process, where reviewers do not know the names and affiliations of paper authors, and paper authors do not know the names and affiliations of reviewers. All submissions (articles, teaching tips, and teaching cases & notes) to the journal will be refereed by a rigorous evaluation process involving at least three blind reviews by qualified academic, industrial, or governmental computing professionals. Submissions will be judged not only on the suitability of the content but also on the readability and clarity of the prose.

Currently, the acceptance rate for the journal is under 35%.

Questions should be addressed to the editor at editorcppj@iscap.us or the publisher at publisher@iscap.us. Special thanks to members of ISCAP who perform the editorial and review processes for CPPJ.

2024 ISCAP Board of Directors

Jeff Cummings
Univ of NC Wilmington
President

Amy Connolly
James Madison University
Vice President

Eric Breimer
Siena College
Past President

Jennifer Breese
Penn State University
Director

David Gomillion
Texas A&M University
Director

Leigh Mutchler
James Madison University
Director/Secretary

RJ Podeschi
Millikin University
Director/Treasurer

David Woods
Miami University
Director

Jeffry Babb
West Texas A&M University
Director/Curricular Items Chair

Tom Janicki
Univ of NC Wilmington
Director/Meeting Facilitator

Paul Witman
California Lutheran University
Director/2024 Conf Chair

Xihui "Paul" Zhang
University of North Alabama
Director/JISE Editor

Copyright ©2024 by Information Systems and Computing Academic Professionals (ISCAP). Permission to make digital or hard copies of all or part of this journal for personal or classroom use is granted without fee provided that the copies are not made or distributed for profit or commercial use. All copies must bear this notice and full citation. Permission from the Editor is required to post to servers, redistribute to lists, or utilize in a for-profit or commercial use. Permission requests should be sent to editorcppj@iscap.us.

CYBERSECURITY PEDAGOGY AND PRACTICE JOURNAL

Editors

Anthony Serapiglia
Co-Editor
Saint Vincent College

Jeffrey Cummings
Co-Editor
University of North Carolina
Wilmington

Thomas Janicki
Publisher
University of North Carolina
Wilmington

2024 Review Board

Cheryl Beauchamp
Regent University

Ulku Clark
Univ of NC Wilmington

Peter Draus
Robert Morris
University
Jeff Landry
Univ of South Alabama

Nick Giacobe
Penn State University

Mike Hills
Penn State University

Li-Jen Lester
Sam Houston State
Univ

Jim Marquardson
Robert Morris University

Stan Mierzwa
Kean University

Etezady Nooredin
University of New
Mexico

Ron Pike
Cal Poly Pomona

RJ Podeschi
Milliken University

Samuel Sambasivam
Woodbury University

Kevin Slonka
Saint Francis University

Geoff Stoker
Univ of NC Wilmington

Paul Wagner
University of Arizona

Ping Wang
Robert Morris University

Tobi West
Coastline College

Johnathan Yerby
Mercer University

Consumer Acceptance of Biometric Credit Cards as an Identify Proofing Mechanism

Laura Poe
laurapoe@verizon.net
Longwood University
Farmville, VA 23909

Abstract

Biometric credit cards have entered the marketplace as an enhancement to the chip card technology for authenticating consumers when making a purchase at a credit card terminal. Consumers using a physical credit card have the capability to provide authentication using a pre-registered fingerprint stored on the card that is compared with the fingerprint used at the time of purchase. The success of the biometric advancement will be impacted by marketplace user acceptance. Cyber vulnerabilities on biometrics through similarity-based attacks and other methods are explored in relation to the impact on consumers' data privacy. After making purchases with a test product of a biometric credit card, consumer attitudes and reactions were measured using a survey instrument to determine the acceptance of biometric credit cards in the marketplace. The results of the research indicated that overall, consumers find the biometric credit card to add to the financial security of physical credit card transactions and are not a privacy concern. This research provides a quantitative analysis of user attitudes towards fraud, reaction to biometric credit cards, and predictive analysis of consumer acceptance of biometric cards for identity proofing.

Keywords: biometrics, identity proofing, consumer attitudes, fraud.

Recommended Citation: Poe, L., (2024). Consumer Acceptance of Biometric Credit Cards as an Identify Proofing Mechanism. *Cybersecurity Pedagogy and Practice Journal*; v3 (n2) pp 4-11.
<https://doi.org/10.62273/JUPA7528>

Consumer Acceptance of Biometric Credit Cards as an Identify Proofing Mechanism

Laura Poe

1. INTRODUCTION

Behavior-driven algorithms are commonly used for fraud detection in the financial services industry. Credit card fraud losses will total approximately \$165.1 billion by 2033, impacting all age groups across the United States (Nilson, 2023). Numerous fraud detection and prevention methods exist to detect fraudulent transactions before they occur, but fraud detection models are generally considered company proprietary information, making the analysis of the various methods challenging. Regardless of the fraud detection model used, the consumer is typically blind to the process until a purchase triggers a fraud alert.

An alternative to fraud detection models are biometric credit cards that are embedded with a registered fingerprint, providing real-time authentication and identity proofing. Studies indicate a reduction in physical card fraud from 31.5% of card transactions to less than 2% when using biometric cards (Poe, 2021). While this is a significant reduction in projected fraud, the success of biometric cards will be impacted by consumers' willingness to provide biometrics on a transactional basis. Consumer-facing methods, such as signature-based purchases are not widely leveraged. The most widely used physical card authentication method was the introduction of the (Europay, Mastercard, Visa) EMV chip card, which added the additional element of security by generating a unique code for the transaction that replaces the actual card number. However, the EMV chip can be combined with additional user authentication to prevent unauthorized users from initializing the transaction. Given most credit card purchases are made with physical credit cards, the lack of adequate identity proofing is a known security gap. Leveraging biometric cards provides the benefit of both the EMV chip and a strong identity proofing mechanism. This research evaluates the user response and attitudes to biometric credit cards after using a biometric credit card for a purchase.

2. BACKGROUND AND RELATED WORK

Despite the efforts to prevent fraud through EMV chip cards and back-end fraud detection models, verifying the identity of a person at transaction initiation remains the most difficult but most important step in preventing fraud. The introduction of the iPhone 5 and the capability to lock and unlock the phone using a fingerprint was instrumental in cultural changes and the way biometric fingerprint authentication was viewed. Google Wallet and ApplePay capitalized on this feature, providing a way for financial transactions to be secured by using the fingerprint in the phone for authorization. ApplePay's growth in the marketplace shows rapidly growing consumer demand in using biometrics when performing financial transactions. Since the release of the iPhone 8, facial recognition is used to authenticate the consumer. Facebook's DeepFace, which uses a neural network approach with a high-capacity model, obtained an accuracy of 97.35% on LFW benchmark (Taigman, 2014). However, in-store purchases require a less bulky system for purchase transactions.

Financial institutions have the opportunity to capitalize on the culture shift and utilize a biometric-enabled physical credit card device to enhance the current physical card, leading to significant reductions in the amount of fraud related to counterfeit and lost/stolen cards. Mobile payments, both in-apps and in-retail stores, have been a major contributor to the adoption of biometrics. The need for authentication speed coupled with the ability to include payment authentication to contactless payments has resulted in fingerprint biometrics becoming the standard (Goode Intelligence 2015).

Leveraging biometrics for credit card purchase authentication is achieved by embedding the fingerprint into the credit card. A study previously conducted on biometric credit cards indicated 0.71% lower error rates than indicated in previous studies of fingerprint accuracy, and an overall the biometric card yielded a flat reduction in fraud of 30.5% of physical credit card fraud (Poe, 2021). While various types of

fraud detection models attempt to catch fraud as it occurs, successful identity proofing prevents the ability for a lost/stolen card to be used by an unauthorized user. The effectiveness of the biometric-enabled credit card must be greater than the overall concerns of consumers and must demonstrate the ability to prove the identity of the card holder at the time of purchase.

Digital forensics and the use of fingerprints impact the individual's privacy beyond the consumer standpoint (Kaye, 2003). The concern of privacy is further impacted by the ability to protect the fingerprint from cyberattacks during the comparison of the live fingerprint with the stored template. Similarity-based attacks, leveraging Kerckhoff's principle of public knowledge of both the function and template, the template itself is not protected and is vulnerable to attack. Based on the similarity index used in the comparison of the stored template and the actual fingerprint, a determination of authenticity is established. If the similarity index is too broad, the digital biometric can be reverse engineered, revealing the original template.

Corporate Responsibility to Consumers

Privacy protection of the data is the responsibility of the party collecting the information from the individual/consumer and are a serious concern in the design of biometric identity proofing and authentication systems. The uniqueness of the traits increases the criticality of protecting the data. When considering privacy, the value of security and convenience typically supersedes the value in safeguarding biometric data. During the authentication process, a user claims an identity by providing biometric information to a system for comparison against the stored references. In the case of surveillance applications, the process differs only that the system initiates the comparison rather than the user (Krishnan, 2012).

Companies, such as Busch Gardens and Disney Theme Parks, collect biometric data for entrance to the park in effort to track members and limit the membership fraud resulting from sharing of annual membership cards. The membership systems store the member's fingerprint data as well as photographs. Upon park entry, a member must scan the same finger each time, which is compared to the fingerprint in the system for a match. If authenticated successfully, the member gains entry into the park. Additionally, photographs of the members are stored to ensure the photograph on the account record matches the person entering the park. In these cases, the theme parks have the responsibility

for storing and protecting the biometric data of their members. Their cybersecurity measures become critical components in the protection of this data.

Numerous governmental programs utilize biometric data, specifically fingerprint data, such as First Capture. First Capture is a multi-agency governmental program working to develop technology designed to capture ten rolled-equivalent fingerprints in less than 15 seconds. The focus is to ensure high quality of the fingerprint image with a device that is portable. The Integrated Automatic Fingerprint Identification System (IAFIS) contains over 47 million fingerprints and includes the electronic exchanges of fingerprints (Melodia, 2015). Governments have equal responsibility in maintaining and protecting biometric data.

3. CASE STUDY AND RESEARCH OBJECTIVES

The purpose of this study was to determine the user perception of consumers using a biometric card as a means for reducing credit card fraud and the applicability of a biometric card for the general credit card market. The study will provide specific data related to the survey of users after using a biometric credit card.

Research Methods

The study was performed by conducting a survey of 200 participants after their first-time use of a biometric credit card. The survey instrument was designed to measure the following categories: consumer perceptions of credit card fraud, ease of use of the biometric card, consumer attitudes towards risk, consumer attitudes towards identity proofing, and consumers attitudes towards privacy. The questions were divided based on these categories, but the question sequence was inconsequential and deemed to have no impact on the outcome of the responses based on the utilization of Likert scale-based questions (Weng and Cheng, 2000).

Performing the survey simultaneously with the biometric card experiment facilitated the timely capture of information while, also, providing the same participant base for both the experiment and the survey. The chances of survey participation were nearly 100%, since the survey was completed immediately following the credit card experiment. The goal of the survey was to find correlations among the categories. The survey was self-administered immediately following the successful registration of the fingerprint to the biometric credit card, a

successful purchase, and an attempted fraudulent purchase.

Each participant completed the survey, but three percent of participants were unable to successfully register their biometric card during the initial card experiment, which could result in a negative experience, reflected in the survey results.

Population and Sample

The population completing the survey consisted of a convenience sample of 200 people from a shopping mall located in Glen Allen, Virginia without regard to demographic criteria. Participants in the study were selected to use a biometric credit card and take a survey following the use of the card. The shopping location provided a strong sample of the population who would be using a physical credit or debit card and could be potential users of the biometric card.

Survey Instrument & Criteria

The survey instrument was created following Creswell’s (2013) strategy for the development of a mixed methods study to incorporate the qualitative analysis with the quantitative analysis. This study focused on the consumer’s attitudes towards fraud and the use of biometrics as a means to prevent credit card fraud in conjunction with the actual fraud detection rate when using the biometric card. The survey instrument was designed to measure the following categories: consumer perceptions of credit card fraud, ease of use of the biometric card, consumer attitudes towards risk, consumer attitudes towards identity proofing, and consumers attitudes towards privacy. The questions were divided based on these categories, but the question sequence was inconsequential and deemed to have no impact on the outcome of the responses based on the utilization of Likert scale-based questions (Weng and Cheng, 2000).

Performing the survey simultaneously with the biometric card experiment facilitated the timely capture of information while, also, providing the same participant base for both the experiment and the survey. The chances of survey participation were nearly 100%, since the survey was completed immediately following the credit card experiment. The goal of the survey was to find correlations among the categories. The survey was self-administered immediately following the successful registration of the fingerprint to the biometric credit card, a successful purchase, and an attempted fraudulent purchase.

The survey was categorized into five main areas to support the research objectives, Table 1: fraud perceptions, ease of use, attitude toward identity proofing, attitudes towards risk, and attitudes towards privacy. Using this criteria, previous individuals who identified as victims of fraud were evaluated as well as age categories.

Variable Name	Variable Type
Positive Fraud Experience	Dependent
Age Category	Dependent
Credit Card Ownership	Dependent
Biometric Card Device False Positive Result	Dependent
Attitude Toward Identity Proofing	Independent
Fraud Perceptions	Independent
Ease of Use	Independent
Consumers’ Attitudes Toward Data Privacy	Independent

Table 1: List of Variables

Data Analysis Methods & Design

A convenience sample was used for both the biometric card experiment and the survey instrument in order to maximize participation and target brick and mortar shoppers. The results of the study may limit the transferability of results to other geographic locations. However, based on the cost of obtaining the biometric cards and the coordination with the registration of those cards by an industry subject matter expert, the convenience sample provided the most feasible solution for obtaining the data. The data analysis is unaffected by the convenience sample, although noted for informational purposes. The study was guided by research questions employing quantitative analyses through the experiment followed by survey results. External reporting provided sustainable comparisons of existing successful fraud rates for the prior two years.

The following research question was evaluated as part of the study.

RQ: Are consumers attitudes towards biometric credit cards supportive in order to reduce credit card fraud?

RO1 Evaluate the consumers’ attitudes towards corporate responsibility in reducing fraud.

RO2: Evaluate the consumers' attitudes towards using a biometric credit card for purchases

4. RESULTS

Research Question: Are consumers attitudes towards biometric credit cards supportive to reduce credit card fraud?

In addition to the results of the physical biometric card experiment, the consumer perception of fraud and the biometric card is important to explore the themes and issues to be addressed by financial institutions seeking to utilize biometric credit cards as a future product. The use of closed questions, indicated by selecting a response provided by the researcher, were applied to determine categorical variables. In order to measure the consumer attitudes towards cards, the following categories were measured: attitude towards identity proofing, fraud perceptions, ease of use, and consumers' attitudes towards data privacy. These were evaluated against the following dependent variables: previous positive fraud experience, age category, credit card ownership, and the corresponding biometric card device false positive result for the participant.

A three-way ANOVA was used to determine the relationship between previous fraud experience, age, and the biometric card device false positive result from the experiment to the consumers' attitudes towards identity proofing, fraud perception, and ease of use of the biometric card. Additionally, descriptive statistics were analyzed, comparing the ease of use scores to the participants' age. Age was evaluated for statistical significance in ease of use and the perceived reduction of fraud.

Each of these data points provided quantitative evidence of the viability of biometric credit cards to provide high authenticity identity proofing and the expected future impact on the rate of successful credit card fraud transactions for lost/stolen physical cards. Additionally, the data provides statistical evaluation of the consumers' perceptions of fraud and prospective use of biometric cards.

Categorical questions were included in the survey instrument to gain the consumer's perspectives on fraud, biometrics, privacy, and corporate responsibility. Participants were asked to designate a score for each question based on the levels 1-5, as follows: 1 – mostly disagree, 2

– disagree, 3 – neutral, 4 – agree, and 5 – mostly agree. Additional demographic data, such as gender and nationality, was captured to determine a relationship to age and previous fraud victims.

In each of the five categories, the mean results were calculated based on the total participant pool of 200. The mean score was further evaluated against age and victim identity. In the first category of fraud perceptions, Table 2, the overall majority found that fraud is a growing problem but did not feel confident that current strategies in the financial services industry are successful in preventing fraud. Participants seem to be somewhat neutral when determining if banks should use biometric identity proofing for fighting fraud. The majority agreed that successful prevention of fraud increases their trust in the financial institution. Evaluating successful prevention is difficult, since most consumers are unaware of the percentage of fraud attempts blocked by financial institutions. The theme of the first section is the participants held the belief that fraud is a problem, and prevention is expected from financial institutions to earn the consumer's business.

The second category, ease of use, Table 3, demonstrated that biometrics were not cumbersome, and a slight majority found them both easy and safe. The third category focused on how the consumer views biometrics as a form of identity proofing. Most participants believe that biometrics make it more difficult for fraudsters to steal identities. At the time of this study, fingerprint and facial recognition were a method for cell phone users to make purchases using ApplePay or GooglePay. This is reflected in the comfort level of using a mobile phone's biometric for making purchases. The overall attitude towards biometrics is slightly positive.

The last two categories focused on risk and privacy. More participants believed there is increased risk in credit card fraud as compared to risk in providing biometric data in a transaction, Table 4. Participants did not find the collection of biometric data to be invasive or a violation to their privacy.

FRAUD PERCEPTIONS						EASE OF USE	
Fraud Growing Problem	Successful Strategies	Bank Selection	Use Biometrics	Fraud Trust	Prevention Trust	Biometric Cumbersome	Biometric Easy Safe
4.25	2.56	3.67	3.18	3.84	4.13	2.49	3.81

Table 2. Fraud perceptions and ease of use

ATTITUDE TOWARD IDENTITY PROOFING						
Financial Security	Biometric Decreases Fraud	Mobile	Biometric Identity	Biometric Deters	Biometric Increases Identity Theft	Biometrics Use Stolen Identities
	3.59	3.60	3.90	3.84	3.40	2.57
						3.85

Table 3. Consumer attitudes towards biometrics

ATTITUDE TOWARD RISK						ATTITUDE TOWARD PRIVACY
Consumer Comfort	Consumer Comfort Mobile	institution Trust Bio Data	Biometric Risks	Bio Less Risk Fraud	Back-Up Verification	Biometric Invasive
	3.71	3.59	2.27	3.35	2.93	3.97
						2.63

Table 4. Consumer attitudes towards risk and privacy

To evaluate the consumers’ attitudes towards corporate responsibility in reducing fraud, a multivariate analysis of variance (MANOVA) test was performed to determine the relationships between age, victims of fraud, attitudes towards credit card bank selection based on a company’s ability to combat fraud, and attitudes towards recommending that companies use biometrics as a means of identity proofing for preventing fraud. The solid distribution between age categories and fraud victims allowed for an analysis by each age group and victim category.

Further analysis through Box’s test of equality of covariance proves statistical significance of age and victims of fraud with corporate responsibility for fraud prevention and utilizing biometrics. Levene’s test, Table 3, supports the results of Box’s test with statistically significant results. Ages 18-24 and 45-54 had the highest scores of Agree to Strongly Agree that fraud prevention tactics by a company were crucial in selecting a bank. Regardless of fraud victimization, the outcomes of the analysis postulate that a bank’s ability to prevent fraud is important in the selection of a financial institution for obtaining a credit card. However, the lack of a statistical relationship between age and selection of a financial institution ruled out any predictive relationship. The overall results remain an indicator that nearly all age groups found fraud prevention a consideration in credit card company selection.

RO2: Evaluate the consumers’ attitudes towards using a biometric credit card for purchases.

The majority of participants who were victims of fraud held a positive attitude towards using biometrics for identity proof during credit card purchases. A linear regression analysis was performed to analyze only the relationship between fraud victims and attitudes towards using biometrics for identity proof during credit card purchases. The relationship between those who think fraud is a growing problem to biometrics as a deterrent was statistically significant. Removing the variable fraud victim allowed for more targeted calculations to determine if a predictor relationship exists. The R square of fraud victims was not high enough to substantiate a predictor relationship.

Nearly unanimously, respondents believed that biometric cards are easy and safe to use for making a purchase transaction. However, respondents were not as agreeable to requiring biometric data as part of a credit card application. The resistance was not in providing biometric data but more focused on capturing the biometrics during the application process. In this study, no credit card application was required to register and use the card.

The majority of the participants responded in disagreement that requiring the biometric print during the application process would be too cumbersome. Recognizing the user’s perception of the level of difficulty using the biometric card is essential to determine the influence between the user’s perception and the marketability of the card. As a sample population of credit card holders, the level of resistance to using credit cards based on any complexity with capturing fingerprints seems minimal in impact. Organizations can, however, focus on reducing the impact to card holders by developing a seamless registration process and potentially utilizing devices that allow consumers to register cards from their own homes.

More than half of respondents had experienced some form of credit card fraud. However, recommendations to use biometrics as a means for fraud prevention remained neutral, eliminating a causal relationship between the two variables. When evaluating demographics, age was a significant factor in relation to biometrics as an invasion of privacy. Participants over the age of 45 had privacy concerns on the collection of biometric data, though they felt that fraud was a growing problem, and the requirement of a fingerprint would lead to reduced numbers of fraud occurrences.

Concerns over personal privacy were addressed as part of the survey, and overall, participants did not feel that a biometric credit card violated their privacy. The 35-44 age group were neutral on privacy, and the age categories of 18-24, 25-34, and 45-54 equally disagreed on the question addressing "requiring biometric fingerprint data is invasive and violates the right to privacy". While ages 35-44 were neutral on privacy, they disagreed that risks involved in providing fingerprint data are less than the risks of fraud. The 45-54 age group felt different and agreed that risks of using biometric data were less than the risk of fraud. The younger group, from 18-34 were neutral, indicating they did not believe the biometrics were any riskier than existing chances of fraud.

5. CONCLUSIONS

The purpose of this study was to evaluate the viability of the biometric credit card by exploring the perceptions and attitudes of consumers towards using biometrics for identity proofing during a credit card transaction. The study provided evidence of consumer's acceptance of the biometric credit card. In general, consumers feel that fraud is a growing problem and believe that using biometrics will result in credit card fraud reduction. Biometric cards were found to be easy to use and more secure than current means of authentication. Numerous statistical analyses were performed to determine the relationship between fraud victims and biometrics as well as age and biometrics. Linear regression analyses were performed as well as a multivariate analysis to determine predictability associations.

The survey conducted gathered data for understanding consumers' attitudes towards fraud and using biometrics to combat fraud. The study was further evaluated based on age group and consumers who had or had not experienced credit card fraud. Fraud victims believe fraud is a growing problem and are more likely to believe that biometrics should be used to combat fraud. Consumer perceptions were measured based on age category for fraud prevention tactics. Ages 18-24 and 45-54 had the highest scores of Agree to Strongly Agree that fraud prevention tactics by a company were crucial in selecting a bank. Those who believe fraud is a growing problem believe biometrics reduce fraud. Little resistance to biometrics could be found, as nearly all participants responded in favor of biometrics for identity proofing and declared a biometric card as easy to use. All age groups found a company's approach to and guarantee of fraud protection and prevention important when selecting a credit card. The ease of use with the biometric

card was reflected in the survey responses by the participants.

The results of this study are applicable to biometric card industry in determining user acceptance and usability for implementing biometric credit cards into the marketplace. Additionally, the results provide further evidence of the market demand for enhanced security measures for the banking and financial industry.

6. RECOMMENDATIONS FOR FURTHER STUDY

The survey was conducted as a joint experiment with the biometric credit card. The conjoining of these two aspects of the study did not allow for survey respondents who had never seen or experienced a biometric credit card and could have concluded differing results. Additionally, the participants in the study opted to take part based on their interest in biometrics. Those uninterested in utilizing biometric cards were more unlikely to participate, creating some level of bias. The bias created an inherent limitation to the survey results. Future research could be conducted from a random sample of participants with no knowledge or experience using biometric credit cards. Place before the references.

7. REFERENCES

- Goode Intelligence. 2015. Biometrics for Banking; Market and Technology Analysis, Adoption Strategies and Forecasts 2015-2020. Retrieved from <http://www.goodeintelligence.com/report-store/view/biometrics-for-banking-market-technology-analysis-adoption-strategies-forecasts-20152020>
- Kaye, D. (2003). *The Nonscience of Fingerprinting: United States v. Llera-Plaza*. Retrieved from Law Review Library: [https://www.qu.edu/prebuilt/pdf/SchoolLaw/LawReviewLibrary/43_21QLR1073\(2001-2003\).pdf](https://www.qu.edu/prebuilt/pdf/SchoolLaw/LawReviewLibrary/43_21QLR1073(2001-2003).pdf)
- Krishnan, A. P., & Sy, B. K. (2012). SIPPA-2.0 – Secure Information Processing with Privacy Assurance (version 2.0). *Tenth Annual International Conference on Privacy, Security and Trust*, 25-34.
- Melodia, M., Bond, P., & Angelovska-Wilson, A. (2015). Legal Risks and Rules of the Move to Biometrics. *New York Law Journal*, 1-2.

- Nilson Report. (2023, April) *Acquisitions and Financing Deals in Payment Cards—2022*. Retrieved April 18, 2023, from https://nilsonreport.com/publication_newsletter_archive_issue.php?issue=1239#
- Papadimitriou, O. (2015, October 17). *Where Does Apple Pay Stand On Its First Birthday*. Retrieved from Tech Crunch: <http://techcrunch.com/2015/10/17/where-does-apple-pay-stand-on-its-first-birthday/>
- Poe, Laura F. (2021) Case Study: Empirical Evaluation of a Biometric Credit Card for Fraud Reduction, *Cybernetics and Systems*, DOI: 10.1080/01969722.2021.2005873
- Spraggs, D. (2007, February 1). *How to Lift Fingerprints*. Retrieved March 27, 2018, from <http://www.policemag.com/channel/patrol/articles/2007/02/how-to-lift-fingerprints.aspx>
- United States Census Bureau. (2017, July 1). *Quick Facts United States*. Retrieved 17 November, 2018, from <https://www.census.gov/quickfacts/fact/table/US/PST045217>.
- Weng, L-J., & Cheng, C-P. (2000). Effects of Response Order on Likert-Type Scales. *Educational and Psychological Measurement*, 60 (6). 908-924.
- Y. Taigman, M. Yang, M. A. Ranzato and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification", *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1701-1708, 2014.
- X. Dong, Z. Jin and A. T. B. Jin, "A Genetic Algorithm Enabled Similarity-Based Attack on Cancellable Biometrics," *2019 IEEE 10th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, Tampa, FL, USA, 2019, pp. 1-8, doi: 10.1109/BTAS46853.2019.9185997.

Teaching Public Key Cryptography: A Software Approach

David Carlson
david.carlson@stvincent.edu
Computing & Information Systems Department
Saint Vincent College
Latrobe, PA 15650 USA

Abstract

Whether you are just starting to teach cryptography, or you teach it as a stand-alone course for computer science majors or as part of a complete major in cybersecurity, the question of how to provide hands-on experience is an important one. Some software may be too expensive, while other schemes only allow students to use small, toy examples. Here, a solution using a software package called bigint is examined. It can allow students to implement, try out, and try to break C++ implementations of most common public key cryptographic algorithms. Better yet, bigint is free and will run under Linux, which is often free as well. Thus, with this free software, students can implement common cryptographic algorithms, using large numbers instead of tiny ones, time how long the computations take, and investigate where the algorithms fail to work well – the sort of exercises that help students more fully understand this technical and rapidly-changing field.

Keywords: Cryptography, Teaching, Cybersecurity Education, Public Key, Software.

Recommended Citation: Carlson, D.E., (2024). Teaching Public Key Cryptography: A Software Approach *Cybersecurity Pedagogy and Practice Journal*, v3 (n2) pp 12-20. <https://doi.org/10.62273/JJIP7451>

Teaching Public Key Cryptography: A Software Approach

David Carlson

1. INTRODUCTION

The purpose of this article is not to teach the reader public key encryption and decryption algorithms. This is done in many good textbooks (Elbirt, 2009; Katz & Lindell, 2020; Stallings, 2020; Stinson & Paterson, 2018; Trappe & Washington, 2006). Rather, the purpose is to show that the bigint software (McCutchen, 2010) can be used to produce programs that carry out these cryptographic algorithms well. Note that Mathematica would be a good alternative, but bigint is free.

Trying these algorithms allows students to deepen their understanding beyond what is possible by simply reading about them in a textbook. Also, because the software is free and runs on common Linux systems, this part of a cryptography course can be done cheaply and relatively easily. Understanding the algorithms is a different matter, not addressed in this paper, and requires some background in mathematics, especially abstract algebra and number theory. Many of the available textbooks present the needed background in these areas.

Although bigint is useful for trying out cryptography, it does not do everything a good course should do. For example, it would be useful to show students where a browser displays the digital certificate used on a website. It might also be good to show where digital certificates are installed on a web server, if such a server is accessible to the class.

Cryptography is heavily used. For example, every time someone uses https to connect to a website, cryptography is used to encrypt the data passed between the user and the website. Essentially, cryptography keeps private data safe from prying eyes. Even if the encrypted data is captured and examined, without the cryptographic key, the data cannot be decrypted.

There are two main types of cryptography: public key cryptography and private key cryptography. Both allow data to be encrypted and later decrypted, perhaps after the encrypted data has been sent to some recipient. Public key

cryptography, the type being discussed here, uses a public encryption key and a private decryption key. The fact that only someone having the private decryption key can decrypt the message is why this method is so useful.

Unfortunately, public key cryptography is slower than private key cryptography, where both sender and receiver use the same private key to encrypt and decrypt. Generally, public key cryptography is used to securely distribute a shared key for private key cryptography. Once this shared key has been shared with the recipient, messages encoded with the shared key can be safely sent to the recipient and decrypted. Other parties cannot decrypt the messages since they don't have the shared key.

This article is about helping students to learn public key cryptography by trying it out using the bigint software package. This can be done in two distinct ways: Students with good programming backgrounds can write some of the code for doing this type of cryptography. Students who need to know some of the characteristics of public key cryptography but who do not have sufficient programming backgrounds can run prewritten programs that demonstrate some of the features of this type of cryptography.

Bigint is a free package that allows programs to be written that work with arbitrarily large integers. Much of public key cryptography uses extremely large integers, the kind that do not fit into an ordinary integer variable. Thus, bigint is an excellent package for trying out the algorithms of public key cryptography. Bigint's creator, Matt McCutchen, has put it in the public domain. Of course, the larger the integers are, the slower any computations will run. Still, it is often possible to work with a hundred or more decimal digits in a reasonable amount of time. Although bigint is no longer maintained, it works well for the example algorithms discussed here. The author of this paper has successfully used bigint on a Linux server for over a decade as a significant part of a cryptography course taught to cybersecurity and computer science majors.

2. SHORT BIGINT EXAMPLES

The bigint package uses .cc files containing C++ code and filenames with a .hh extension for its header files. It provides a test program that tries out and illustrates the functionality that is provided for these big integers. Most of the constructs are familiar items from C++ and C. What is new is using these items on arbitrarily large integers as well as some of the operations that can be done on these integers.

There are two data types, `BigInteger` and `BigUnsigned`, for large integer variables. You can simply assign an ordinary `int` into a variable having one of these new types. You can also calculate a huge value having one of these two types and assign it into a `bigint` variable. To copy a value from a `BigInteger` to an ordinary `int` variable, a `toInt()` conversion function must be used. An exception is thrown if the value is too large to fit into an ordinary `int`. You can also convert a string of digits into a `BigInteger` by using the `stringToBigInteger` conversion function. Here are some examples of putting a number into a `BigInteger` variable:

```
BigInteger a; // a contains 0 by default
int b = 2047;
a = b; // Converts int into a BigInteger
b = a.toInt(); // Converts BigInteger to
// ordinary int (if it will fit)
BigInteger c(a); // Copy a to BigInteger c
// Put int literal into BigInteger d:
BigInteger d(-314159265);
```

There is no `BigInteger` literal, but you can convert a string of digits to a `BigInteger` as follows:

```
string s("3141592653589793238462643383");
BigInteger e = stringToBigInteger(s);
cout << e << endl;
```

You can also work in hexadecimal, which is sometimes convenient. Note how this example switches to hexadecimal, does some bitwise Boolean operations and shifts, and then returns to decimal for future output:

```
BigUnsigned i(0xFF0000FF), j(0x0000FFFF);
cout.flags(ios::hex | ios::showbase);
// The << operator uses these flags.
cout << (i & j) << // Boolean AND is &
    << (i | j) << // Boolean OR is |
    << (i ^ j) << // Exclusive OR is ^
    // Shift distances are unsigned ints
    << (j << 5) << // Shift left 5 bits
    << (i >> 3) << // Shift right 3 bits
```

```
cout.flags(ios::dec); // Now to decimal
```

In the last lines of the above example, note that `j = 0000FFFF = 1111111111111111` and that we then shift this number left by 5 bits.

You may have noticed that we are seeing some of the math that is needed to do public key cryptography. In particular, we need operations on huge integers, operations such as powers, greatest common divisor (`gcd`), modular exponentiation (`modexp`), and modular inverse (`modinv`). `Bigint` has all of these.

We say that numbers `a` and `b` are congruent mod `m` if their difference is divisible by `m`. Thus `a ≡ b (mod m)` means that `a - b = km`, for some integer `k`. Congruence mod `m` is heavily used in public key cryptography.

Now try an example where we calculate powers of 274 by repeated multiplication:

```
int max = 10;
BigUnsigned x(1), big274(274);
for (int power = 0; power <= max; power++)
{
    cout << "274^" << power << " = " << x
        << endl;
    x *= big274; // A bigint assignment
}
```

In public key cryptography, it is very common to calculate a number to a power, but done mod some third integer. That's called modular exponentiation. `Bigint` has a built-in `modexp` function for this. Here is some `BigInteger` code to calculate a greatest common divisor, a modular inverse, and a modular exponentiation, though the numbers could be much larger than what is used here:

```
cout << gcd(BigUnsigned(60), 72) << '\n'
    << modinv(BigUnsigned(7), 11) << '\n'
    << modexp(BigUnsigned(314), 159, 2653)
    << endl;
```

Notice that `modinv(y, n)` finds a value, which when multiplied by `y`, produces 1 (mod `n`). Thus, the example with 7 and 11 should produce 8 as the inverse, since `7 * 8 = 56` and `56 ≡ 1 (mod 11)`, where `≡` indicates congruence. If we convert that congruence to an equation, it would say that `56 = 1 + 11k` for some integer `k`. (In fact, it is clear that `k = 5` makes this statement to be true.) Note that `modexp(r, s, n)` finds `r` to the `s` power, with the result reduced mod `n`.

3. USING BIGINT

Installing `bigint` on a Linux system involves copying folders of files to each user's directory, including files whose names fit the patterns `BigInteger*` or `BigUnsigned*`, as well as ones named `sample.cc`, `Makefile`, and a few others. `Makefile` uses your Linux installation's `g++` compiler for C++. (If you don't have C++ installed, you need to install it for `bigint` to work.)

The simplest way to start is to edit the `sample.cc` file, which contains a test program, and replace the test code with your own code. Then enter `make` at the command line to compile your `sample.cc` program. If there are error messages, edit `sample.cc` to make corrections. Otherwise, run your program by entering `./filename` but with `filename` replaced by the name of your compiled program. By default this will be the name of your `.cc` file with the `.cc` omitted.

You can also name your C++ file something other than `sample.cc` as long as you make two small changes to `Makefile`. In that file, find the comment "Components of the program". On the next two lines, change `sample` and `sample.o` to use the actual name of your C++ file. For example, you might use `program4` and `program4.o` instead of the original names. Once your program compiles, you can run it by entering `./program4` at the command line. Use the actual name of your compiled program, of course. When you are ready to create another `bigint` program, just make a copy of the entire folder containing the current program, change the code in the new folder, run `make`, and if `make` is successful, use `./filename` to run your program.

4. BASIC NUMBER THEORY FUNCTIONS

Let's consider again these three functions: `gcd`, `modinv`, and `modexp`. Public key cryptography uses them extensively. This gives us hope that we can implement cryptographic routines such as those in RSA and elliptic curve cryptography. However, we need a few other key items, such as the ability to generate large primes. Here are a few of the routines. They use the C++ `rand()` function which generates a random integer between 0 and the constant `RAND_MAX`. The reader can refer to cryptography textbooks for explanations of these and other related functions. The purpose of including these functions here is not to teach public key cryptography but simply to show that we have all of the machinery needed to do public key cryptography. Students can then run cryptography exercises on a computer and

not simply read the descriptions of public key cryptography in a text. Hands-on work is quite possible! Let's try some here.

First, we have some functions that produce random digits:

```
char RandomDigit(void)
{
    int digit;
    digit = (9.999 * rand()) / RAND_MAX;
    return digit + '0';
}

char RandomNonzeroDigit(void)
{
    int digit;
    digit = (8.999 * rand()) / RAND_MAX + 1;
    return digit + '0';
}

char RandomOddDigit(void)
{
    int digit;
    digit = (4.999 * rand()) / RAND_MAX;
    return 2 * digit + 1 + '0';
}
```

Next is a function that generates a positive integer:

```
void GeneratePosIntPlain(BigUnsigned &
PosInt, int NumDigits)
{
    int k;
    string s;

    s = RandomNonzeroDigit();
    for (k = 1; k < NumDigits; k++)
        s = s + RandomDigit();
    PosInt = stringToBigUnsigned(s);
}
```

Then we have a function to generate a random string of digits. The string that is produced has a set number of digits.

```
string GenerateString(int NumDigits)
{
    int k;
    char nonzero[2];
    nonzero[0] = RandomNonzeroDigit();
    nonzero[1] = '\0';
    string s(nonzero);
    for (k = 2; k < NumDigits; k++)
        s = s + RandomDigit();
    return s + RandomOddDigit();
}
```

The strategy for getting a prime number of a

desired length is the same one that is used elsewhere in public key cryptography: generate strings of digits and then use appropriate primality tests as many times as needed to find a string that represents a prime number with probability as close to 1 (that is, 100%) as desired. The Fermat primality test (Trappe & Wahington, 2006) is shown here:

```
// Global constants for speed:
const BigUnsigned BigOne = BigUnsigned(1);
const BigUnsigned BigTwo = BigUnsigned(2);
const BigUnsigned BigThree =
    BigUnsigned(3);
const BigUnsigned BigFive =
    BigUnsigned(5);
const BigUnsigned BigSeven =
    BigUnsigned(7);
```

The following is a function to generate a prime number with a set number of digits:

```
// Assumes srand(time(NULL)) was done to
// seed the random number generator.
BigUnsigned GeneratePrime(int NumDigits)
{
    BigUnsigned candidate;
    string s = GenerateString(NumDigits);
    candidate = stringToBigUnsigned(s);
    // Fermat primality tests:
    while(! PassPrimalityTests(candidate))
    {
        string s=GenerateString(NumDigits);
        candidate = stringToBigUnsigned(s);
    }
    return candidate;
}
```

Next is a function to run tests to see if integer m is probably prime:

```
bool PassPrimalityTests(const BigUnsigned
& m)
{
    if (! FermatPrimalityTest(BigTwo, m))
    // Uses global constant defined above.
        return false;
    if (! FermatPrimalityTest(BigThree, m))
        return false;
    if (! FermatPrimalityTest(BigFive, m))
        return false;
    if (! FermatPrimalityTest(BigSeven, m))
        return false;
    return true; // Hope it really is prime!
}
```

The previous function uses the following function that performs the Fermat primality test.

```
// Base a to see if m is probably prime.
bool FermatPrimalityTest(const
    BigUnsigned & a, const BigUnsigned & m)
```

```
{
    if (modexp(a, m - 1, m) == BigOne)
    // Check if a^(m-1) gives 1, mod m.
        return true;
    else
        return false;
}
```

Miller-Rabin primality testing can be implemented (Trappe & Washington, 2006) in a similar way. By applying this test for n values of variable a, we can guarantee that the probability that a number is not prime when it passes n of the checks described in the Miller-Rabin algorithm is less than $(1/4)^n$. Thus, we can make this probability arbitrarily small.

5. PUBLIC KEY CRYPTOGRAPHY

We can now try out cryptographic algorithms such as RSA, elliptic curve encryption and decryption, etc. Here is a typical way in which a textbook might describe RSA encryption and decryption:

Bob's computer picks large secret primes p and q.

It calculates $\phi = (p - 1) * (q - 1)$.

The computer calculates $n = p * q$.

It further chooses an integer encryption exponent e so that $\text{gcd}(e, \phi) = 1$.

It goes on to find an integer decryption exponent d with $d * e \equiv 1 \pmod{\phi}$.

The sender's computer makes n and e public, but keeps p, q, phi, and d private.

Alice encrypts a numeric message m as $c = m^e \pmod{n}$ and sends it, c, to Bob.

Bob decrypts c by using $m = c^d \pmod{n}$.

It would seem to be difficult to translate the above into code that would run on a computer, but with bigint, it is fairly easy. That's why it is useful for examples in the classroom and for student homework.

Here is what we saw above in outline, now written using C++ and bigint code that students can run. In this example, the one program is both the sender and receiver, but the two pieces of functionality could be split out and given to a pair of students, one of whom sends the encrypted message to the other, who then decrypts it. Some minor details have been omitted. The initialization stage looks like this:


```
int PrimeLength;
// Seed the random number generator:
srand(time(NULL));
BigUnsigned m, c, e, d, decrypted;
cout << "Enter number of decimal digits"
      << "for primes p and q (e.g. 120): ";
cin >> PrimeLength;
BigUnsigned p =
    GeneratePrime(PrimeLength);
BigUnsigned q =
    GeneratePrime(PrimeLength);
BigUnsigned n = p * q;
BigUnsigned phi = (p - 1) * (q - 1);
```

Next, the program generates a 4-digit prime and the encryption exponent e and decryption exponent d :

```
e = GeneratePrime(4);

// We need to have gcd(e, phi) = 1.
// Try until you we get one that works.
while (gcd(e, phi) != 1)
    e = GeneratePrime(4);

d = modinv(e, phi);
```

The next step is to ask the user (Alice) for a short message to be encrypted. It is written as a number.

```
string data;
cout << "Enter a message string: " << endl;
// For example, 010203 to represent ABC
cin >> data;
m = stringToBigUnsigned(data);
```

Now we encrypt the message m to produce the ciphertext c and then decrypt it so that Bob can read it:

```
c = modexp(m, e, n);
decrypted = modexp(c, d, n); // Decrypt c.
cout << "Decrypted message m = " <<
      << decrypted << endl << endl;
```

Finally, we check to see if the decrypted message matches the original plaintext message and report on the results:

```
cout << "original m & decrypted message "
if (m == decrypted)
    << cout << "match" << endl;
else
    cout << "do NOT match" << endl;
```

Similarly, bigint can be used to write programs that do elliptic curve cryptography, ElGamal encryption and decryption, etc.

6. CRYPTOGRAPHIC EXERCISES

Many types of cryptographic exercises can be tried with bigint. For students who have a good background in programming and math, it is quite possible to have them write programs in bigint that encrypt and decrypt messages using RSA, ElGamal, and elliptic curve cryptography, much like the brief RSA example just presented here. Textbooks are available that provide the details of the algorithms and the math on which they are based. Students would have to translate those algorithms into bigint code. They might also write bigint programs to try to break some encrypted message and perhaps time how long it takes. It would be useful to have as a parameter the minimum number of bits the primes have in them. That way, the difficulty of cracking the code can be easily varied. If the number of bits is too large, there is little chance of extracting the message, while too small of a number of bits would result in a very fast calculation of the plaintext message.

In some cryptography courses, the students might not have the math or programming background to do the type of exercises just discussed. There are other types of exercises that would be more appropriate. Some of these are timing exercises where you find out how long it takes to do something. For example, students could use a program that finds how long it takes to encrypt a message (supplied as input to the program). Then the same type of timing exercise could be used for the process of decrypting the ciphertext.

To be more specific, consider a timing exercise that begins by calculating $n = p * q$ as in RSA. By printing a message just before and after this calculation, the program can show that this multiplication happens quickly. (An alternate approach is to have the program access the system time right before and after the calculation. Then the difference of those two times gives the approximate time for that calculation.

In contrast, if n is large enough, the factoring of n as p times q can take much longer, with a very noticeable or even prohibitive delay. This factoring could be done by trying successive values in a large table of primes to see if any of them divide n . The difference of two squares method could also be tried. In the latter, you need a large table of squares. You then add a square to n to see if you get a square. If that

works, you have $n + b^2 = a^2$, so that $n = a^2 - b^2 = (a + b)(a - b)$. In both methods, the table might be too small to factor n . However, even if the table is large enough, the checking of many table values may mean that the factoring of n takes a large amount of time. If RSA is implemented well and n is large enough, none of these factoring methods will succeed unless you can devote a large amount of time (perhaps years) to the factoring attempts. Students can try RSA with moderate values of n to see how long it takes to break the scheme by factoring n . They might also graph how this time increases as they slowly raise the number of digits in n .

Other exercises might have students encrypt a message, decrypt it, and verify that the original message is obtained. Still others might report how many collisions some cryptographic hash function (or even a non-cryptographic hash function) produces with a certain set of data. (A collision is when two data values hash to the same result.) As a general rule, the fewer collisions there are, the better the hash function is. Another type of exercise is to see how many hash function values you have to check to get a collision. Collisions are inevitable if you hash enough values, but they should be rare.

One strategy is to use a test program based on the birthday paradox. It has been shown that if the values fit the range $[0, n-1]$, and some other technical conditions are met, having about the square root of n as the number of data items to be hashed is enough to have better than a 50% chance of a collision. A variation is to make two tables of hash values and look for a value in one table that also occurs in the other. A related exercise is to use the Baby Step, Giant Step algorithm (Trappe & Washington, 2006) that uses two lists to try to break ElGamal encryption, which is based on discrete logs. If students lack the math and programming background to create the software for this, they could simply be given the software and asked to see if they can use the software to get one or more collisions.

We next look at some of the specifics of the testing of a hash function using the two list approach. Below is a partial listing of a program that looks for collisions when testing a simple non-cryptographic hash function. Unless you have really strong students, you would probably want to give students the program and simply ask them to use it to find hash function collisions.

This program makes two lists of hash values for

somewhat random inputs and looks for matches between the lists. The details of the matches (such as what input hashed to what value) and the total number of matches are printed. Both lists are 45000 items long. The hash values are 30 bits long. Using the analysis by Trappe and Washington (Trappe & Washington, 2006), $N = 2^{30} = 1073741824$ possible hash values. Then $\text{sqrt}(N) = 2^{15} = 32768$. Since each list is almost 50% longer than $\text{sqrt}(N)$, namely length 45000, we expect a very good chance of a collision, a match. Note that $\lambda = (45000^2) / (2^{30}) = 1.886$ roughly. Then the chance of a match is approximately $1 - e^{-\lambda} = 0.848 = 84.8\%$. Thus, a match is quite likely.

We begin with some initialization so that we can create the first table of hash values.

```
NumDigits = 61;
GeneratePosIntPlain(FirstRunStart,
    NumDigits);
cout << "FirstRunStart is " <<
    FirstRunStart << endl;

cout << "Creating a table of hash values"
    << endl << endl;
BigUnsigned Largest =
    stringToBigUnsigned("100truncated");
```

That last line should use 10^{61} , which is 1 followed by 61 zeros. It can't fit here, but a string containing a 1 followed by 61 zeros will work in the code.

Next, we produce the table of hash values, all of which have 61 bits:

```
for (k = 0; k < MAX; k++)
{
    Increment = modexp(Two, k, Largest);
    k3 = FirstRunStart + Increment;
    k3 = k3 % Largest;
    table[k] = hash(k3);
}
```

The next step is to set up to create the second table of hash values. These will all have 63 bits.

```
NumDigits = 63;
GeneratePosIntPlain(SecondRunStart,
    NumDigits);
num = SecondRunStart;
cout << "SecondRunStart is " <<
    SecondRunStart << endl << endl;
```

The last section generates the new hash values and checks if any of them match a hash value in the first table.

```

for (m = 0; m < MAX; m++)
{
    GeneratePosIntPlain(num, NumDigits);
    value = hash(num);
    // Sequential search table for value
    index = SequentialSearch(value);
    if (index >= 0)
    {
        matches++;
        bigindex = BigUnsigned(index);
        index3 = FirstRunStart +
            modexp(Two, bigindex, Largest);
        index3 = index3 % Largest;
        cout << "Matching values found at "
        << index3 << " and " << num << endl;
        cout << "hash of first one: " <<
            hash(index3) << endl;
        cout << "hash of second one: " <<
            hash(num) << endl << endl;
    }
    num++;
}
cout << "Number matches: " << matches
    << endl << endl;

```

For those who teach ElGamal encryption, a useful exercise is to have students write, or simply use, a small program that prints the powers of the numbers 1 through n-1 for some positive number n, where the powers are reduced modulo n (Stallings, 2020). It makes it easy to see which of the numbers 1 through n-1 is a primitive root of n, since in such a row all of the computed values are distinct. Note that a primitive root might be better named a "multiplicative generator". Students could have their program print an asterisk after each row in which the starting number is a primitive root of n. The following is the output of this program when 11 is chosen as the prime. It flags 2, 6, 7, and 8 as the primitive roots mod 11.

Enter a prime number less than 44: 11
Table of powers of a modulo 11 where the powers for a are on the next line:

```

1 2 3 4 5 6 7 8 9 10
1 1 1 1 1 1 1 1 1 1
2 4 8 5 10 9 7 3 6 1*
3 9 5 4 1 3 9 5 4 1
4 5 9 3 1 4 5 9 3 1
5 3 4 9 1 5 3 4 9 1
6 3 7 9 10 5 8 4 2 1*
7 5 2 3 10 4 6 9 8 1*
8 9 6 4 10 3 2 5 7 1*
9 4 3 5 1 9 4 3 5 1
10 1 10 1 10 1 10 1 10 1

```

This table can also be read backwards to find discrete logarithms. For example, if we look in the row starting with 2 for the item in column 4, we get a 5. That indicates that 2^4 gives 5 when working modulo 11. That's correct since $2^4 = 16$, which is congruent to 5 when we are working mod 11. We can also read this calculation backwards to say that the discrete log of 5 (when using base 2 and modulus 11) is 4.

7. CONCLUSIONS

This article has made the case that the bigint software package is a very useful tool in the teaching of cryptography. It can easily deal with integers having one hundred or so decimal digits, which is much more realistic than using ordinary numbers of type int. It is also free, which is a big help when budgets are tight. Of course, other types of exercises are likely to be needed in addition to the ones that use bigint.

The presentation here does not try to cover all areas of public key cryptography. Rather, it simply shows some representative examples. Professors who teach cryptography are welcome to contact the author for bigint examples and homeworks. These would be best placed on a Linux system where students could copy these items to their personal folders. Some of the examples that have students run a program several times and record what was produced can be done in less than an hour; ones that involve coding might take several hours. Students almost always succeed at that first type of problem, but some find those involving coding to be challenging. Still, the majority of students succeed on this. Since the author's class is typically small, any statistics on this beyond this general trend are probably meaningless.

It should be noted that public key cryptography is expected to be phased out around 2030. By then, it is likely that quantum computers will be approaching the power needed to break a lot of public key cryptography. The U.S. National Institute of Standards and Technology (NIST) has already published its first group of algorithms for quantum-resistant cryptography (NIST, n.d.; NIST, 2023). Many are based on the closest vector problem or the shortest vector problem in a high-dimensional space. That would be a quite different approach to cryptography. Public key cryptography is expected to still be taught but will likely be done more for historical interest. The main emphasis will instead be on private key methods (thought to be quantum-resistant) and new types of quantum-resistant cryptography.

8. REFERENCES

- Elbirt, A. (2009). *Understanding and Applying Cryptography and Data Security*. Taylor & Francis Group, LLC.
- Katz, J., & Lindell, Y. (2020). *Introduction to Modern Cryptography*, 3rd ed. CRC Press.
- McCutchen M. (2010), C++ Big Integer Library. Retrieved August 21, 2023 from <https://mattmccutchen.net/bigint/>.
- National Institute of Standards and Technology (n.d.). National Cybersecurity Center of Excellence. Migration to Post-Quantum Cryptography. Retrieved August 21, 2023 from <https://www.nccoe.nist.gov/crypto-agility-considerations-migrating-post-quantum-cryptographic-algorithms>.
- National Institute of Standards and Technology. Computer Security Resource Center. Selected Algorithms 2022. Last modified August 14, 2023. Retrieved August 21, 2023 from <https://csrc.nist.gov/Projects/post-quantum-cryptography/selected-algorithms-2022>.
- Stallings, W. (2020). *Cryptography and Network Security: Principles and Practice*, 8th ed. Pearson Education, USA.
- Stinson, D., & Paterson, M. (2018). *Cryptography: Theory and Practice*, 4th ed. CRC Press.
- Trappe, W., & Washington, L. (2006). *Introduction to Cryptography with Coding Theory*, 2nd ed. Pearson Prentice Hall.

Teaching Case

Digital Forensics and Incident Response (DFIR): A Teaching Exercise

Jennifer L. Breese
jzb545@psu.edu
Penn State Greater Allegheny
Penn State University
McKeesport, PA, U.S.

Maryam Roshanaei
mur45@psu.edu
Penn State Abington College
Penn State University
Abington, PA, U.S.

J. Andrew Landmesser
jal620@psu.edu
Penn State Brandywine
Penn State University
Media, PA, U.S.

Brian Gardner
bkg113@psu.edu
Penn State Schuylkill
Penn State University
Schuylkill Haven, PA, U.S.

Abstract

Cybersecurity requires practical knowledge related to protecting electronic information systems and, more importantly, hands-on skill sets for students. To prepare cybersecurity students for effective workforce contributions, experiential practice in a modern, secure environment is essential. An ideal and cost-effective way to provide this environment for both institutions with funding limitations and students with starved resources is to establish a live virtual isolated lab environment that acts as a sandbox for performing cybersecurity-related exercises, including ethical hacking, penetration testing, offensive and defensive security, information risk assessment and management, and malware analysis. This teaching exercise provides suggestions and resources, including free training by reputable cybersecurity companies offering services to the broader industry community, as excellent options to include in student coursework. Additionally, this teaching exercise offers three lessons and a full learning module to include in a variety of introductory cyberforensics, information systems, and other related disciplines to both provide hands-on learning and engage students pursuing a major in cyber studies.

Keywords: cybersecurity, digital forensic incident response, cyber education, LinkedIn Learning

Recommended Citation: Breese, J.L., Landmesser, J.A., Roshanaei, M., Gardner, B., (2024). Digital Forensics and Incident Response (DFIR): A Teaching Exercise. *Cybersecurity Pedagogy and Practice Journal*; v3 (n2) pp 21-34. <https://doi.org/10.62273/EPXU4458>

Digital Forensics and Incident Response (DFIR): A Teaching Exercise

Jennifer L. Breese, Maryam Roshanaei, J. Andrew Landmesser and Brian Gardner

1. INTRODUCTION

Rapid technological advancements have led the entire world to shift towards the digital realm, which increased exponentially during the COVID-19 pandemic. The transition has resulted in the emergence of cybercrimes and security breach incidents that threaten the privacy and security of users and organizations overall. Alghamdi (2020) examined the use of digital forensics in countering cybercrimes, which has been a critical development in cybersecurity. Security vulnerabilities and breaches have galvanized the developments in digital forensics, requiring data extraction from digital devices to be used as evidence in both criminal and civil legal proceedings. To understand the importance of digital forensics, Lallie et al. (2021) thoroughly discusses current trends, potential threats, and opportunities of digital forensics in cybersecurity with a focus on the impact of COVID-19 changes to the overall security landscape. Research has also identified specific threats to digital forensics, which include technical, operational, and personnel-related challenges (Easttom et al., 2022; Lallie et al., 2021; Srinivas & Kumar, 2019; Whitman & Mattord, 2021). Both statistics and analytics have shown that the exponential growth of cyber threats and attacks necessitate a corresponding need for forensic experts and forensic researchers for automation procedures in the digital realm (Joseph & Norman, 2018). According to Cyberseek (2021), cybersecurity workforce preparation is an integral part of closing the skill gap required for combating threats found in mobile apps, networks, and phishing in mobile applications.

Cybersecurity expertise requires a solid understanding of policies related to protecting electronic information systems, but again more hands-on skill sets for students is an important focus. Job descriptions related to cybersecurity, even at the entry-level, state a need for three to seven years of prior cybersecurity-related experience. To prepare cybersecurity students, hands-on practice in a modern, secure environment is essential, and the best way to provide this is to establish a dedicated physical and live virtual lab environment. The physical lab is typically a dedicated room/classroom that

houses its own servers and dedicated workstations to learn physical/logical networking and perform digital/computer forensics. Normally this setup can be too costly, even for institutions with larger resources to deploy. Again, the isolated lab environment can provide a safe, fully virtualized, and sandboxed platform to perform ethical hacking, penetration testing, offensive & defensive security, information risk assessment & management, malware analysis, etc. While a physical lab is cost prohibitive, live virtual lab environments that include free training from reputable cybersecurity companies are great options to include in student coursework. Although some of these free options do not have broad certification recognition in the employment market, they could still be a differentiating factor when included in a student's resume submission and on their LinkedIn profile.

Providing learning opportunities and garnering interest in lower-level courses are key to developing a robust student pipeline in Cybersecurity Analytics and Operations (CYAOP). We are utilizing materials provided by NIST (National Institute of Standards and Technology) and other free online courses offered by industry providers to develop foundational modules for learning, interest, and excitement among our student population, potentially even drawing additional students to Information Sciences and Technology (IST) and CYAOP majors.

What is Cyberforensics?

Marcella (2021) describes cyberforensics as the discipline focused on identification, preservation, examination, and analysis of digital evidence using scientifically accepted and validated processes. Digital evidence can come from many diverse sources, including personal computing devices, networking devices, servers, cloud computing environments, and Internet of Things (IoT) devices such as vehicles and surveillance cameras with most personal computing and IoT devices networked to cloud resources. According to NIST (2022), hundreds, if not thousands, of individual digital forensic techniques might need to be used in a complete digital forensic examination. NIST identifies several useful models, each with a different emphasis, for digital forensic examinations. Further, these digital

forensic standards differ depending on the scene, nature and type of evidence being handled. For the successful prosecution and admissibility in court, certain accepted procedures must be properly followed. Digital forensic examiners use different methods and tools to accomplish the same job during digital investigations and with the changing world of digital technology these tools and methods are variable to change (Mabuto & Ventor, 2011). Beardall (2023) navigated the symbiotic relationship between cybersecurity and digital forensics, exploring the key role of digital forensic methodologies play in addressing cyber incidents while also recognizing the issues with the lack of investigative standardization.

2. PROJECT PURPOSE

The goal of this teaching exercise is to develop an instructional learning tool in cyberforensics to enhance the learning experience of students pursuing a degree in Cybersecurity Analytics and Operations (CYAOP), Information Sciences and Technology (IST), or an equivalent degree. The instructional learning module provides materials on many of the latest digital forensics frameworks and their related subjects. This teaching exercise also provides skills that students need to fully comprehend the security strengths and weaknesses of digital forensics, including mobile devices, platforms (e.g., Apple iOS and Android), and their functionalities. The instructional learning module and tool includes a step-by-step, hands-on approach that uses many current industry tools and techniques to help gain a basic understanding about Digital Forensic Investigation Response (DFIR) with additional elements to demonstrate mastery.

The development of this exercise has been a trial-and-error process through a collective of four campus faculty from the overall seven campus cyber consortium of the Penn State college campus community. Additional steps were added to the exercise to create a coordinated set of modules that can be used across the overall curriculum. These additions have served to connect other course knowledge progressions rather than a one-day or a one-off session for student learning.

3. BACKGROUND

NIST (2022) provides best practices in digital investigation techniques based on computer science methods from peer-reviewed sources, academic and classroom materials, and technical guidance from professional organizations. NIST

describes options for acquiring and analyzing data from a mobile device that are explored in this case leveraging the Paraben Electronic Evidence Examiner (E3) platform. NIST (2022) also specifies a seven-step process with the first three focused on data collection and the last four on data interpretation.

1. Step one protects collected evidence from modification typically using write blocking.
2. Step two performs data acquisition as an image copy of the original data.
3. Step three in data collection ensures the integrity of the acquired data typically using cryptographic hashing techniques.
4. The fourth step (also the initial one in data interpretation) attempts to recover any deleted data. NIST discusses three commonly used techniques for data recovery of metadata-based file recovery, file carving, and deleted record recovery.
5. The next step performs navigation through the acquired data typically supported by validated forensics software tools.
6. The sixth step identifies and extracts data relevant to the investigation using criteria of interest like specific text of date-time intervals.
7. The last step in data interpretation analyzes all the extracted data artifacts to develop a narrative and timeline of events for inclusion in a final report.

What are mobile forensics and why is it important?

Mobile forensics is a subset of cyberforensics focused on analyzing digital evidence from mobile devices in a forensically sound manner. Since mobile devices are networked with other digital devices, evidence from mobile devices can often provide clues to additional digital resources to investigate. Most mobile users conduct email and social media interactions from their mobile devices. With social media applications encouraging users to share personal data, mobile users often leave significant personal data on their mobile devices without being aware (Casey, 2011). Investigators must have the skills needed to overcome challenges including potential remote device wiping, encryption, and physical imaging of various file systems. While there are many aspects of DFIR to consider in the ever-changing DFIR IoT landscape as discussed by Beardall (2023) this exercise focuses on mobile forensic discovery and analysis.

Pennsylvania State University (PSU) Commonwealth Campuses of Abington, Brandywine, and Greater Allegheny received 2022 software grants for Paraben Electronic

Evidence Examiner (E3), an all-in-one platform for adding forensic data through a collective interface for analysis and improving student understanding of mobile device forensics via hands-on lab exercises. The Paraben Unified Police Support (PUPS) grant helps resource constrained organizations, including educational institutions, to enhance the ability to process digital evidence with software and training costs for the E3 Fundamental Fast Track and Mobile Fast Track for one year including access to the Paraben Online Training Academy and access to all the courses and labs (see <https://paraben.com/dfir-le-grant/> for details on grant application). Once Paraben E3 licenses are granted, students download Paraben E3 platform installer from <https://paraben.com/paraben-downloads/> to install on individual computers. If opening E3 without Admin privilege, E3 prompts for Admin login but you can click No to continue in E3 without Admin privileges; however, some types of evidence will be unavailable. Once the Activation wizard opens, select the Internet License option and click Activate. Once the Connect to Web License Server dialog displays, enter the generic student user login and password supplied by Paraben then click Connect. E3 opens a dialog with four options of 1) Acquire Device, 2) Import Data, 3) Add Evidence, or 4) Open Case. If you close the dialog, you can still perform these functions from E3 menu options.

Paraben also provides a free online version of their DFIR tool for download with limited capabilities; educational tutorials on the site are also useful (see <https://paraben.com/free-dfir-tools/> for the free download) (*free download*, n.d.). The free download is available after a trial is completed. Paraben also has a YouTube channel that provides educational videos on the use of the tool at <https://www.youtube.com/user/ParabenForensics> (YouTube, n.d.).

Our previous hands-on labs in our undergraduate cyberforensics courses of Information Sciences and Technology (IST) 453 Legal, Regulatory, Policy Environment of Cyber Forensics, and IST 454 Computer and Cyber Forensics utilized only free forensics tools primarily available in Kali Linux distributions. Kali Linux provides Autopsy as the GUI tool for Sleuth Kit under the Forensics menu option, carving tools like *scalpel*, and *guymager* under Forensic Imaging Tools to support multiple image file formats. With the importance of mobile forensics in current investigations, this software grant enabled our students to gain experience specifically with Android and iPhone devices. Paraben also issued

lab instructor and student lab manuals for Android and iPhone devices. However, the student manuals assumed a greater level of Paraben tool-specific training than our students have the experience to walk through self-guided, so we added more specific detailed Android lab steps for our undergraduate students. These steps are detailed further in section 5 titled 'Teaching Exercise Learning Steps' under Additional Steps/Suggestion(s). Also, the Paraben Android and iPhone student labs used an existing Paraben evidence file starting at step four in the NIST seven-step digital forensics process. Students need to understand in the advanced levels how to extract the data file which is not provided by Paraben as the data file is furnished. Fortunately, we provided an earlier course lab focused on the first three NIST steps that included manually performing Linux filesystem acquisitions using the *dcfldd* command with required options. *dcfldd* is an enhanced version of disk dump command that includes features useful for forensics and security including hashing input data during transfer to ensure data integrity and logging of output.

This case was developed for a module initially for IST 453 Legal, Regulatory, Policy Environment of Cyber Forensics and has been adopted in other courses, specifically Security Risk and Analysis (SRA) 221 Overview of Information Security, and IST 454 Computer and Cyber Forensics. This teaching exercise initially provides five steps to developing a module on student learning and includes a free three-hour DFIR (Digital Forensic Incident Response) training and certificate through Cyber Triage (*Online incident response training with Brian Carrier, 2022*). The Cyber Triage certificate titled *Intro to DFIR: Divide and Conquer* may not be widely recognized in industry but attempts to be tool agnostic focusing on breaking down the large investigative questions into smaller questions. Smaller questions can be answered through the artifacts uncovered in an investigation.

The ability to add sections to the module exists; however, skipping steps is not recommended or advised. Again, additional steps and labs have been adopted in various courses for the development of advanced student knowledge.

1. Textbook background chapters for reading and comprehension are assigned, and open access articles can be substituted as a student cost-effective alternative.
2. Students are required to complete the free DFIR training through Cyber Triage. The training, which takes approximately three hours, instills the ability to frame an

investigation before the students complete the exercise. Students are asked to submit the completed certificate issued through the industry provider through the Learning Management System Dropbox for credit. Further, students can and should add the DFIR training to their LinkedIn to differentiate them from peers in similar fields when they enter industry.

4. STUDENT/PROGRAM BENEFITS

The material in parts or in the additional exercises seeks to facilitate the following for students' success:

1. Provide students essential and valuable skills along with the opportunity to explore offensive security and ethical hacking methodologies.
2. Ensure that the projects will prepare students for the workforce and allow students to work from their own location with minimal computer requirements.
3. Give students in-depth theory and practical knowledge of different digital forensics, tools, platforms, and their functionalities.
4. Provide students with cost-effective and hands-on experience using open-source tools.
5. Support the incoming CYAOP major at our Commonwealth Campuses.
6. Increase instructor effectiveness in the classroom and in a hybrid course sharing environment.
7. Enhance overall faculty instructional effectiveness in SRA/CYBER/IST courses.

5. TEACHING EXERCISE LEARNING STEPS

NIST has several other mobile device image files that can be downloaded from the link below:
<https://www.cfreds.nist.gov/mobile/index.html>

The questions below were asked of the students in the IST 453 course based on the image file provided by the guest speaker, Brett Creasy, from the cyberforensics management team at *bit-x-bit* and led to a further understanding of suggestions by NIST (2020).

1. What is the theme of the criminal investigation?
 - a. What data did you review to determine this?
2. What general areas (location) was the phone used in?
 - a. What data might you review to determine this?
3. Are there any 3rd party apps installed on the phone?

- a. Where might you look to determine this?
4. If a raid was performed on the hotel they are staying in, what specific additional electronic device would be of interest in the investigation?
5. What are some of the interesting terms you uncovered in your review of the information (ex: What were some of the terms Brett mentioned in his speech? Do some additional research if needed).
6. What other information did you uncover? (ex: Were there "cover" names that seemed strange or odd in the conversations?)

The specific image file for this exercise used has since been removed from the NIST site. Many images are placed on the site and subsequently removed by educators; there are currently (as of 2023) twenty images and some access to archived images are available. Although we are unable to provide the original link to the cited image for this exercise, the NIST site cited above has images for download to create your own step and develop similar questions for discovery as those mentioned above. The upload for analysis is the free Cellebrite Reader software <https://cellebrite.com/en/cellebrite-software/cellebrite-reader/>. Cellebrite is a digital forensics tool that preserves the integrity of evidence data throughout the investigation process, which is known as validation. Cellebrite as a tool is designed to assist in the validation of forensic evidence so that it holds admissibility in court and decreases the time spent acquiring data from a mobile device by leveraging an aimed approach (Wilson & Chi, 2018).

Additional Steps/Suggestion(s)

The opportunity to provide additional step(s) to the "module" was created through an educational institution grant from Paraben Corporation, who provided licenses at no cost to our often resource-challenged students. This grant made it possible to create instructions suited for college students with little or no knowledge of the DFIR process or software.

Paraben

The Paraben E3 Android evidence file uses a scenario with an Android device confiscated from a 16-year-old user involved in a possible drug ring incident at a local high school. After installing Paraben E3 and activating using Internet licenses, students download the **SCH-R740C-AndroidEDU1.0.ds** evidence file from a shared course location, confirming the SHA256 hash of the downloaded Paraben E3 Android evidence file.

After starting Paraben E3, students click "Add Evidence" and enter a case name. In the **Add New Evidence** wizard, students select "Paraben Tools" from the category list and then select "E3 mobile data case file/DS case file" from the source type list. Finally, students navigate to the downloaded **SCH-R740C-AndroidEDU1.0.ds** evidence file containing data acquired from the Android device. Since Paraben E3 cases can have multiple evidence files, students are asked to name this new evidence before E3 provides the Case Content in leftmost frame allowing students to navigate evidence by double-clicking on items to drill-down in case hierarchy. Selecting any specific item in the case hierarchy displays that item properties in rightmost frame of E3.

The original Paraben student lab manual simply lists questions to answer by finding from the evidence file. We requested that our undergraduate students answer these questions in the context of a complete forensic report as output from the NIST-recommended forensic process. We wanted to confirm that our students could complete step seven of data interpretation by analyzing all the extracted data artifacts to develop a narrative and timeline of events relevant to the case scenario from the evidence. Students are instructed to answer the following eight questions derived from the original Paraben E3 Android student lab manual within their created forensic case report for the lab scenario investigation:

1. Mobile device information: What is the model of this device? Is a subscriber identity module (SIM) card present in the device? What is the device International Mobile Equipment Identity (IMEI)? What firmware is the device running?
2. Is there an email account set up on the device? If so, what is the email address? How many contacts are on the device?
3. Is there a Secure Digital (SD) card present in the device? How many photos are on the device? Are any photos relevant to the case?
4. How many Apps are there on the device? Which Apps have relevant information to this case? How can you obtain information from these Apps?
5. How many Short Message Service (SMS) are there on the device? How many Multimedia Messaging Service (MMS) are there on the device?
6. Does the device contain any user-entered calendar entries? If so, do any relate to this case?
7. What internet searches have been executed on the device?

8. Who is the owner of the device? Is the owner involved in the drug ring? If so, what role does the owner play?

6. CONCLUSION AND FUTURE ADDITIONS

Instead of being a one-day, one-off course, this exercise has been an attempt to create a cohesive module within the overall curriculum that connects with other course knowledge progressions. According to NIST (2022), digital investigation techniques require knowledge of how tools function and how they are limited. Forensic tool functionality and limitations are impacted by the types of devices being investigated, with increased focus on mobile devices being used to conduct and collaborate illegal activities and to build timelines at locations of key events. Students need hands-on exercises conducting digital forensics on mobile devices with industry tools to better prepare them to provide accurate, timely investigations including internal organization investigations upon graduation. The Paraben education grant was instrumental in allowing our participating Penn State University Commonwealth Campuses to bring this experience with mobile device forensic tools and to reduce the learning curve with these tools after graduation. As we continue to improve our lab activities based on receiving future Paraben education grants, some areas that we plan to enhance include supporting direct Android data acquisition from an Android virtual machine and conducting iPhone mobile device forensics.

7. REFERENCES

- Alghamdi, M. I. (2021). Digital Forensics in Cyber Security—Recent Trends, Threats, and Opportunities. In (Ed.), *Cybersecurity Threats with New Perspectives*. IntechOpen. <https://doi.org/10.5772/intechopen.94452>
- Beardall, D. (2023). Unveiling the Digital Shadows: Cybersecurity and the Art of Digital Forensics. *Cyber Operations and Resilience Program Graduate Projects*. 5. https://scholarworks.boisestate.edu/cyber_gradproj/5
- Casey, E. (2011). *Digital evidence and computer crime: Forensic science, computers, and the internet*: Academic press.
- Cyberseek. NIST. (2021, November 19). Retrieved July 25, 2022, from <https://www.nist.gov/itl/applied-cybersecurity/nice/cyberseek>

- Easttom, C. (2022). *Digital Forensics, Investigation, and Response* (4th ed.). Jones & Bartlett Learning.
- Free Digital Forensic Tools*. Paraben Corporation. (n.d.). <https://paraben.com/free-dfir-tools/>
- Lallie, H. S., Shepherd, L. A., Nurse, J. R., Erola, A., Epiphaniou, G., Maple, C., & Bellekens, X. (2021). Cyber security in the age of COVID-19: A timeline and analysis of cyber-crime and cyber-attacks during the pandemic. *Computers & Security, 105*, 102248. <https://doi.org/10.1016/j.cose.2021.102248>
- Joseph, D. P. & Norman, J. (2019). An Analysis of Digital Forensics in Cyber Security. *Advances in Intelligent Systems and Computing First International Conference on Artificial Intelligence and Cognitive Computing:701–708*. https://doi.org/10.1007/978-981-13-1580-0_67
- Marcella, A.J. (Ed.). (2021). *Cyber Forensics: Examining Emerging and Hybrid Technologies* (1st ed.). CRC Press. <https://doi.org/10.1201/9781003057888>
- Mabuto, E. K., & Venter, H. S. (2011). State of the Art of Digital Forensic Techniques. In *Information Security South Africa 2011 Conference Proceedings*.
- National Institute of Standards and Technology (NIST) (2022, May). *Digital Investigation Techniques: A NIST Scientific Foundation Review*, NISTIR 8354-DRAFT. U.S. Department of Commerce: Gaithersburg, MD. Retrieved from <https://doi.org/10.6028/NIST.IR.8354-draft>
- Online incident response training with Brian Carrier*. Cyber Triage. (2022, April 15). Retrieved July 5, 2022, from <https://www.cybertriage.com/training/>
- Srinivas, J., Das, A. K., & Kumar, N. (2019). Government regulations in cyber security: Framework, standards and recommendations. *Future generation computer systems, 92*, 178-188. <https://doi.org/10.1016/j.future.2018.09.063>
- Whitman, M. E., & Mattord, H. J. (2021). *Principles of Information Security*. Cengage Learning.
- YouTube. (n.d.). *Paraben Forensics*. YouTube. <https://www.youtube.com/user/ParabenForensics>
- Wilson, R., & Chi, H. (2018). A framework for validating aimed mobile digital forensics evidences. *Proceedings of the ACMSE 2018 Conference*. <https://doi.org/10.1145/3190645.3190695>

APPENDIX A
Using Mobile Forensic Software Survey - Pre

Using Mobile Forensic Software Survey - PRE

* Indicates required question

How do you feel about Learning Mobile Forensic Software?*



Choose

What is your intended major?*

- IST
- Cybersecurity






Are you pursuing an SRA minor?*

- Yes
- No
- Maybe

What academic semester are you in currently?*






- 3rd
- 4th
- 5th
- 6th
- 7th
- 8th
- 9th or higher

Do you feel prepared based on your PREVIOUS COURSES to complete a mobile forensic investigation?*

1	2	3	4	5
				
<p>Scared/sad</p> <p>I am not prepared to do this</p>	<p>Uneasy</p> <p>I do not know if I can do this</p>	<p>Unsure</p> <p>A little nervous</p>	<p>Somewhat Confident</p> <p>I just need a little more information</p>	<p>Confident</p> <p>I know I can do this</p>






Choose

How comfortable do you feel using mobile forensic software to complete an investigation?*

1	2	3	4	5
				
<p>Scared/sad</p> <p>I am not prepared to do this</p>	<p>Uneasy</p> <p>I do not know if I can do this</p>	<p>Unsure</p> <p>A little nervous</p>	<p>Somewhat Confident</p> <p>I just need a little more information</p>	<p>Confident</p> <p>I know I can do this</p>

Choose

What was your comfort level with the course material BEFORE the use of mobile forensics software was introduced?*

1	2	3	4	5
				
<p>Scared/sad</p> <p>I am not prepared to do this</p>	<p>Uneasy</p> <p>I do not know if I can do this</p>	<p>Unsure</p> <p>A little nervous</p>	<p>Somewhat Confident</p> <p>I just need to read and study</p>	<p>Confident</p> <p>I know I can do this</p>

Choose

Did you complete the DFIR two-hour training yet?*

- Yes
- No

What can be done to increase your comfort level completing mobile forensic investigations?*






Your answer

APPENDIX B
Using Mobile Forensic Software Survey – Post

Using Mobile Forensic Software Survey - POST

* Indicates required question

How do you feel about Learning Mobile Forensic Software?*

1	2	3	4	5
				
Scared/sad I am not prepared to do this	Uneasy I do not know if I can do this	Unsure A little nervous	Somewhat Confident I just need to read and study	Confident I know I can do this

Choose

What is your intended major?*

- IST
- Cybersecurity

Are you pursuing an SRA minor?*






- Yes
- No
- Maybe

What academic semester are you in currently?*

- 3rd
- 4th






- 5th
- 6th
- 7th
- 8th
- 9th or higher

Do you feel prepared based on the CURRENT COURSE INFORMATION to complete a mobile forensic investigation?*

1	2	3	4	5
				
Scared/sad I am not prepared to do this	Uneasy I do not know if I can do this	Unsure A little nervous	Somewhat Confident I just need a little more information	Confident I know I can do this






Choose

How comfortable do you feel using mobile forensic software to complete an investigation?*

1	2	3	4	5
				
Scared/sad I am not prepared to do this	Uneasy I do not know if I can do this	Unsure A little nervous	Somewhat Confident I just need a little more information	Confident I know I can do this

Choose

What was your comfort level with the course material BEFORE the use of mobile forensics software was introduced?*

1	2	3	4	5
				
Scared/sad I am not prepared to do this	Uneasy I do not know if I can do this	Unsure A little nervous	Somewhat Confident I just need to read and study	Confident I know I can do this

Choose

Did you complete the DFIR two-hour training yet?*

- Yes
- No

What can be done to increase your comfort level completing mobile forensic investigations?*

Your answer

Which program are you more comfortable using:

- Cellebrite
- Paraben
- Other: _____

Phishing: Gender Differences in Email Security Perceptions and Behaviors

Jie Du
dujie@gvsu.edu

Andrew Kalafut
kalafuta@gvsu.edu

Gregory Schymik
schymikg@gvsu.edu

School of Computing
Grand Valley State University
Allendale, MI 49401, USA

Abstract

Information security is a major concern for everyone nowadays. While substantial research exists on gender differences in education and technology, there appears to be very little research on gender differences in information security and that research examines a broad list of self-reported information security behaviors in a single study. Our research adds to the literature by examining in more depth one specific area of information security behavior: peoples' behavior relating to phishing attacks. This research attempts to investigate gender differences in email security perceptions and behaviors by surveying students, faculty, and staff at one midwestern public, master's granting university. The survey questions are developed based on the Health Belief Model. 414 usable survey response sets were collected and analyzed. The findings suggest that men and women have different perceptions on self-efficacy, vulnerability, barriers, cues to action, and self-reported security behaviors. While the Health Belief Model provides a relatively good fit in explaining email security behaviors for both men and women, each group appears to value each of the underlying factors differently. The findings shed light on how to design and conduct security training to increase adoption of protective email behaviors.

Keywords: Security behaviors, security perceptions, gender difference, phishing, health belief model, email.

Recommended Citation: Du, J., Kalafut, A., Schymik, G., (2024). Phishing: Gender Differences in Email Security Perceptions and Behaviors. *Cybersecurity Pedagogy and Practice Journal*, v3(n2) pp 35-47. <https://doi.org/10.62273/PELX2965>

Phishing: Gender Differences in Email Security Perceptions and Behaviors

Jie Du, Andrew Kalafut and Gregory Schymik

1. INTRODUCTION

Anyone paying attention to current events in academia has seen multiple reports of cyber-attacks on universities and many members of the academy can say, with confidence, that their institution has experienced some sort of significant cyber-attack resulting in some sort of network access by unauthorized persons with intentions to profit in some way from their attacks. A 2018 report on cybersecurity in education identifies the education industry as the lowest cybersecurity performer compared to all other industries (SecurityScorecard, 2018). According to atlasVPN (2022), over 80% of malware attacks around the world were found to be targeting the education industry. In 2021 alone, ransomware attacks against US schools and colleges have been estimated to have cost at least \$3.5 billion in downtime (Bischoff, 2022). Most of these attacks are the result of successful phishing attacks.

Given the constant funding pressures faced by universities, they are forced to seek out lower cost solutions to security challenges. The most obvious of these low-cost solutions to problems caused mostly by human behaviors (responses to phishing attacks) is training the users of the systems to be vigilant against these phishing attacks. To do this, institutions must understand the drivers of those behaviors. A part of that understanding must include understanding the differences between security perceptions and behaviors between genders, assuming they occur, so that training can be better targeted to specific people to address their specific needs and tendencies. This paper attempts to aid in that understanding.

We follow others in the IS literature and apply the Health Belief Model - a theoretical model built off of the Technology Acceptance Model, the Theory of Planned Behavior, Protection Motivation Theory, and Expectancy-Value Theory - to the examination of the email security behaviors of students, faculty, and staff at a Midwest, master's-granting university in the United States. We previously investigated gender differences in email security perception and behaviors in male and female students. This paper expands on that

earlier work (Du & Schymik, 2018) with a broader sample of students, faculty, and staff in the institution and aims to investigate the gender-related determinants to people's email security behavior, including whether these determinants are different between students and faculty/staff.

2. LITERATURE REVIEW

Gender Differences in Cybersecurity

Gender differences in a variety of cybersecurity issues have been investigated in previous literature. Anwar et al. (2017) investigated gender differences related to a wide variety of security issues, including email, malware, social media privacy, passwords, backups, and protection of sensitive information. They found significant differences in self-reported behaviors. However, since they considered a variety of behaviors in a single construct it is difficult to determine what specific behaviors are influenced by gender. McGill and Thompson (2021) studied gender differences in security and privacy behaviors, finding differences in 40% of the behaviors studied, with men practicing more preventive security and privacy behaviors than women.

Farooq et al. (2015) examined information security awareness, knowledge, and behaviors among university students, finding male students to have better awareness knowledge, and behaviors than female students. Conversely, McCormac et al. (2017) found better information security awareness among females, in a study of working Australians. Combined, these indicate that gender differences among students may not match those of employees, motivating our effort to study gender differences in the two populations separately.

More specific to email, the literature on gender differences and phishing attacks shows some mixed results on the subject. Sheng et al. (2010) found that women fell for phishing attacks more often than did men while Diaz et al. (2020), and Benenson et al. (2017) found no differences in susceptibility. Verkijika (2019) found no differences in mobile phishing avoidance motivation and behavior, but they did find that gender was a significant moderator of the effect

of anti-phishing self-efficacy on both of those factors.

Health Belief Model

In healthcare, preventive healthcare behaviors are behaviors that will lessen the harmful effects of diseases, such as vaccination, diet, and exercise. In computer security, protective security behaviors are those behaviors that will lessen the harmful effects of security incidents, such as using antivirus software or checking URLs before clicking on them. Although the fields of healthcare and security are very different, these ideas are significantly similar: both are behaviors that people can follow in order to protect themselves from potential harm. Therefore, it is reasonable to consider that similar theories may explain both.

The Health Belief Model (HBM) was originally developed to explain preventive health behaviors (Rosenstock, 1974). In early versions of the model, a person’s attitude towards preventive health behaviors was considered a function of the perceived susceptibility and perceived severity of the illness, as well as the perceived benefits and perceived barriers to performing the preventive

health behavior. Later work, a decade after the original, added three additional variables: self-efficacy, cues to action, and general health orientation (Janz, 1984).

Relying on the similarity of the preventive health and protective security behaviors, the HBM has since been applied to explain protective security behaviors. Such applications have covered several security domains, including the use of email (Ng et al., 2009), the adoption of computer security software (Claar et al., 2013), and how to prevent unauthorized access to computers (Williams et al., 2014), and the use of antivirus software (Dodel & Mesch, 2017).

Health Belief Model and Gender Differences

In the healthcare field, differences in HBM factor significance by gender have been observed in several studies. For example, perceived barriers and self-efficacy were significant determinants of oral hygiene behaviors among males, while only self-efficacy was significant among females (Zetu et al., 2014). As a further example, gender has been shown to be a common modifying factor in applying the HBM to COVID-19 vaccine hesitancy (Limbu et al., 2022).

Index	Name	Definition
BEN	Perceived Benefits	A user’s belief in the perceived effectiveness of practicing email security behavior
BAR	Perceived Barriers	A user’s perceived cost and inconvenience of practicing email security behavior
EFF	Self-Efficacy	A user’s self-confidence in his skills/ability in practicing email security behavior
VUL	Perceived Vulnerability	A user’s perceived likelihood of an email security incident occurrence
CUE	Cues to Action	Experiences that would activate a user to practice email security behavior
EXP	Prior Experience	A user’s previous experience with email security incidents
SEV	Perceived Severity	A user’s perceived seriousness of an email security incident
BEH	Self-reported Behavior	A user’s self-reported behavior when using emails

Table 1: The Definitions of Constructs in the Research Model

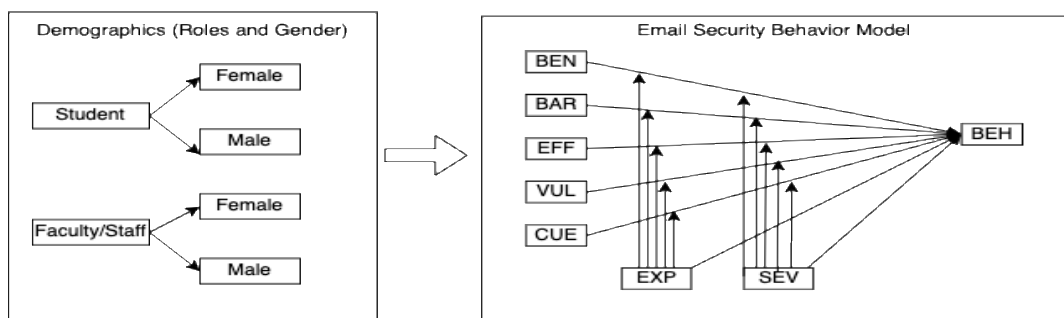


Figure 1: The Research Model

In the security field, gender differences in the HBM have been less explored. However, the problem has been approached by Anwar et al. (2017), who found significant gender differences on self-efficacy, prior experience, and computer skills among a group of employees. Their study does not include students. Fatokun et al. (2019) explored gender differences in the cybersecurity behaviors of students at Malaysian universities, using a model based off of the HBM combined with protection motivation theory. They found higher self-efficacy among males and higher perceived severity among females. Although they used a new model based on the health belief model, not the HBM itself, since these factors are also used in the HBM they may imply gender differences in the corresponding HBM constructs.

Although not much previous work appears to have been done on investigating gender differences in the application of the health belief model to cybersecurity, the fact that gender differences have been observed in cybersecurity, and the fact that gender differences have been observed in the application of the health belief model to healthcare, indicates that this may be a promising avenue to gain insight into the causes of gender differences in cybersecurity.

3. RESEARCH MODEL AND HYPOTHESES

This study aims to investigate gender-related determinants to peoples' email security behaviors and whether these determinants are different between their professional roles (e.g., Student vs faculty/staff). We hope that knowledge of such differences may shed light on how to design and conduct security training to increase adoption of beneficial email security behaviors. The research model used in this study is based upon the research model from our prior study (Schymik & Du, 2018) (see Figure 1). Demographics including professional role and gender are taken into consideration when examining people's email security behaviors. The model measures the main effects of the constructs and the interaction effects moderated by prior experience and perceived severity. Table 1 describes each construct.

Our research begins by asking two broad questions of the dataset collected in prior research. The first examines the latent variables in the model and asks if there are difference between the 8 latent variables generated by our subjects' survey responses between the genders.

- Q1: Do gender differences exist in our subjects' security perceptions and behaviors?

The second question looks at the results of the regression analysis performed using the latent factor scores and asks if gender impacts which factors in the model are significant.

- Q2: Do the factors that impact people's email security behavior differ by gender?

Gender differences were identified in our previous study (Du & Schymik, 2018). This study expands that research by surveying a larger sample including students, faculty, and staff at a university with aim of investigating the gender-related determinants to their email security behaviors. This larger sample population provides us the opportunity to explore whether professional roles (student vs faculty/staff) have an impact on email security perceptions and behaviors. It has been noted that the higher education workforce has higher levels of information security awareness than do students and might perceive information security compliance as a protective measure that may prevent security-breach-related disasters (Sari et al., 2016). Pursuing the suggestion that other underlying factors may have a significant effect on email security behavior (Greitzer et al., 2021) and noting that our latest work on this data set indicated differences between the two professional roles (Authors 2023, under review), we extend the initial research questions to investigate the impact professional role may have on gender differences in email security perceptions and behaviors:

- Q3: Do gender differences exist within the professional roles (student vs faculty/staff)?
- Q4: Do the factors that impact people's email security behaviors within professional roles differ by gender?

We originally posited two hypotheses:

- H1: There are gender differences in people's email security perception and behaviors.
- H2: Determinants of men's and women's security behaviors are different.

Acknowledging that these hypotheses may be too broadly defined, we focus on Q3 and Q4 above, and expand our two initial hypotheses to explore the within group gender differences:

- H3: Gender differences in people's email security perception and behaviors exist among students.
- H4: Gender differences in people's email security perception and behaviors exist among faculty/staff.
- H5: Determinants of security behaviors differ between genders among students.

- H6: Determinants of security behaviors differ between genders among faculty/staff.

Please note that due to page length limitations, the research questions and hypotheses are stated generically instead of listing one for each of the latent variables in the research model where appropriate.

4. METHODOLOGY

Participants and Procedure

The target population for this study was students, faculty, and staff at one midwestern public, master's-granting university in the United States. A random sample of the target population was contacted via email and invited to participate by completing an anonymous online questionnaire. The email provided the purpose and procedure to participants in this study along with other information required by the university. All participants were 18 or over and consented to participate.

Survey Development

An electronic Likert-scale questionnaire was developed to measure the constructs in our research model (see Appendix A). The first section of the survey contains 8 items on demographics. The second section of the survey contains 32 items to measure the participants' security perceptions and behaviors validated from prior research (Ng et al., 2009) (Claar et al., 2013). The items to measure the model constructs are anchored on a 5-point scale, which ranged from strongly disagree (1) to strongly agree (5) for most of the items. The scales used for the items that measure security behavior and prior experience are different. The 5-point Likert scale ranged from never (1) to every time (5) for security behavior and from never (1) to a great deal (5) for prior experience. The survey was

anonymous and administered using the Qualtrics online survey platform. The Internal Review Board (IRB) of the university approved the study.

We sent out the survey to a random sample of students and a random sample of faculty/staff at the aforementioned university. After removing responses with missing data, the date collection yielded 417 remaining survey response sets. The gender question in the demographic section of the survey presented three options: male, female, and other. Only three responses chose "other" for this question. Due to the very small number of responses in this category, these responses were excluded from further analysis, resulting in 414 total response sets used in our analysis. The whole dataset contains 149 male subjects and 265 female subjects. Our sample shows a similar female to male ratio (64%/36%) to the gender distribution in the population (61%/39%) and thus is representative of the university community. Figure 2 shows the gender distribution in our datasets.

Data collected in this study were analyzed in three different datasets. The whole dataset contains all 414 response sets. Based on the participants' role at the university, the whole dataset was divided into a student dataset of 100 students and an employee dataset of 314 employees. Each dataset was further split into two gender groups.

To test the hypotheses for each dataset, a three-step analysis was conducted. First, an exploratory factor analysis was conducted to extract factors that impact the participants' email security behaviors. As a result, eight factors were extracted; these factors are consistent with the eight constructs in our research model. Cronbach Alpha coefficients were calculated for each latent variable..

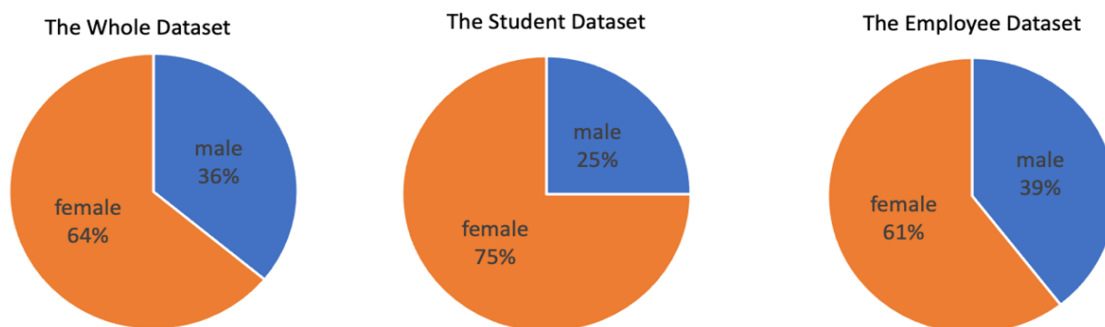


Figure 2: Gender Distribution Data Analysis

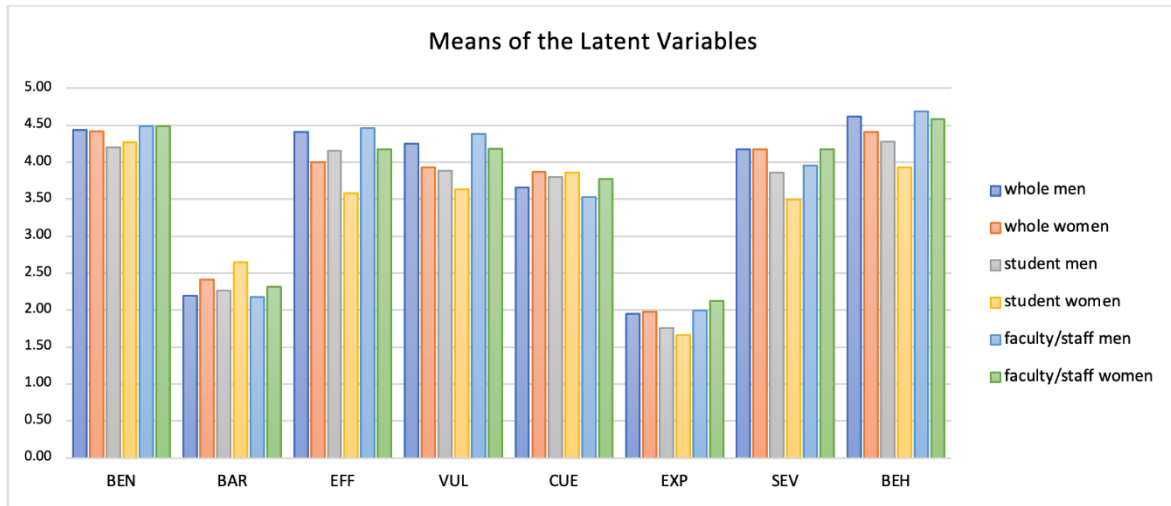


Figure 3: Means of the Latent Variables

All constructs exhibited acceptable construct validity and reliability, allowing us to proceed with our regression analysis. Second, an independent-samples two-tailed *t*-test was conducted to compare means of each latent variable in the two gender groups (see Figure 3) to investigate whether significant differences exist on any latent variable based on gender. Finally, a regression analysis (main effects first and interaction effects second) for each gender group in each dataset was conducted to examine whether women and men value the factors impacting their security behaviors differently. All the data were analyzed using IBM Statistical Package for Social Science (SPSS) version 25. All statistical tests were conducted with an alpha level of <0.05.

5. RESULTS

Gender differences were detected in all 3 datasets. In this section, the data analysis results were organized based on the dataset. In each dataset, the results of the *t*-test on the mean of each variable are reported first. An independent two-tailed *t*-test was run on each of the three datasets, resulting in 3 examinations (see Tables 2 – 4). Followed were the regression analysis results. The research model was assessed separately for each gender group in each of the three datasets, thus resulting in 6 examinations (see Tables 5 - 6).

Whole Dataset

For the whole dataset, the mean values between women and men were statistically different on five variables: EFF, VUL, BAR, CUE, and BEH. No gender differences existed on other factors. These results indicate support for H1. Table 2

shows the *t*-test results on the whole dataset. These results suggest that men have higher levels of perceived self-efficacy and perceived vulnerability, and self-report better security behaviors than women do. In contrast, women have higher levels of perceived barriers and cues to action than do men.

	Men		Women		t	P	Sign. Diff?
	(N=149)		(N=265)				
	Mean	SD	Mean	SD			
BEN	4.44	0.70	4.42	0.62	0.271	.787	ns
BAR	2.19	0.74	2.41	0.77	-2.828	.005	**
EFF	4.41	0.62	4.00	0.78	5.506	.000	***
VUL	4.25	0.73	3.93	0.81	4.079	.000	***
CUE	3.66	0.81	3.87	0.62	-2.849	.005	**
EXP	1.95	0.86	1.99	0.80	-0.437	.662	ns
SEV	4.17	0.97	4.17	0.85	-0.101	.919	ns
BEH	4.62	0.48	4.41	0.68	3.751	.000	***

ns: non-significant; **p*<.05, ***p*<.01, ****p*<.001.

Table 2: Comparison of Means in the Whole Dataset

The results of the regression indicated that: (1) in the main effect model (see Table 5), while women’s email behaviors are influenced by EFF, VUL, EXP, and BAR, men’s email behaviors are only influenced by EFF, and (2) in the interaction effect model (see Table 6), EFF, EXP, VUL, BEN, EXPxBEN, EXPxVUL, and SEVxEFF are the significant determinants on women’s email behavior while EFF remains the only significant factor that impact men’s email behavior. These results support H2. These results suggest that while men’s decision-making process regarding practicing protective email security behaviors is mainly influenced by their self-efficacy, women’s decision-making process regarding the same behaviors is impacted by several factors (including self-efficacy, like men).

Student Dataset

For the student dataset, there was a significant difference in the mean values of EFF and BAR for men and women. No gender differences existed on other factors. H3 was supported. Table 3 shows the *t*-test results on the student dataset. These results suggest that the male students have higher self-efficacy to practice protective email behaviors than their female peers while the female students have higher perceived barriers of practicing protective email behaviors than their male peers.

	Men		Women		t	P	Sign. Diff?
	(N=25)		(N=75)				
	M	SD	M	SD			
BEN	4.20	0.91	4.27	0.66	-0.425	.672	ns
BAR	2.26	0.77	2.65	0.77	-2.201	.030	*
EFF	4.16	0.72	3.58	0.95	2.814	.006	**
VUL	3.89	1.33	3.63	1.00	1.061	.291	ns
CUE	3.80	1.01	3.86	0.92	-0.275	.784	ns
EXP	1.76	0.82	1.66	0.71	0.601	.549	ns
SEV	3.86	1.25	3.49	1.14	1.36	.177	ns
BEH	4.28	0.68	3.93	1.01	1.623	.108	ns

ns: non-significant; *p<.05, **p<.01, ***p<.001.

Table 3: Comparison of Means in the Student Dataset

The results of the regression show that: (1) in the main effect model (see Table 5), EFF and EXP are the two significant determinants for the email behavior in the female student group while EFF and BAR are the significant determinants to the male students' email behaviors, and (2) in the interaction effect model (see Table 6), EFF, EXP, and SEVxBAR are the significant determinants to the female students' email behaviors while no factor was found to significantly impact the male students' email behaviors. H5 was supported. These results in the main effect model suggest that self-efficacy is a factor that impact both the male and female students regarding their email behaviors. Besides self-efficacy, the female students are more influenced by prior experience

while the male students are more focused on perceived barriers when practicing email behaviors. It is interesting to note that none of the factors impacted the male students' behaviors when the interaction effects were taken into consideration. This will be addressed in the discussion section.

Employee Dataset

There was a significant difference in the mean values of BEH, EFF, CUE, and VUL for the male and female employees. Table 4 shows the *t*-test results on the employee dataset. H4 was supported. These results suggest that the male employees had higher levels of self-efficacy and perceived vulnerability and self-reported better email security behaviors than their female peers. On the other hand, the female employees were more influenced by cues to action when practicing such behaviors than their male peers.

	Men		Women		t	P	Sign. Diff?
	(N=124)		(N=190)				
	M	SD	M	SD			
BEN	4.49	0.64	4.49	0.60	0.094	.925	ns
BAR	2.18	0.74	2.32	0.75	-1.624	.105	ns
EFF	4.46	0.59	4.17	0.63	4.129	.000	***
VUL	4.38	0.51	4.18	0.63	3.053	.002	**
CUE	3.53	0.91	3.77	0.75	-2.481	.014	*
EXP	1.99	0.86	2.12	0.79	-1.352	.177	ns
SEV	3.96	1.32	4.17	1.04	-1.517	.131	ns
BEH	4.69	0.41	4.58	0.46	2.100	.037	*

ns: non-significant; *p<.05, **p<.01, ***p<.001.

Table 4: Comparison of Means in the Employee Dataset

The results of the regression indicated that in the main effect model (see Table 5), EFF is the only determinant for the female employees' email behaviors while EFF and EXP are two significant determinants for the male employees' email behaviors.

Variables	Whole		Students		Employees	
	Male	Female	Male	Female	Male	Female
BEN	0.083	0.078	0.103	0.112	0.003	0.093
BAR	-0.065	-.108*	-0.458*	-0.044	-0.067	-0.095
EFF	0.483***	0.521***	0.778**	0.508***	0.401***	0.362***
VUL	0.126	.160**	-0.288	0.074	0.032	0.044
CUE	0.083	0.034	0.149	0.057	0.142	-0.021
EXP	-0.032	0.185***	-0.381	0.353***	-0.202*	-0.023
SEV	0.074	-0.001	-0.379	-0.163	0.134	-0.017
R ²	0.358	0.515	0.695	0.542	0.238	0.218
adjusted R ²	0.326	0.502	0.569	0.494	0.192	0.188

*p<.05, **p<.01, ***p<.001.

Table 5. Regression Results on Main Effects Model

Variables	Whole		Students		Employees	
	Coefficient		Coefficient		Coefficient	
	Male	Female	Male	Female	Male	Female
BEN	0.137	0.123*	0.572	0.079	0.066	0.145
BAR	-0.062	-0.1	-0.667	-0.019	-0.056	-0.126
EFF	0.498***	0.455***	0.695	0.5***	0.380**	0.311***
VUL	0.121	0.150**	-0.753	0.099	0.018	0.033
CUE	0.045	0.026	0.187	0.123	0.078	-0.024
EXP	-0.065	0.162***	-0.088	.340**	-0.318**	-0.049
SEV	0.183	0.004	-0.249	-0.099	0.275	-0.012
EXPXBEN	-0.159	-0.128*	0.76	-0.03	-0.171	-0.133
EXPXBAR	-0.088	-0.005	0.388	0.043	0.024	-0.035
EXPXEFF	-0.02	-0.079	-0.598	-0.043	0.129	-0.106
EXPXVUL	-0.009	-0.101*	-0.309	-0.086	0.087	0.094
EXPXCUE	-0.052	0.013	-0.085	-0.035	-0.033	-0.027
SEVXBEN	0.128	0.007	0.67	-0.183	0.024	0.116
SEVXBAR	0.057	-0.049	0.279	-0.310**	-0.159	0.091
SEVXEFF	-0.082	-0.115*	-0.406	-0.123	-0.185	-0.028
SEVXVUL	0.116	0.07	0.28	0.075	-0.188	0.127
SEVXCUE	0.007	-0.087	0.118	0.151	-0.119	-0.109
SEVXEXP	0.141	0.018	0.19	0.017	0.187	0.002
R ²	0.402	0.59	0.908	0.668	0.293	0.292
adjusted R ²	0.319	0.56	0.632	0.561	0.171	0.218

*p<.05, **p<.01, ***p<.001.

Table 6: Regression Results on Interaction Effects Model

6. DISCUSSION

The results of the regression on the interaction effect model (see Table 6) are consistent with the findings of the main effect model: EFF is the only significant determinant on the female employees' email behavior while EFF and EXP are two significant determinants on the male employees' email behaviors. Therefore, H6 was supported. These results suggest that self-efficacy is a factor that impact both female and male employees' email behaviors. However, the male employees are also influenced by their experience of prior email security incidents even though they reported lower levels of prior experience than the female employees did.

Gender Differences on Security Perceptions and Behaviors

The results of this study show evidence of gender differences in people's security perceptions and behaviors. For the whole sample, there are statistically significant differences in terms of self-efficacy, vulnerability, barriers, cues to action, and self-reported email security behavior based on gender. Our findings of men having higher levels of self-efficacy and security behaviors are consistent with prior research (Anwar et al., 2017). Men are more confident of their capability to practice protective email behaviors and self-

report higher levels of email security behaviors than women do. We also found that women perceive higher levels of barriers to practicing email security behaviors than men do. This might be related to women's lower levels of computer skills (Anwar et al., 2017) or less knowledge of computers (He & Freeman, 2009). It is interesting to find that our female participants reported a significantly lower likelihood of receiving unsafe email than the male participants did. The female participants did not feel more vulnerable to phishing email. Previous studies have found that women have a greater susceptibility to phishing email (Jagastic et al., 2007). Jagastic et al. (2007) utilized several roleplay tasks to measure people's susceptibility to phishing and they found that women are more likely than men to click on phishing links and giving information to phishing websites due to less technical training and less technical knowledge than men. The gap in women's perception of vulnerability and their actual security behaviors needs further exploration.

For the employee sample, the results are consistent with the findings in the whole sample except for perceived barriers. There is no significant difference in perceived barriers between our female and male employee participants. A possible explanation is that

Variables	Whole		Students		Employees	
	Coefficient		Coefficient		Coefficient	
	Male	Female	Male	Female	Male	Female
BEN	0.137	0.123*	0.572	0.079	0.066	0.145
BAR	-0.062	-0.1	-0.667	-0.019	-0.056	-0.126
EFF	0.498***	0.455***	0.695	0.5***	0.380**	0.311***
VUL	0.121	0.150**	-0.753	0.099	0.018	0.033
CUE	0.045	0.026	0.187	0.123	0.078	-0.024
EXP	-0.065	0.162***	-0.088	.340**	-0.318**	-0.049
SEV	0.183	0.004	-0.249	-0.099	0.275	-0.012
EXPXBEN	-0.159	-0.128*	0.76	-0.03	-0.171	-0.133
EXPXBAR	-0.088	-0.005	0.388	0.043	0.024	-0.035
EXPXEFF	-0.02	-0.079	-0.598	-0.043	0.129	-0.106
EXPXVUL	-0.009	-0.101*	-0.309	-0.086	0.087	0.094
EXPXCUE	-0.052	0.013	-0.085	-0.035	-0.033	-0.027
SEVXBEN	0.128	0.007	0.67	-0.183	0.024	0.116
SEVXBAR	0.057	-0.049	0.279	-0.310**	-0.159	0.091
SEVXEFF	-0.082	-0.115*	-0.406	-0.123	-0.185	-0.028
SEVXVUL	0.116	0.07	0.28	0.075	-0.188	0.127
SEVXCUE	0.007	-0.087	0.118	0.151	-0.119	-0.109
SEVXEXP	0.141	0.018	0.19	0.017	0.187	0.002
R ²	0.402	0.59	0.908	0.668	0.293	0.292
adjusted R ²	0.319	0.56	0.632	0.561	0.171	0.218

*p<.05, **p<.01, ***p<.001.

Table 6: Regression Results on Interaction Effects Model

perceived barriers might be moderated by prior experience with email security incidents. Our female employee participants reported higher levels of prior experience on email incidents than that of their male peers. The greater experience of email incidents might reduce our female employees' perceived barriers and possibly mediate the gender effect on it. It is also possible to argue that prior experience with email incidents is more likely to confound gender differences (Venkatesh et al., 2000).

For the student sample, the male students had higher self-efficacy while the female students had higher levels of perceived barriers when practicing email security behaviors. These findings are consistent with prior research (Anwar et al., 2017). It is interesting to note that several gender differences that existed in the whole dataset and the employee dataset were not present in our student dataset. Compared to the more diverse employee sample, our student sample is more homogeneous. It is possible that gender differences in perception could be potentially confounded by other demographic variables, such as education level and organization level (Venkatesh et al., 2000). This needs to be explored further.

Our findings indicate that women report higher levels of cues to action than men, which is

inconsistent with prior research (Anwar et al., 2017). However, Anwar et al. (2017) used a form of cues to action question that asked only whether participants have received such queues. Using this type of CUE questions, they found that women are less sensitive to cues to action. We used a different form of cues to action in this study. The CUE questions used in our study focus on predicted responses to hypothetical situations (Claar et al., 2013), such as asking if the subject would be more careful about email security if his friend told him an email incident. The difference in the form of cues to action might explain the contradictory finding on cues to action. Our CUE questions lean more toward subjective norm, which refers to the perceived social pressure to perform and not to perform the behavior (Ajzen, 1991). Our findings that women more influenced by cues to action (e.g., weight the opinions of others' more highly than men) are supported in (Venkatesh et al., 2000; Venkatesh & Morris, 2000; McGill & Thompson, 2021).

Gender Differences in Determinants of Security Behaviors

The results of the regression in all 3 datasets indicate that the factors that impact men's email security behavior are different from the ones that impact women. The results in the whole dataset suggest that while men are more focused on their self-efficacy in their decision-making process

regarding practicing protective email security behaviors, women are more balanced on their decision-making process, which is impacted by several factors, including perceived vulnerability, self-efficacy, prior experience, and perceived benefits. This is further supported by the similar variance in self-reported email security behaviors (W: 0.423, M: 0.498) explained by the significant determinants among women (VUL, EFF, EXP, BEN, EXPxBEN, EXPxVUL, and SEVxEFF) and men (EFF). Our finding is in line with the fact that women are more balanced in the adoption and usage decisions (Venkatesh et al., 2000).

When we examined the employee sample, self-efficacy was a significant factor that impacts both men's and women's email security behavior. Prior experience was a significant factor that impact men's security behavior but not women even though they reported higher levels of prior experience of email incidents. It is notable that gender differences in the employee dataset were somehow less compared to those in the whole dataset. Other important demographic variables, such as organization level and education could potentially confound gender differences (Venkatesh et al., 2000).

For the student sample, self-efficacy and prior experience significantly impact the female students' security behavior. A significant moderating effect of perceived severity on perceived barriers was also found in the female student group. It is interesting to note that we did not find any significant factors in the male student group by using the interaction effect model. When interaction effects were omitted, self-efficacy and perceived barriers are significant factors that impact our male students' security behavior.

As mentioned before, the focus of this paper is gender differences, and our results supported our hypotheses. It is notable that there are some unexpected findings in this study, such as the unexpected coefficient direction of prior experience in the whole dataset as well as the employee dataset. This needs further investigation.

Implications

Gender differences on security perceptions need to be considered when designing security training or awareness programs. The fact that women are more influenced by cues to action and have higher levels of perceived barriers of adopting security behaviors suggest that training differences should be targeted. For instance, incorporating interaction with close family

members or friends in the training could help reduce perceived barriers in females (Dixon, 2014).

To maximize adoption of security behaviors, training might be tailored to emphasize factors that are salient to each gender group. For example, training needs to emphasize self-efficacy for men, while offering women a more balanced approach that includes self-efficacy, perceived benefits, perceived vulnerability, and prior experience, all of which are more important to women. Self-efficacy is a factor that impacts both men's and women's security behavior, while men have a higher reported self-efficacy than women do. This finding shed light on the importance of boosting women's self-efficacy via training programs to improve their security behaviors. The fact that women don't feel more vulnerable to security threats even though they have more prior experience of email incidents suggest a gap between their security perceptions and behaviors. Deploying phishing simulations and educating those who fail such simulations might help close this gap.

Substantial differences between students and faculty/staff were detected in our study. Self-efficacy was the only construct that shows gender-related differences in the student sample while multiple constructs (EFF, VUL, CUE, and BEH) show gender-related differences in the employee sample. Compared to our relatively young student sample, the employee sample has more diverse demographics such as age and education. Also, the employees in the university have gone through several security training sessions. These factors might contribute to the differences and warrant future investigation.

Limitations and Future Research Directions

Our research on gender differences of email security perceptions and behaviors focuses on the university setting in western culture. These issues should be addressed in other settings where email poses security threats. Another limitation of this study is the relatively small sample size for our male student group. While the sample size is acceptable (de Winter, 2013), a much larger sample size would give more reliable statistical results. Another limitation in the current work is the measurement of cues to action. The cues to action questions need to be refined in the future work to clearly measure subjects' responses to whether they have received such cues. Subjective norm, a factor that impact women more (Venkatesh et al., 2000; Venkatesh & Morris, 2000; McGill & Thompson, 2021) needs to be considered into the research model in the future.

Future research is necessary to fully understand gender differences by refining the current gender variable and adding more demographic variables. Future work should investigate gender as a psychological factor based on femininity and masculinity (Venkatesh et al., 2000; Venkatesh & Morris, 2000). Examining age, education level, and organization level in people's security perceptions and behaviors could be explored in the future work that might provide interesting insights (Fatokun et al., 2019).

7. CONCLUSION

This study examined gender differences in email security behaviors from two different measures. The findings reveal different perceptions on self-efficacy, vulnerability, barriers, cues to action, and self-reported security behaviors. The two genders appear to value each of the underlying factors differently. Several factors impact women's security behaviors, including self-efficacy, prior experience, perceived vulnerability, and perceived benefits while self-efficacy is the only factor that impacts men's security behaviors. The gender differences identified in this study provide evidence that gender plays a vital role in shaping people's security perceptions and thus impacting their behaviors. Security training and awareness programs can be designed and conducted more efficiently with these gender differences taken into consideration.

8. REFERENCES

- Ajzen, I. (1991). The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179-211, doi: 10.1016/0749-5978(91)90020-T.
- Anwar, M., He, W., Ash, I., Yuan, X., Li, L., & Xu, L. (2017). Gender Difference and Employees' Cybersecurity Behaviors. *Computers in Human Behavior*, 69, 437-443, doi: 10.1016/j.chb.2016.12.040.
- atlasVPN. (2022). Over 80% of Malware Attacks Target Education Sector as Back-to-School Season Nears. Retrieved Feb. 27, 2023, from <https://atlasvpn.com/blog/over-80-of-malware-attacks-target-education-sector-as-back-to-school-season-nears>.
- Benenson, Z., Gassmann, F., & Landwirth, R. (2017). Unpacking Spear Phishing Susceptibility. In *Proceedings of Financial Cryptography Workshops*, doi: 10.1007/978-3-319-70278-0_39.
- Bischoff, P. (2022). Ransomware Attacks on Us Schools and Colleges Cost \$3.56bn in 2021. Retrieved Feb. 27, 2023, from <https://www.comparitech.com/blog/information-security/school-ransomware-attacks/>.
- Claar, C., Shields, R., Rawlinson, D., & Lupton, R. (2013). College Student Home Computer Security Adoption. *Issues in Information Systems*, 14(2), 139-148, doi: 10.48009/2_iis_2013_139-148.
- de Winter, J. (2013). Using the Student' S T-Test with Extremely Small Sample Sizes. *Practical Assessment, Research, and Evaluation*, 18(1), 10, doi: 10.7275/e4r6-dj05.
- Diaz, A., Sherman, A., & Joshi, A. (2020). Phishing in an Academic Community: A Study of User Susceptibility and Behavior. *Cryptologia*, 44, 53-67, doi: 10.1080/01611194.2019.1623343.
- Dixon, L.J., Correa, T., Straubhaar, J., Covarrubias, L., Graber, D., Spence, J., & Rojas, V. (2014). Gendered Space: The Digital Divide between Male and Female Users in Internet Public Access Sites. *Journal of Computer-Mediated Communication*, 19(4), 991-1009, doi: 10.1111/jcc4.12088.
- Dodel, M., & Mesch, G. (2017). Cyber-Victimization Preventive Behavior: A Health Belief Model Approach. *Computers in Human Behavior*, 68, 359-367, doi: 10.1016/j.chb.2016.11.044.
- Du, J., & Schymik, G. (2018). Gender Differences on Students' Email Security Behaviors. In *Proceedings of the 24th Americas Conference on Information Systems*, 5, <https://aisel.aisnet.org/amcis2018/SocialInclusion/Presentations/5>.
- Farooq, A., Isoaho, J., Virtanen, S., & Isoaho, J. (2015). Information Security Awareness in Educational Institution: An Analysis of Students' Individual Factors. In *Proceedings of IEEE Trustcom/BigDataSE/ISPA. Helsinki, Finland*, 352-359, doi: 10.1109/Trustcom.2015.394.
- Fatokun, F., Hamid, S., Norman, A., & Fatokun, J. (2019). The Impact of Age, Gender, and Educational Level on the Cybersecurity Behaviors of Tertiary Institution Students: An Empirical Investigation on Malaysian Universities. *Journal of Physics: Conference Series*, 1339, doi: 10.1088/1742-6596/1339/1/012098.
- Greitzer, F. L., Li, W., Laskey, K.B., Lee, J., & Purl, J. (2021). Experimental Investigation of Technical and Human Factors Related to Phishing Susceptibility. *ACM Transaction on*

- Social Computing*, 4(2), 1-48, doi: 10.1145/3461672.
- Janz, N., & Becker, M. (1984). The Health Belief Model: A Decade Later. *Health Education Quarterly*, 11(1), 1-47, doi: 10.1177/109019818401100101.
- Jagatic, T., Johnson, N., Jakobsson, M., & Menczer, F. (2007). Social Phishing. *Communications of the ACM*, 50(10), 94-100, doi: 10.1145/1290958.1290968.
- Limbu, Y., Gautam, R., & Pham, L. (2022). The Health Belief Model Applied to Covid-19 Vaccine Hesitancy: A Systematic Review. *Vaccines*, 10(6), 973, doi: 10.3390/vaccines10060973.
- McCormac, A., Zwaans, T., Parsons, K., Calic, D., Butavicius, M., & Pattinson, M. (2017). Individual Differences and Information Security Awareness. *Computers in Human Behavior*, 69, 151-156, doi: 10.1016/j.chb.2016.11.065.
- McGill, T., & Thompson, N. (2021). Exploring Potential Gender Differences in Information Security and Privacy. *Information & Computer Security*, 29(5), 850-865, doi: 10.1108/ICS-07-2020-0125.
- Ng, B.-Y., Kankanhalli, A., & Xu, Y. (2009). Studying Users' Computer Security Behavior: A Health Belief Perspective. *Decision Support Systems*, 46, 815-825, doi: 10.1016/j.dss.2008.11.010.
- Rosenstock, I. (1974). The Health Belief Model and Preventive Health Behavior. *Health Education Monographs*, 2(4), 354-386, doi:10.1177/109019817400200405.
- Sari, P., Nurshabrina, N. & Candiwan (2016). Factor analysis on information security management in higher education institutions. In *Proceedings of the International conference on cyber and IT service management (2016)*, 1-5, doi: 10.1109/CITSM.2016.7577518.
- Schymik, G. & Du, J. (2018). Student Intentions and Behaviors Related to Email Security: An Application of the Health Belief Model. *Journal of Information Systems Applied Research*, 11(3), 14-24, <http://jisar.org/2018-11/> ISSN: 1946-1836.
- Authors, 2023. Under review.
- SecurityScorecard. (2018). 2018 Education Cybersecurity Report. Retrieved Feb. 27, 2023, <https://explore.securityscorecard.com/rs/797-BFK-857/images/SSC-EducationReport-2018.pdf>.
- Sheng, S., Holbrook, M., Kumaraguru, P., Cranor, L., & Downs, J. (2010). Who Falls for Phish? A Demographic Analysis of Phishing Susceptibility and Effectiveness of Interventions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Atlanta, Georgia, USA: Association for Computing Machinery*, 373-382, doi: 10.1145/1753326.1753383.
- Venkatesh, V., & Morris, M.G. (2000). Why Don't Men Ever Stop to Ask for Directions? Gender, Social Influence, and Their Role in Technology Acceptance and Usage Behavior. *MIS Quarterly*, 24(1), 115-139, doi: 10.2307/3250981.
- Venkatesh, V., Morris, M., & Ackerman, P. (2000). A Longitudinal Field Investigation of Gender Differences in Individual Technology Adoption Decision-Making Processes. *Organizational Behavior and Human Decision Processes*, 83(1), 33-60, doi: 10.1006/obhd.2000.2896.
- Verkijika, S. (2019). If You Know What to Do, Will You Take Action to Avoid Mobile Phishing Attacks: Self-Efficacy, Anticipated Regret, and Gender. *Computers in Human Behavior*, 101, 286-296, doi: 10.1016/j.chb.2019.07.034.
- Williams, C., Wynn, D., Madupalli, R., Karahanna, E., & K. Duncan, B. (2014). Explaining Users' Security Behaviors with the Security Belief Model. *Journal of Organizational and End User Computing*, 26(3), 23-46, doi: 10.4018/joeuc.2014070102.
- Zetu, L., Zetu, I., Dogaru, C., Duța, C., & Dumitrescu, A. (2014). Gender Variations in the Psychological Factors as Defined by the Extended Health Belief Model of Oral Hygiene Behaviors. *Procedia - Social and Behavioral Sciences*, 127, 358-362, doi: 10.1016/j.sbspro.2014.03.271.

APPENDIX A Survey Questions

	Age verification— I verify that I am at least 18 years old.
	What is your age?
	Gender (choose the one you identify with most)
	What is your primary role at GVSU?
	How many years have you attended GVSU?
	Are you an undergraduate or graduate student?
	Which of these most closely matches your job function?
	What is the highest degree or level of school you have completed? If currently enrolled, highest degree received.
BEH1	Before opening an email, I first check if the subject and sender make sense. (every time/never)
BEH2	Before opening an email attachment, I first check if the file name of the attachment makes sense. (every time/never)
BEH3	Before clicking on a link in an email, I first check to see if the URL for the link makes sense. (every time/never)
BEH4	Before opening an email attachment, I first check to see if the contents and sender of the email make sense. (every time/never)
BAR1	Being on the alert for unsafe emails is time consuming. (agree/disagree)
BAR2	The expense of being on the alert for unsafe emails is a concern for me. (agree/disagree)
BAR3	Being on the alert for unsafe emails would require changing my email habits, which is difficult. (agree/disagree)
BAR4	Being on the alert for unsafe emails would require substantial investment in effort other than time. (agree/disagree)
EFF1	I am confident I can recognize unsafe emails. (agree/disagree)
EFF2	I am confident I can recognize unsafe email attachments. (agree/disagree)
EFF3	I am confident I can recognize unsafe links in emails. (agree/disagree)
EFF4	I can recognize unsafe emails even if no one was around to help me. (agree/disagree)
CUE1	If I saw a news report or read a newspaper or magazine article about a crime related to unsafe emails, I would be more concerned about opening or clicking links within emails. (agree/disagree)
CUE2	If a friend were to tell me of a recent experience with identity theft related to a suspicious email, I would be more conscious of opening emails or clicking links within emails. (agree/disagree)
CUE3	If my computer started behaving strangely, I would be concerned I had improperly handled unsafe emails. (agree/disagree)
CUE4	If I became a victim of identity theft, I would be concerned I had improperly handled unsafe emails. (agree/disagree)
CUE5	If I received an email from the HelpDesk of my university about risks posed by unsafe emails, I would be more concerned about opening emails or clicking links within emails. (agree/disagree)
EXP1	How often do you receive unsafe emails in your inbox(es)? (daily/never)
EXP2	How frequently have you been affected by unsafe emails? (daily/never)
EXP3	How recently have you been affected by unsafe emails? (in the last week/never)
EXP4	The level of impact I have experienced due to receiving unsafe emails is (major impact/no impact)
VUL1	There is a good chance that I will receive an unsafe email. (agree/disagree)
VUL2	There is a good chance I will receive an email with an unsafe email attachment. (agree/disagree)
VUL3	There is a good chance I will receive an email containing links to phishing sites. (agree/disagree)
BEN1	Being on the alert for unsafe emails is effective in preventing viruses from infecting my computer. (agree/disagree)
BEN2	Checking if the sender and subject make sense before opening an email is effective in preventing viruses from infecting my computer. (agree/disagree)
BEN3	Checking if the file name of the attachment makes sense before opening an email is effective in preventing viruses from infecting my computer. (agree/disagree)
BEN4	Exercising care before opening email attachments is effective in preventing viruses from infecting my computer. (agree/disagree)
BEN5	Exercising care before clicking on links in emails is effective in preventing viruses from infecting my computer. (agree/disagree)
SEV1	Having my computer infected by a virus as the result of unsafe email practices is a serious problem for me. (agree/disagree)
SEV2	Putting the school's network at risk because of unsafe email practices is a serious problem for me. (agree/disagree)
SEV3	If my computer is infected by a virus as the result of unsafe email practices, my daily work/schoolwork could be negatively affected. (agree/disagree)

Feasibility of Creating a Non-Profit and Non-Governmental Organization Cybersecurity Incident Dataset Repository Using OSINT

Stanley J. Mierzwa
smierzwa@kean.edu

Iassen Christov
hristovj@kean.edu

Center for Cybersecurity
Kean University
Union, New Jersey 07083, USA

Abstract

Organizations of all types are prone to cybersecurity and information security attacks. Non-Profit Organizations (NPOs) and Non-Governmental Organizations (NGOs) are not exempt from using information technology solutions and, thus, have been the recipient victims of cyber attackers. There exist many areas and venues where data are collected to report back annually on the status and numbers of cybersecurity attacks against many sectors of our society. The Department of Homeland Security (DHS) Cybersecurity and Infrastructure Security Agency (CISA) catalogs sixteen critical sectors that are considered vital to the United States. However, finding where the NPO and NGO community should reside with regard to a categorized sector is challenging. The cybersecurity incident data collected by many agencies does place a focus on the sixteen sectors. This effort and write-up will focus on the NPO and NGO communities and provide the process followed to create the data repository, categorization of attacks taxonomy, fields captured, outlets, and areas where data that is relevant to historical cybersecurity incidents in these types of agencies is available. In addition, the beginning of a running log and dataset for the NPO and NGO community will emerge to determine if this activity is feasible and can continue. A desired outcome for this effort is to make available a dataset that can be referenced by researchers, students, and leaders investigating cybersecurity risk management and analysis of the NPO and NGO sectors. In addition, this effort was started with the purpose of providing students from varied academic disciplines the opportunity to engage in practical pedagogical cybersecurity and cybercrime activity.

Keywords: Pedagogy; Cyber-attacks; Cybersecurity; Cybercrime; Non-Profit; Non-Governmental Organization

Recommended Citation: Mierzwa, S., Christov, I., (2024). Feasibility of Creating a Non-Profit and Non-Governmental Organization Cybersecurity Incident Dataset Repository Using OSINT . *Cybersecurity Pedagogy and Practice Journal*; v3 (n2) pp 48-57. <https://doi.org/10.62273/EUFO3601>

Feasibility of Creating a Non-Profit and Non-Governmental Organization Cybersecurity Incident Dataset Repository Using OSINT

Stanley Mierzwa and Iassen Christov

1. INTRODUCTION

There exists an unmet need or gap with regard to cybersecurity and cybercrime incident historical data from attacks against the non-profit and non-governmental organization arenas or communities. The role of sharing data is instrumental in cybersecurity research and will benefit organizations aiming to predict and protect against malicious attacks (Kouper & Stone, 2024). Understanding that organizations may opt not to disclose if and when a cyber-attack strikes them, open source avenues still exist where relevant information related to such breaches and attacks against the NPO and NGO sectors has been made available in the public domain. The Non-Profit Cyber Incident Repository (NPCIR), created as part of this activity, partnering with students, faculty, and staff, allows individuals the opportunity to study, refer, and distill a variety of data point insights related to cybersecurity attacks. The qualities or data elements provided include the country or region of the affected NPO or NGO, the general type of attack, the type of industry-focus organization, relation to the Confidentiality-Integrity-Availability (CIA) triad, the year of the attack, and many other fields.

Having historical data on previous cybersecurity attacks against NPOs and NGOs can be a valuable asset that can be referred to by critical stakeholders in these organizations so that they can better prepare to defend against upcoming attacks. Additionally, understanding the landscape of what types, frequency, and other critical cyber-attack qualities are utilized against these outfits may be helpful for NPOs and NGOs that wish to place greater emphasis and obtain approval to align more significant cybersecurity resources.

2. THEORETICAL MODEL AND FRAMEWORK

To guide the effort to envision, initiate, build, and pursue this novel research activity, the use of Open Source Theory was followed as a framework. The theory was introduced as a form of human behavior, generating the capabilities and potential for sharing information early and

frequently, given the fast emergence and phenomenon of the Internet (Glassman, 2013). Key pillars followed by this relatively new theory include the development of new concepts, using tools commonly available to the masses, creation through practice, co-development of ideas, and focusing on a problem to solve (Glassman, 2013). Other studies have included the Open Source Theory in practice, including an investigation of this evolving framework (Choi & Glassman, 2017; Kim et al., 2015; Kuznetcova & Glassman, 2018).

3. LITERATURE REVIEW

Non-Profit and Non-Governmental Organizations

Understanding what constitutes a non-profit or non-governmental organization requires a brief definition. Although there may not be one wholly agreed-upon set of definitions, in considering the operation of a non-profit organization, it will consist of five essential characteristics, including being institutionalized as an organization, existing privately and separate from government, not returning any profits and being non-profit distributing, being of a self-governed model and configured with its internal governance, and in some degree allow for voluntary participation (Morris, 2000; Salamon & Anheier, 1992). Non-profits are more prone to not having ample budgets and resources to help safeguard their information technology and systems assets. They are considered a primary concern for their leadership and critical stakeholders (Ermicio & Liu, 2022).

Non-Governmental Organizations can be involved in a variety of missions or focus activities. For one example, the role of NGOs has been formidable in the development and delivery of vaccines and vaccine development (Walton, 2017). Historically, the nomenclature or non-governmental organization term generally took shape during the time of the emergence of the United Nations, during the mid-1940s (Willets, 2002). Many different mission-focused and variably structured NGOs exist in different parts of the globe and of varying sizes. This global landscape can introduce unique cyber challenges. As a form, the organization should be

independent of government management or control, without the aim to challenge governments as a political party, not for profit, and noncriminal (Willetts, 2002).

Existing Cybersecurity Incident Datasets

Known cybersecurity incidents have continued to rise substantially in organizations of many types, and this may only be tipping the case since not all breaches may be reported. Related to the NPO and NGO communities, an estimated 80% of nation-state attacks were directed against the broad swath of think tanks, government agencies, and NGOs (Lambert, 2021; Sobers, 2024). Over time, many different publicly accessible and available reports and datasets have emerged and, in some cases, continue to be routinely updated – meaning, it is not just a point-in-time snapshot of cyber incident attack information. The Critical Infrastructure Ransomware Attacks (CIRA) dataset and repository include records of attacks that are categorized as ransomware, and this information collection was initiated in 2019 (Rege, 2023). As of the time of this research, the CIRA database contains 1,654 ransomware attacks (Rege, 2023).

One other very broad-based, interactive, searchable cybersecurity event database is the CISSM Cyber Attacks Database, which is routinely updated (Harry & Gallagher, 2018). The solution permits quickly searching cyber-attack information with category types of actor type, attack type, and location. Another specific resource made available by the CommunityIT Innovators (2023) which is a non-profit organization that provides technology expertise to the NPO community. The resource provides a yearly snapshot view of NPO cybersecurity incident information within their client base (2023 Community IT Nonprofit Incident Report, 2023).

The United States Federal Bureau of Investigation (FBI) makes available and routinely provides reports via its Internet Crime Complaint Center (IC3 - Internet Crime Complaint Center, 2022). The IC3 is seen as the nation's central hub for providing the ability to report cybercrimes and review historically collected data via the availability of an annual report. The IC3 system does not provide the details and data to perform a deep dive into providing individual company or organization names. However, it does report on the ranking of the most frequent attack types and trends, along with an overarching landscape view of the previous year's internet breaches and cybercrimes reported into their portal.

Cybersecurity Focus on the NPO and NGO Environment

Relevant and excellent resources exist to aid the NPO and NGO sectors with activities, tasks, and methods to prevent cybersecurity attacks. Specific to the global public health NPO and NGO environments, Mierzwa et al. (2020) outlined ways these sectors could integrate strategies to minimize cyber risks. An abundance of cybersecurity frameworks exists in both the public domain and private (requiring purchase of subscription) domains. For some small businesses, which smaller or start-up NPOs may be categorized, the use of the freely available public domain genre of frameworks can be evaluated (Mierzwa & Klepacka, 2023).

Lack of Cybersecurity Incident Reporting Requirements on the NPO and NGO Community

In 2022, President Biden of the United States signed into law an act named the Cyber Incident Reporting for Critical Infrastructure Act of 2022 (CIRCIA). The act requires that organizations considered to be covered entities report to the United States Cybersecurity and Infrastructure Security Agency (CISA) when cybersecurity incidents or ransomware payments are made (CISA.GOV, 2022a). CISA is tasked with developing and issuing the regulations, which are expected to require covered entities to report cyber incidents within 72 hours from the time a cybersecurity incident occurs (CISA.GOV, 2022b). The development of CIRCIA is yet another notice requiring agencies and organizations to report cybersecurity incidents of notice potential. Thus, this research on the subject will further contribute to the knowledge and information available on this topic. In its annual Data Breach Investigations Report, Verizon states that incident reporting can be influenced by factors such as reporting laws and partner visibility, and some industries, such as entertainment or construction, may have low numbers reported because of a lack of requirements (Verizon, 2022). However, the sectors of NPOs and NGOs may not necessarily sit within any of the known entities and, as such, may not be required to report cyber-incidents.

Recent bipartisan legislation has emerged during 2022 that mandates the reporting of cyber incidents for specific critical infrastructure sectors of our nation, and the Cybersecurity and Infrastructure Security Agency (CISA) has been provided the opportunity to document the rule mechanics (Friedman & Mitchell, 2022). The importance of reporting can include many factors, but one essential product is to permit the timely

sharing of cyber threat activities that can benefit our critical infrastructure sectors.

Specific sectors, such as healthcare, are certainly affected by cyber-attacks but are widely under-reported. Specific reasons many health-related organizations under-report cyber incidents include not knowing what to report and how to do so (Cyber Peace Institute, 2021). This is despite an increase in such attacks as ransomware, which has doubled between 2016 and 2021, creating disruptions or inability to use healthcare systems and even canceling scheduled care (Neprash et al., 2022). Even if considered to be possibly underestimated, the Neprash et al. (2022) study also found that one in five ransomware attacks were not reported or available in the US Department of Health and Human Services Office of Civil Rights (HHS OCR) Data Breach Portal (Neprash & Rozenshtein, 2023).

4. METHODS

Data Collection for Repository

The ability to gain deeper knowledge and insight on a topic has become more accessible than ever and with incredible speed with the use of openly and freely available resources and information as a result of the growth of the Internet, specifically within the World Wide Web. Some of the earliest forms of essential Open Source Intelligence (OSINT) can be attributed to the Second World War, with the advent of investigations of free and available data and information and connecting with findings to incorporate actions (Glassman & Kang, 2012; Schaurer & Jorger, 2010). With a growing source of information publicly available, OSINT has taken on a greater urgency with the increases in nodes of information sources that can be unrestricted via the use of the surface and deep web (Glassman & Kang, 2012).

The introduction of OSINT to undergraduate and graduate students pursuing the information technology, cybersecurity, or criminal justice fields provides an excellent opportunity to engage with practical training. For those students pursuing law enforcement, the activity of OSINT can be a powerful approach to training future investigators (Larsen et al., 2023). OSINT techniques can include searching resources found in the surface web, deep web, social media, image and reverse-image, and mapping searches (Larsen et al., 2023).

The decision to not explore or search for relevant breach, attack, or hacking information on the Dark Web relevant to the NPO and NGO sectors was made purposely. The Dark Web may have

ample amounts of relevant and unpublished information related to hacking that has transpired. However, the Dark Web also contains illicit information, knowledge, skills, and data related to hacking, including breaching accounts, such as social media and websites, access to malware, and steps to perform such attacks as Denial-of-service (Choi & Lee, 2022). This research activity includes academics, professionals, and students, and as such, by not entertaining the idea of scouring the Dark Web for open-source information on cyber-attacks, an ethical step is being demonstrated to students and users of this finalized dataset regarding using more formal published materials.

Alert Source	Purpose
Google Alert	"NGO" and "breach"; "NGO" and "hacking"; "non-profit" and "breach"; "non-profit" and "hacking"; "nonprofit" and "breach"; "nonprofit" and "hacking"; "relief agency" and "breach"; "humanitarian" and "breach"; "humanitarian" and "hacking"; "non-governmental" and "breach"; "non-governmental" and "hacking"; "charity" and "hacking"; "charity" and "cyber-attack"
INOREADER	"non-profit cyber-attack"; "non-profit cybercrime"; "non-profit breach"; "non-governmental organization cyber-attack"

Table 1: Automated Alert Terms Utilized

For this research activity in creating an available dataset, publicly available sources were utilized to search, discover, analyze, and determine if an entry was viable for inclusion in the database. The critical sources included Google Search, Google Scholar, INOREADER News Feed Creator, the digital library repositories of publications at the Kean University Digital Library, and the University of Cumberland's Digital Library. The use of the Google Alerts feature and function as a mechanism to be automatically notified of new publicly available information relevant to this

specific research effort was created to automate the process. The digital alerts were configured using key search terms outlined in Table 1. It is expected that as more sources are identified to contribute further to this research, the repositories will be added.

Initiation and Creation of an NPO and NGO Cyber Incident Reporting Dataset

The creation of the dataset was inspired by the lack of such a resource available in the public domain. Many sources of information related to cybersecurity and cybercrime incidents can be found via such resources as the FBI Internet Crime Complaint Center (IC3.GOV) web portal. Additionally, focused datasets related to specific types of cybersecurity incidents are available, as well as snapshots of activities in a variety of reports over the past many years. However, having a place for NPO and NGO organizations to reference historical records of attacks was found to be lacking. Hence, the vision of this instantiation of a data repository is to continue to grow and add value to these sectors. The expected importance of such data to quickly reference can be referred to by the NPO and NGO sector, as well as by those organizations that assess or audit this sector and by third-party partners that provide technology security services to protect them.

Dataset Collection Fields

The collection of historical data commenced in December 2023 and continues to the time of this writing. It is expected to continue going forward, with a different cohort of students trained and integrated each semester as new interns or academic assistants assigned to the Kean Center for Cybersecurity. As information was gained related to cybersecurity incidents that pertained to or affected non-profit and non-governmental organizations, the data was added to a secure web-based data portal. The initial set of fields utilized when recording the information found in the open source or publicly available avenues include the items found in Table 2. In addition to general demographics, fields related to the cyber incident were captured, and attempts to align with the 16 critical DHS sectors were made, as well as connections to TAG Infosphere’s Cyber taxonomy (TAG Infosphere, 2024). The TAG Cyber taxonomy includes an organized list of significant categories aligned with cybersecurity approaches (TAG Infosphere, 2024). Additionally, we leverage and make a correlation between cyber-attacks and the Confidentiality-Integrity-Availability (CIA) triad. The CIA triad has helped define information system security in literature. It provides definitions of the three pillars that

organizations can follow in the effort to safeguard their assets and technology infrastructure (Shiou et al., 2023). In relation to our categorizing which of the CIA triad elements relate to the NPO or NGO breach or cyber-attack, an assessment was made by the investigative team to make the category assignment.

Dataset Field	Purpose
Contributor First Name	Record contributor name.
Contributor Last Name	Record contributor name.
Contributor Email	Record contributor email.
Contributor Organization	Record contributor organization.
Incident Unique Identifier	We assigned a unique identifier.
Date of Incident	Date or approximate/estimated date of the incident.
Year	Year of incident – YYYY format.
Non-Profit/NGO Name	Name of NPO or NGO.
Non-Profit/NGO Country(s)	Country location(s) of NPO or NGO breached.
DHS CISA Critical Sector Targeted (if applicable)	Organization category alignment or included in one of the sixteen DHS CISA Critical Sectors.
Outside Sector (if applicable)	Organization category if outside of DHS CISA.
CVE Targeted (if available)	Targeted CVE (if applicable and available).
Attack Type	Type or genre of cyber-attack.
CIA Triad Element Affected	Recording of whether the Confidentiality, Integrity, or Availability (CIA) pillar was affected.
Location of OSINT Knowledge Source	Link or location of open-source intelligence article or information evident from the attack.

Dataset Field	Purpose
TAG Cyber Taxonomy	Record which category of the TAG Cyber reference taxonomy was affected or associated and included in the attack.
Upload of artifact about the incident	Additional artifact or evidence of attack.
Other Non-Profits/NGOs Targeted	Recording of other NPOs or NGOs affected by this attack.
Additional Information	Other miscellaneous information about the attack.

Table 2: NPCIR Dataset Fields (Codebook)

The process followed to discover and catalog the NPO or NGO cyber incident record included several steps. After first enabling and establishing automated alerts to when such information was newly available, the research team also did manual historical searches through the select information portals. Once discovered, a thorough read of the open source and publicly available articles or evidence was made, and it was determined whether the found information was relevant for inclusion in the dataset. Prior to adding the record to the dataset, members of the team manually considered each data element as outlined in Table 2 prior to addition to the repository. As part of the evaluation, considerations were made as to whether the attack was limited to the organization itself or other NPOs and NGOs.

Access to the resulting dataset is made available to individuals and companies after making a formal request via completing a web-based inquiry form. Upon receipt of the completed request form, members of the research and investigative team analyze the use case for utilizing the data repository. The repository is to be used for research defensive investigation, academic research, or potentially by the press if found useful given a specific applicable news story. The dataset is not permitted to be used to charge a fee or for any profit-making enterprise. Additionally, as the dataset continues to expand, it will be stamped with an increasing version label number and made available upon request and approval at least two to three times per year and provided in a PDF report.

5. RESULTS

In this first incarnation (version 1.0), the resulting data is presented after seven months of open-source intelligence gathering to determine the feasibility of the NPCIR data repository. It is anticipated that after each instantiation of the data repository, the dataset will grow, and it is expected that at least two versions with the most updated results included will be generated annually and provided for open access. A total of $N = 168$ cybersecurity incidents have been discovered thus far focused on the NPO and NGO organizations. The recorded year with the most documented and discovered attacks was 2023, and the earliest documented was 2011. The most attacked NPO and NGO country was the United States, with an abundant number of countries being recorded. The CIA pillars most affected in the recorded attacks, which could include multiple pillars, included (1) Confidentiality at 81%, (2) Integrity at 49%, and (3) Availability at 41%. The most targeted or DHS CISA-aligned sector affected by cyber-attacks included those from the Healthcare and Public Health sectors, reporting 45.5% of all incidents discovered. The resulting summarized data can be found in the following Figures: 1, 2, 3, 4, and 5 with additional descriptive statistics.

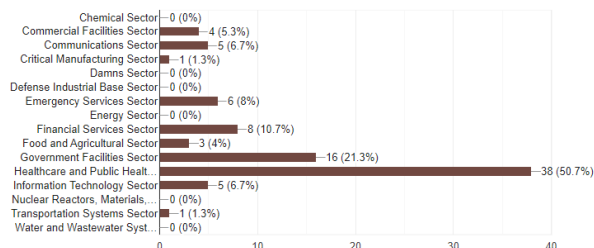


Figure 1 DHS CISA Targeted Sectors

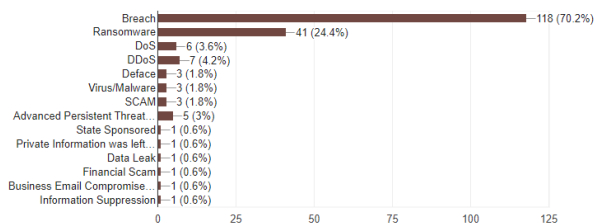


Figure 2 Attack Type

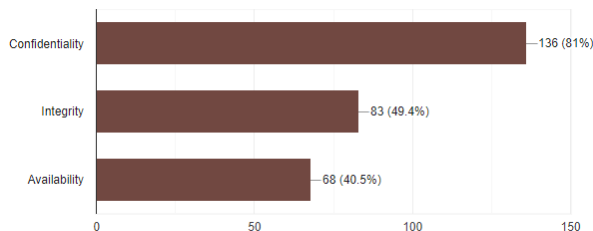


Figure 3 CIA Affected

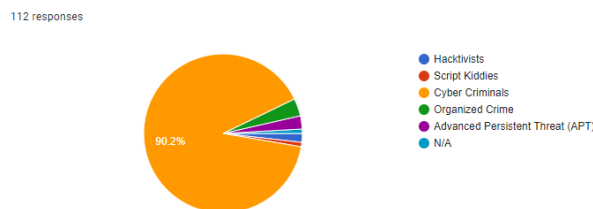


Figure 4 Threat Actor Type

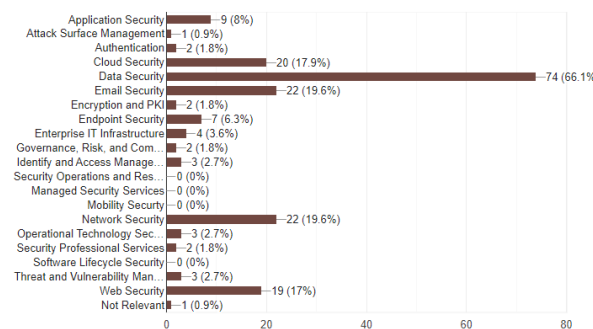


Figure 5 TAG Cyber Reference Taxonomy

One of the descriptive findings was that the information collected and discovered regarding non-profits that have been hacked was more likely to be categorized as “breached” as a hacking approach in the attack, resulting in over 70% of the discovered attacks. This finding supports the previous finding of a quantitative study that resulted in a significantly higher number of breaches due to unauthorized access to non-profits via breaches, and for-profits experienced breaches in higher volume due to theft (Ignatovski, 2023). The next most common attack type was found to be “ransomware” at 24% (Ignatovski, 2023).

6. DISCUSSION AND IMPLICATIONS

In relation to operating an organization, NPOs and NGOs are no different than most for-profit businesses. Organizations with an NPO or business mission will aim toward being efficient, innovative, and effective by considering the use of information technology solutions in their

operations. Just the same as larger corporations, the NPO and NGO environment will be involved in generating technology solutions and using modern cloud-based solutions, such as Microsoft Azure (Souidi et al., 2015). Additionally, as part of the data collection research efforts in NPOs and NGOs, innovative self-report survey systems or other novel, groundbreaking solutions may be offered (Falb et al., 2016; Mierzwa et al., 2016; Savel et al., 2014). In essence, every time a new information technology or web-based or Internet-connected product is provisioned, there will be the potential for cybersecurity attacks.

Several essential learning items emerged as a result of this exercise. In one case, many NPOs and NGOs may not necessarily align with the CISA's sixteen critical infrastructure areas, and as a result, a field to categorize areas outside the sectors was added. As this activity and research continue, these critical areas outside the CISA sector may grow. At this time, the additional NPO- or NGO-focused categories include Humanitarian Aid, Education, Food Security, Support Democracy, Political Organizations, Scientific Services, Technical Services, and Arts/Entertainment Services. For students who may be unsure of the relevance of cybersecurity to varied sectors of industry and focus, this activity provided excitement for them in understanding the practicality of this approachable activity. It is envisioned that a new set of interns will be assigned each semester to carry this effort through a sustained model. Without providing routine attention to this OSINT activity, there is the potential to no longer be viable.

7. CONCLUSION

It can be overwhelming for a cybersecurity student, researcher, incident responder, cyber risk manager, or defensive technical cyber expert to zero in on relevant information affected in the specific sector one operates. With cybersecurity and cybercrime being very fast-moving targets, having valuable resources that are more attuned to one's industry or sector can be an appreciated asset in information security and cybersecurity awareness. The NPO and NGO environment has often relied on general-purpose cybersecurity repositories to help them contend with the aftermath of attacks and defend against upcoming potential attacks. Cybersecurity analysts and researchers will rely on mixed or more than one dataset when aiming to protect their organizations from attacks, and this activity helps to draw attention to those in the NPO and NGO communities.

This first version and initial student-focused activity and research were started to help students involved in a cybersecurity internship. The students were from a variety of disciplines, including information technology, computer science, and criminal justice. This practical investigation activity has helped the students to grasp better and understand the landscape of research for the purpose of preventing cybersecurity and cybercrime attacks. The knowledge, skills, and abilities provided to the students included understanding the DHS CISA sector categories, introduction to OSINT, tracking and tracing threat intelligence, practical assessment of the CIA triad, concepts of security ethics, and using nontechnical skills of collaboratively working on a tangible project. This effort was found to be feasible and will continue to be pursued with future students.

8. LIMITATIONS

The discovered and added cybersecurity incidents in the NPO and NGO sectors or communities were, for the most part, done manually by humans, with the exception of automated alerts. This is a limitation and demonstrates that the process may be inefficient given the steps necessary. This human-level interaction creates a limitation, and perhaps future evolutions to the product could be completed with artificial intelligence bots or automated scripts. Additionally, there is the potential that previous evidence of cyber-attacks against the NPO or NGO communities was removed from the public domain or surface web and not able to be cataloged, which would not appear in this repository dataset.

9. FUNDING

The author (s) received no financial support during the course and activity of this research and publication.

10. ACKNOWLEDGEMENTS

Part of the inspiration for pursuing this activity was the interest of students who were assigned to the Kean Center for Cybersecurity as part of an internship research project. Heartly thanks to students Heni Patel, Gurdeep Singh and Caitlin Chiodo, for their involvement in the practical cybersecurity exercise and research. This outreach activity is in favor and support of the National Security Agency Center of Academic Excellence Cybersecurity Defense (NSA CAE-CD) outreach action. Such outreach and student-faculty involvement activities are advocated and expected in yearly efforts. The authors are

genuinely grateful to the Kean University Office of Career Services for their excellent outreach to partner students with faculty at the university that align with student interests.

11. REFERENCES

- Choi, M., & Glassman, M. (2017). What it means to be a citizen in the internet age: Development of a reliable and valid digital citizenship scale. *Computers & Education, 107*, 100-112. <https://doi.org/10.1016/j.compedu.2017.01.002>
- Choi, K. S., & Lee, C. S. (2022). In the Name of Dark Web Justice: A Crime Script Analysis of Hacking Services and the Underground Justice System. *Journal of Contemporary Criminal Justice, 39*(2), 201-221. <https://doi.org/10.1177/10439862231157520>
- CISA.GOV. (2022a). CISA Central Reporting Operations Guide. As retrieved on December 26, 2022, from: https://www.cisa.gov/sites/default/files/publications/CISA_Central_Operations_Branch_Slick%20Sheet_508c.pdf
- CISA.GOV. (2022b). Cyber Incident Reporting for Critical Infrastructure Act of 2022 (CIRCIA) Fact Sheet. As retrieved on December 23, 2022, from: https://www.cisa.gov/sites/default/files/publications/CIRCIA_07.21.2022_Factsheet_FINAL_508%20c.pdf
- CommunityIT Innovators. (2023). *2023 Nonprofit cybersecurity incident report*. <https://communityit.com/2023-nonprofit-cybersecurity-incident-report/>
- Cyber Peace Institute. (2021). *Playing with lives: Cyberattacks on healthcare are attacks on people*. <https://cyberpeaceinstitute.org/report/2021-03-CyberPeaceInstitute-SAR001-Healthcare.pdf>
- Ermicioi, N., & Liu, M. X. (2022). Cybersecurity in nonprofits: Factors affecting security readiness during Covid-19. *SAIS 2022 Proceedings, 18*. <https://aisel.aisnet.org/sais2022/18>
- Falb, K., Tanner, S., Asghar, K., Souidi, S., Mierzwa, S., Assaznew, A., & Stark, L. (2016). Implementation of Audio-Computer Assisted Self-Interview (ACASI) among adolescent girls in humanitarian settings: feasibility, acceptability, and lessons learned. *Conflict and Health, 10*(1), 1-8.

- Friedman, S., & Mitchell, C. (2022). Former Federal CISO calls for early CISA workshop as industry leaders seek a place in shaping incident reporting rules. *Inside Cybersecurity*. Arlington. <https://doi.org/10.1016/j.fsidi.2023.301622>
- Glassman, M. (2013). Open source theory. *01. Theory & Psychology*, 23(5), 675-692. <https://doi.org/10.1177/0959354313495471>
- Glassman, M., & Kang, M. J. (2012). Intelligence in the internet age: The emergence and evolution of Open Source Intelligence (OSINT). *Computers in Human Behavior*, 28(2), 673-682. <https://doi.org/10.1016/j.chb.2011.11.014>
- Harry, C., & Gallagher, N. (2018). Classifying cyber events. *Journal of Information Warfare*, 17(3), 17-31.
- Ignatovski M. (2023). For-profit versus non-profit cybersecurity posture: breach types and locations in healthcare organizations. *Health Information Management Journal*, 0(0), 1-8. doi:10.1177/18333583231158886
- Internet Crime Complaint Center. (2022). *Federal Bureau of Investigation Internet Crime Report 2021*.
- Kim, Y., Glassman, M., & Williams, M. S. (2015). Connecting agents: Engagement and motivation in online collaboration. *Computers in Human Behavior*, 49, 333-342. <https://doi.org/10.1016/j.chb.2015.03.015>
- Kouper, I., & Stone, S. (2024). Data sharing and use in cybersecurity research. *Data Science Journal*, 23(3), 1-19. <https://doi.org/10.534/dsj-2024-003>
- Kuznetcova, I., & Glassman, M. (2018). Rethinking the use of multi-user virtual environments in education. *Technology, Pedagogy and Education*, 29(4), 389-405. <https://doi.org/10.1080/1475939x.2020.1768141>
- Lambert, J. (2021). Microsoft Digital Defense Report. Retrieved on March 6, 2024 from: <https://www.microsoft.com/en-us/security/blog/2021/10/25/microsoft-digital-defense-report-shares-new-insights-on-nation-state-attacks/>
- Larsen, O. H., Ngo, H. Q., & Le-Khac, N. A. (2023). A quantitative study of the law enforcement in using open source intelligence techniques through undergraduate practical training. *Forensic Science International: Digital Investigation*, 47, 1-11. <https://doi.org/10.1016/j.fsidi.2023.301622>
- Mierzwa, S. J., & Klepacka, A. (2023). Practical Approaches and Guidance to Small Business Organization Cyber Risk and Threat Assessments. *Journal of Strategic Innovation and Sustainability*, 18(2). <https://doi.org/10.33423/jsis.v18i2.6255>
- Mierzwa, S., RamaRao, S., Yun, J. A., & Jeong, B. G. (2020). Proposal for the Development and Addition of a Cybersecurity Assessment Section into Technology Involving Global Public Health. *International Journal of Cybersecurity Intelligence & Cybercrime*, 3(2), 48-61. <https://www.doi.org/10.52306/03020420BA BW2272>
- Mierzwa, S., Souidi, S., & Savel, C. (2016). On selecting an appropriate customizable electronic self-report research technology. *Procedia Engineering*, 159, 66-71. <https://doi.org/10.1016/j.proeng.2016.08.065>
- Morris, S. (2000). Defining the nonprofit sector: Some lessons from history. *International Journal of Voluntary and Nonprofit Organizations*, 11(1), 25-43.
- Neprash, H., McGlave, C. C., Cross, D. A., Vimig, B. A., Puskarich, M. A., Huling, J. D., Rozenshtein, A. Z., & Nikpay, S. S. (2022). Trends in Ransomware Attacks on US Hospitals, Clinics, and Other Health Care Delivery Organizations, 2016-2021. *JAMA Health Forum*. 3(12). doi: 10.1001/jamahealthforum.2022.4873.
- Neprash, H., & Rozenshtein A. Z., (2023). New Data Quantifies Ransomware Attack on Healthcare Providers. *Lawfare. Institute in Cooperation with Brookings*. As retrieved on January 10, 2023, from: <https://www.lawfareblog.com/new-data-quantifies-ransomware-attacks-healthcare-providers>
- Rege, A. (2023). "Critical Infrastructure Ransomware Attacks (CIRA) Dataset". Version 12.9. Temple University. Online at <https://sites.temple.edu/care/cira/>. ORCID: 0000-0002-6396-1066.
- Salamon, L. M., & Anheier, H. K. (1992). In search of the non-profit sector: The question of definitions. *Voluntas*, 3, 125-151.
- Savel, C., Mierzwa, S., Gorbach, P., Lally, M., Zimet, G., Meyer, K., Souidi, S., & Adolescent Trials Network for HIV, & Interventions, A.

- (2014). Web-based, mobile-device friendly, self-report survey system incorporating avatars and gaming console techniques. *Online journal of public health informatics*, 6(2), e191. <https://doi.org/10.5210/ojphi.v6i2.5347>
- Shiou, W. L., Wang, X., & Zheng, F. (2023). What are the trends and core knowledge of information security? A citation and co-citation analysis. *Information & Management*, 60(3), 1-21. <https://doi.org/10.1016/j.im.2023.103774>
- Sobers, R. (2024). 161 Cybersecurity statistics and trends [updated 2023]. Varonis. Retrieved on March 6, 2023, from: <https://www.varonis.com/blog/cybersecurity-statistics>
- Souidi, S., Boccio, D., Mierzwa, S., & Aguilar, J. (2015). The feasibility of using Microsoft Azure infrastructure for a monitoring and evaluation system solution in Sub-Saharan Africa. *IEEE Global Humanitarian Technology Conference*. IEEE, 226-232.
- TAG Infosphere. (2024). TAG Cyber Taxonomy. As retrieved on February 28, 2024, from: <https://tag-infosphere.com/service/cybersecurity/taxonomy>
- Verizon. (2022). Verizon Data Breach Investigations Report 2008-2022. A retrieved on December 6, 2022, from: <https://www.verizon.com/business/resources/T34a/reports/dbir/2022-data-breach-investigations-report-dbir.pdf>
- Walton, J. (2017). The role of non-governmental organizations in vaccine development and delivery. *International Journal of Health Governance*, 22(3), 152-160. <https://10.1108/IJHG-02-2017-0006>
- Willets, P. (2002). What is a non-governmental organization? *Conventions, treaties and other responses to global issues*, 2(11), 229-248