12-15-2019

Identifying potentially risky insider on-compliance using machine learning to assess multiple protection motivation behaviors

Michael Curry
Oregon State University, michael.curry@oregonstate.edu

Byron Marshall
Oregon State University

Robert E. Crossler
Washington State University

IDENTIFYING POTENTIALLY RISKY INSIDER NON-COMPLIANCE USING MACHINE LEARNING TO ASSESS MULTIPLE PROTECTION MOTIVATION BEHAVIORS

Michael Curry[1]
College of Engineering, Oregon State University,
Corvallis, Oregon, USA

Byron Marshall
College of Business, Oregon State University,
Corvallis, Oregon, USA

Robert E. Crossler
Carson College of Business, Washington State University,
Pullman, Washington, USA

ABSTRACT

Cybersecurity researchers have made significant steps to understand the mechanisms of security policy compliance and unify theories of security behavior. However, due partly to the limitations of traditional variance model statistical methods, these studies by necessity typically focus on a single security policy issue. By contrast, new machine learning algorithms frequently employed by data scientists offer great promise as a new statistical approach for examining robust individualized interpretations of policy and can also identify potentially risky behaviors. This study proposes to explore cybersecurity training impediments of multiple protection motivation behaviors in ransomware prevention training. It demonstrates the feasibility of using machine learning with survey items from the cybersecurity research to predict non-compliance. It also illustrates a potentially novel method to statistically validate research theory through higher levels of ML prediction. This study is a work in progress and we seek feedback on its design and relevance.

[1] Corresponding author. michael.curry@oregonstate.edu

Keywords: Machine learning, policy non-compliance, cybersecurity

INTRODUCTION

Ransomware is currently the most significant organizational malware threat today (Verizon Enterprise Solutions 2018), suggesting that changing insider behavior to prevent multi-vector threats is a pressing issue in cybersecurity. In response to such threats, cybersecurity professionals frequently update recommended best practices for security behavior (e.g. saving data according to retention rules, verifying legitimate emails, and reporting incidents). When users do not comply with recommendations it makes them less safe and puts the organization at a higher security risk. Consequently high participation rates in cybersecurity training to communicate best security practices is important to protect against sophisticated multi-vector attacks such as ransomware (D'Arcy et al. 2009; Jensen et al. 2017).

Significant steps have been made to understand the mechanisms of cybersecurity policy compliance (S. Boss et al. 2015; e.g. , Bulgurcu et al. 2010; Crossler and Belanger 2014; Curry et al. 2019; Johnston and Warkentin 2010; Siponen et al. 2010). These studies generally assume a universally applicable security policy rules matched to correct security behaviors. Thus, one common approach to persuading insiders' participation in cybersecurity training is the use of fear appeals.

Cybersecurity research has begun to unify theories of security behavior (e.g. Moody et al. 2018) using familiar statistical methods where inputs are generally tangible and empirically observable inputs. These and other studies have used traditional variance modeling to validate behavior-driving constructs assessed through psychometric surveys. By implication, individuals who have low levels of these identified constructs are more likely to be non-compliers. However,

machine learning (ML) classifier algorithms employed by data scientists offer great promise as a new statistical approach to combining theoretical constructs. Furthermore, the use of survey items as inputs to ML is not uncommon in practice as organizations use opinion surveys or transaction feedback to guide response protocols. But the use of psychometric items in assessing or identifying cybersecurity risk in machine learning has not been explored in the research literature. Therefore the research question we seek to answer is:

RQ: Are items from behavioral information security research practically useful in an ML context as indicators of the likelihood of non-compliance with recommended security practices?

This study adopts the lens of Posey et al. (2013) who define three dimensions for security topics to help identify impediments to cybersecurity training, and explores the feasibility of the use of machine learning (ML) classifier algorithms to process psychometric survey response items in identifying individuals who are likely to pose elevated levels of non-compliance risk. We seek to develop a new approach for employing theoretical models that combine behavioral information security with data science to produce actionable data relationships for improved cybersecurity behavior.

BACKGROUND

Posey et al. (2013) Identify a taxonomy of protection-motivation behaviors (PMB), which they define as 'volitional insider behavior to protect organizational information and systems from security threats.' These serve as a broad categorization of representative clusters for classifying responsible security behaviors. We itemize six behaviors noted by security professionals as targeting ransomware vectors and their PMB classification (Posey et al. 2013) in Table 1.

Table 1. Potential Ransomware Prevention Protective-Motivation Behaviors (PMB)

Topic	Attack vector targeted	PMB (Category) from Posey et al.

		(2013)
Legitimate and Secure e-mail use	Phishing email messages with a malicious attachment or a link to a payload hosted elsewhere are a primary infection source	An organizational insider only responds to emails which have a legitimate business request (1) An organizational insider opens email attachments only if he/she knows the email's sender and was expecting the email (1) An organizational insider does not open emails that he/she believes have a chance of containing a virus or other potentially malicious components (6). An organizational insider pauses before responding to an email to make certain that he/she is responding to a valid request (6).
Secure Software	Secure use of computer software	An organizational insider immediately applies software updates to his/her computer workstation when notified of the update by an authorized individual or department within his/her organization (6)
Reports any incidents	A prompt organizational response can minimize the consequences of ransomware attack	If an organizational insider identifies something that looks out of the ordinary in his/her work environment, he/she immediately reports it to the proper organizational authorities (3)
Saves information according to retention policy	Appropriate storing of information per policy may support backups and limit sources of malware	An organizational insider stores information only according to the retention policies specified by his/her organization (3)
Changes password according to guidelines	Protecting account from becoming a source of spreading ransomware within the organization	An organizational insider changes his/her passwords according to his/her organization's security guidelines (3)
Backs up important data	Having backups enables recovery from ransomware attack	An organizational insider backs up important data and documents on a regular basis (4)
Mindfulness while working	Being alert and intentional while at work to avoid careless behavior that can lead to insecurity	An organizational insider works at a steady but cautious pace to ensure that he/she performs their job tasks in a secure manner (4)

Posey et al. (2013) also define three dimensions for classifying behavioral topics: the criticality or importance for all to perform on an ongoing basis; the difficulty promoting ongoing adherence

due to burden, focus and effort required; and finally the degree of common sense that leads insiders to naturally understand the logic and rationale. Assessing these three behavioral dimensions across all five PMBs can offer valuable insights on insiders security policy expectations (Siponen 2003) and provide clear targets for improving the most important behaviors for an organization. We posit that low levels of these three behavioral dimensions across all five PMBs may impact insiders understanding of recommended security behaviors, their motivation for following those practices and consequently their actual participation in cybersecurity training. We formalize these in the following hypotheses:

H1: Insiders' assessment of criticality or importance for all to perform a recommended PMB on an ongoing basis will impact motivation to adopt.

H2: Insiders' assessment of difficulty promoting ongoing adherence due to burden, focus and effort required to perform a recommended PMB will impact motivation to adopt.

H3: Insiders' assessment of degree of common sense that leads insiders to naturally understand the logic and rationale to perform a recommended PMB will impact motivation to adopt.

Adapting the Posey et al. (2013) taxonomy to the challenge of multi PMB initiatives, we propose to collect data from participants invited to participate in cybersecurity training and theoretically classify by ransomware vulnerability. Successful strategies for preventing a threat such as ransomware would be an important methodological contribution to cybersecurity professionals by addressing a complex multi-vector threat.

Machine Learning

The use of ensemble classifiers such as the random forest or C5.0 algorithms shows promise in extending cybersecurity research at the organizational level. But the internal logic

captured in the resulting models is largely opaque (Gopal et al. 2011). Because this relatively new approach is not widely understood, management may tend to view machine learning as a 'blackbox' causing difficulty in making a convincing business case that analytical conclusions are reliable and actionable (Müller et al. 2016; Ramasubramanian and Singh 2016; Sharma et al. 2014). By contrast, we posit that combining theoretically inspired indicators of cybersecurity behavior should offer satisfying indicators of likely noncompliance. There is no particular reason to believe that theoretically validated constructs do not interact in complex ways more easily discovered using machine learning than commonly liner-regression-based analysis. This notion is formalized in the following hypothesis.

H4: *Using psychometric cybersecurity indicators of PMB attitudes and intentionality will increase categorization accuracy as compared to explained variance based on regression analysis methods.*

Reducing assessment length by using the 'best items'

The large number of survey items in research studies often makes the proposed instrument impractical in an organizational setting. Feature selection techniques are commonly used to reduce computational cost and increase accuracy in machine-learning classification tasks involving a high number of indicator variables. Feature selection techniques promise to shorten surveys while retaining a large portion of indicative power. In the security domain, for example, Jenkins and Grimes (2014) employed feature selection for keystroke dynamics, e.g., how long keys are held down and time between key presses, to identify password reuse without actually recording passwords, and Liu and Yu (2005) note that feature selection plays an important role in intrusion detection. Ensemble classifiers may be capable of employing complex interactions that have not been carefully explored in the research literature. These interactions may be 'captured'

in the classification model even though the number of assessment items has been reduced using feature selection. We formalize this in the following hypothesis.

H5: The number of psychometric cybersecurity indicators employed in classification can be reduced without significantly reducing prediction accuracy.

METHODOLOGY

We propose to collected data from 250 participants, as summarized in Figure 1.

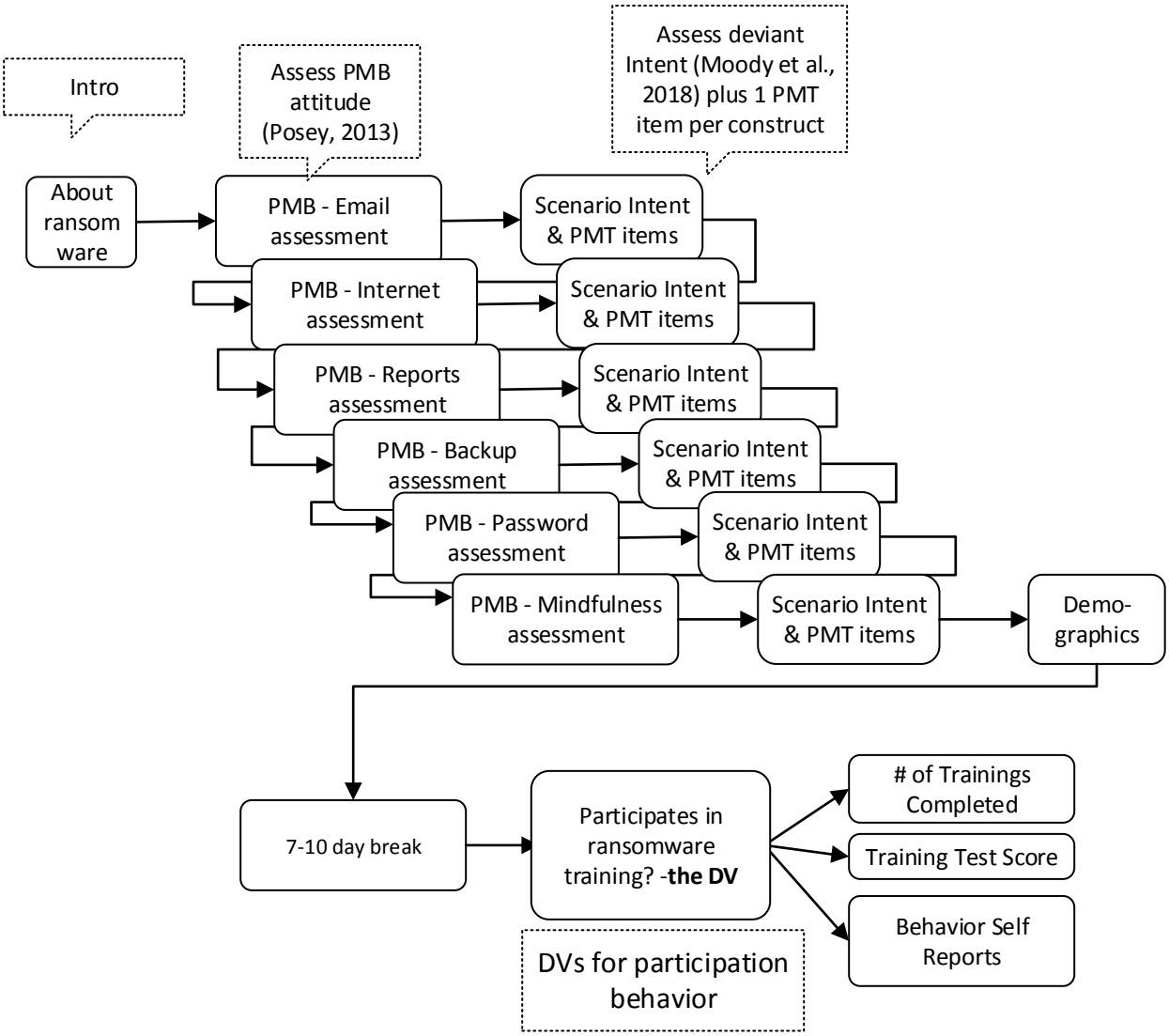

Figure 1. Study design to assess PMB attitudes, deviant intentions and ransomware training participation.

Participants are to be presented with a short motivational introduction about ransomware and the need for multiple behavior changes. This is followed by presenting six different recommended PMBs and assessments of attitudes towards each. To assess drivers of intention, a single protection motivation theory (PMT) (S. R. Boss et al. 2015) indicator for each construct will be asked over the six behaviors. This can be used to identify which PMBs insiders have difficulty with following and may also suggest necessary organizational improvements. We plan to confirm difficulties with the PMB by also assessing deviant intention towards this PMB using scenarios (Moody et al. 2018; Siponen et al. 2010). Finally we assess participation in the follow-on ransomware training as a dependent variable (see Appendix A) to classify each insider's threat.

Discussion

Guided partly by feature selection computations, this study will explore the ability of a ML classifier algorithm to categorize potentially risky individuals using a subset of the items used to assess theoretical drivers of noncompliance. This smaller instrument can also be used by an organization to evaluate the most problematic areas of security policy compliance, and identify risky individuals who may also be a higher security risk then target them for additional treatments or monitoring. This approach charts a path whereby items assessing constructs drawn from behavioral InfoSec studies may be practically employed as inputs to cybersecurity risk assessments. The model can also be reused on future data as insiders participate in the security training.

CONCLUSION

This work in progress seeks to fill the gap in exploring how insiders interpret cybersecurity policy by using psychometric indicators processed with powerful ML algorithms. This work can

contribute to the literature by exploring the feasibility of bringing together constructs and items developed using reliable and familiar theory development techniques with new computational algorithms. Theoretically validated items that are thought to indicate drivers of behavior may or may not translate well into practical indicators in a ML context. But if they do, it would have useful practical implications and may lead to new insights into how security behavior models are theorized and implemented in the future.

REFERENCES

Boss, S., Galletta, D., Lowry, P., and Moody, G. 2015. "What Do Users Have to Fear? Using Fear Appeals to Engender Threats and Fear That Motivate Protective Security Behaviors," *MIS Quarterly (*. (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2607190).

Boss, S. R., Galletta, D. F., Lowry, P. B., Moody, G. D., and Polak, P. 2015. "What Do Systems Users Have to Fear? Using Fear Appeals To Engender Threats and Fear That Motivate Protective Security Behaviours," *MIS Quarterly* (39:4), pp. 837–864. (https://doi.org/10.25300/MISQ/2015/39.4.5).

Bulgurcu, B., Cavusoglu, H., Benbasat, I., and Information, M. 2010. "Information Security Policy Compliance: An Empirical Study of Rationality-Based Beliefs and Information Security Awareness," *MIS Quarterly* (34:3), pp. 523–548. (https://doi.org/10.1093/bja/aeq366).

Crossler, R., and Belanger, F. 2014. "An Extended Perspective on Individual Security Behaviors : Protection Motivation Theory and a Unified Security Practices (USP) Instrument," *The Database for Advances in Information Systems* (45:4), pp. 51–71. (https://doi.org/10.1145/2691517.2691521).

Curry, M., Marshall, B., Crossler, R. . E., and Correia, J. 2019. "InfoSec Process Action Model (IPAM): Targeting Insider ' s Weak Password Behavior," *Journal of Information Systems*, pp. 1–51.

D'Arcy, J., Hovav, A., and Galletta, D. 2009. "User Awareness of Security Countermeasures and Its Impact on Information Systems Misuse: A Deterrence Approach," *Information Systems*. (http://pubsonline.informs.org/doi/abs/10.1287/isre.1070.0160).

Jenkins, J., and Grimes, M. 2014. "Improving Password Cybersecurity through Inexpensive and Minimally Invasive Means: Detecting and Deterring Password Reuse through Keystroke-Dynamics," *Information Technology* …. (http://www.tandfonline.com/doi/abs/10.1080/02681102.2013.814040).

Jensen, M. L., Dinger, M., Wright, R. T., and Thatcher, J. B. 2017. "Training to Mitigate Phishing Attacks Using Mindfulness Techniques," *Journal of Management Information Systems* (34:2), pp. 597–626. (https://doi.org/10.1080/07421222.2017.1334499).

Johnston, A. C., and Warkentin, M. 2010. "Fear Appeals and Information Security Behaviors: An Empirical Study," *MIS Quarterly* (34:3), pp. 549–566.

Liu, H., and Yu, L. 2005. "Toward Integrating Feature Selection Algorithms for Classification

and Clustering," *IEEE Transactions on Knowledge and Data Engineering* (17:4), pp. 491–502. (https://doi.org/10.1109/TKDE.2005.66).

Moody, G., Siponen, M., Quarterly, S. P.-M., and 2018, U. 2018. "Toward a Unified Model of Information Security Policy Compliance," *MIS Quarterly* (42:1).

Müller, O., Junglas, I., Brocke, J. Vom, and Debortoli, S. 2016. "Utilizing Big Data Analytics for Information Systems Research: Challenges, Promises and Guidelines," *European Journal of Information Systems*. (https://doi.org/10.1057/ejis.2016.2).

Posey, C., Roberts, T. L., Lowry, P. B., Bennett, R., and Courtney, J. 2013. "Insiders' Protection of Organizational Information Assets: Development of a Systematics-Based Taxonomy and Theory of Diversity for Protection-Motivated," *MIS Quarterly*.

Ramasubramanian, K., and Singh, A. 2016. *Machine Learning Using R*, Apress.

Sharma, R., Mithas, S., and Kankanhalli, A. 2014. "Transforming Decision-Making Processes: A Research Agenda for Understanding the Impact of Business Analytics on Organisations," *European Journal of Information Systems*. (https://doi.org/10.1057/ejis.2014.17).

Siponen, M., Pahnila, S., and Mahmood, M. A. 2010. "Compliance with Information Security Policies: An Empirical Investigation," *Computer* (43:2), pp. 64–71.

Siponen, M. T. 2003. "Information Security Management Standards: Problems and Solutions."

Verizon Enterprise Solutions. 2018. "2018 Data Breach Investigations Report." (http://www.verizonenterprise.com/resources/reports/rp_DBIR_2018_Report_en_xg.pdf).

APPENDIX A – INSTRUMENT

Assessing Attitudes of Protective-Motivation Behaviors

The following items adapted from Posey et al. (2013) employ a bipolar scale

Obvious: How to protect my account with <recommended PMB> is (is not) obvious.

Reasonable standard: Protecting my account with <recommended PMB> is (is not) a reasonable standard.

My responsibility: Protecting my account with <recommended PMB> is (is not) my responsibility.

Degree of difficulty: Protecting my account with <recommended PMB> is (is not) difficult to perform.

Involves a judgement call: Protecting my account with <recommended PMB> does (does not) involve a judgement call on my part.

Assessing Deviant Security Behavior Intention Scenarios and Questions

We plan to assess deviant intent to not follow the recommended behavior using scenarios where users can indicate agreement with a deviant behavior without fear of self-incrimination (Moody et al. 2018; Siponen et al. 2010). Here are drafts of proposed scenarios

Legitimate and Secure Email Use: Riley is a new employee in a small business and wants to make a good impression. During Riley's second week at this job an email message arrives from the senior manager directing each employee to complete an evaluation by the end of the day. Riley recognizes the manager's email address and signature block. Riley knows it might be a good idea to ask someone else about the email's authenticity, but decides not to bother anyone instead. Riley clicks the "Begin Evaluation" link in the email message.

Secure Internet Use: Peyton uses the computers at work to look for vacation deals. Peyton learns that many experts believe using work computers for non-work activities increases the likelihood of being attacked by malicious software. Peyton is invited to participate in training at work to learn more about safe internet usage. Peyton believes this training may be beneficial and initially agrees to participate. However later Peyton decides not to actually complete the training.

Scenario Reports Potential Ransomware Attack: Blake's work computer displays a "software update needed, click to update" message. Blake clicks the update button and the computer freezes. Blake steps away for a short break. When Blake returns the computer is still not responding. Blake can't even get the computer to restart. Blake decides to stop using the computer until the next day. Despite being aware of a policy to promptly report computer issues Blake leaves work for the day and does not tell anyone in IT.

Scenario Saves information according to retention policy; Backs up important data: Avery is on a companywide project. All twelve project members are from a different part of the

company. Avery takes notes for the project meeting. Avery knows saving information on the company cloud storage is part of the authorized retention policy. Avery prefers a free internet file sharing site and uploads the meeting minutes there. Avery then sends an email message to the other team members with a link to the notes.

Scenario Chooses password according to guidelines: Reese is an employee in the company with a system account. Reese receives a notification that his password is due to expire and needs to be changed. Reese clicks on the password reset. Before making a password choice, Reese reads that sharing a password with another site is not recommended. Reese chooses the same password that he uses for online banking.

Mindfulness while working: Rory is leaving work for the day and in a hurry. Just as Rory is about to leave she receives an email from a senior manager asking if she is still available. Rory realizes that when she is in a hurry there is less time for her to examine an email message and determine if it is legitimate or a potential source for security threats like ransomware. Although Rory does not recognize the email address, she decides to reply to the email anyway to tell the manager to say she is leaving for the day.

Protection Motivation Theory (PMT)

To assess drivers of intention, one PMT question per construct will be asked for each of the six behaviors, adapted from Boss et al (2015).

PMT Items used to assess drivers of intention. These items use a 7 point Likert scale from strongly disagree to strongly agree unless otherwise indicated.

Severity: If I were to do what <scenario named individual> did, a serious information security problem would result.

Vulnerability: if I were to do what <scenario named individual> did an information security problem would likely occur.

Fear: I am afraid of what may happen if I were to do what <scenario named individual> did.

Task Self-Efficacy: I can easily follow recommended guidelines for <recommended security behavior>

Response Efficacy: Information security problems can be reduced by participating in recommended security behavior training on <recommended security behavior> and following recommended guidelines.

Intention: I would act in the same way as <scenario named individual> did if I were in the same situation.

Dependent variables

Multiple indicators will be used to assess actual behavior.

- Participation: does participant initiate recommended security training (yes/no)
- Number of trainings completed
- Training test scores (4-6 questions per PMB block)
- Behavior self-reports ('has your behavior changed…' type questions)