An exploratory study into the aetiology of cybercrime: Comparing the utility of the Routine Activities Theory using a model-comparison approach

Troy Smith
Targeted Evidence-Based Research Solutions Ltd.
E-mail: dr_t_smith@yahoo.com

Abstract

The study took a human-centric approach to exploring cybercrime by comparing the utility of the Routine Activities Theory (RAT) between cyber-dependent (hacking and malware infection) and cyber-enabled (online harassment and unsolicited content) crime victimization. Data were collected using a self-administered survey (N = 200) disseminated using chain referral sampling to Facebook users residing in The Republic of Trinidad and Tobago. The data were analysed using Naïve Bayes, a supervised machine learning algorithm to develop classification models. This research found that the RAT showed greater utility with cyber-dependent crimes than cyber-enabled crimes. Voluntary and involuntary personal information disclosure through social media increased the likelihood of being a target of hacking, online harassment, and unsolicited content. Deviant online activities such as peer-to-peer downloads and watching pornography increased the risk of cyber-dependent crime victimization. Personal guardianship measures such as knowledge of cybercrime victimization and setting accounts to private are associated with cyber-dependent crimes. The study concluded that a difference in aetiology exists between cyber-dependent and cyber-enabled crime in relation to victimization suggesting that all cybercrimes cannot be explained equally with general criminological theories. This implies that a typological approach to the study of cybercrime victimization, prevention strategies and potentially legal frameworks is necessary.

Keywords: routine activities theory, cybercrime, cyber-dependent, cyber-enabled, model comparison, human-centric

Dr Troy Smith *is a graduate of the Doctor of Philosophy programme offered by the Institute of Criminology and Public Safety, University of Trinidad and Tobago. He is a scholar with multiple peer-reviewed articles and ongoing international projects focusing on the areas of cybercrime, problematic social media use and the effect of exogenous shocks on crime patterns. He also actively seeks to enhance research within the social sciences through the use of alternative statistical methods where appropriate; for example, the use of Bayesian analysis and Rasch measurement. He has over fourteen years of experience working in the area of national security in various specialty areas within Trinidad and Tobago. He is currently one of the directors of Research Analysis Inquiry and Development, a research non-profit entity focused on producing quality multidisciplinary research within the Caribbean.*

Introduction

Several studies and reports have highlighted the need for a human-centric approach to cybersecurity, citing the human element as an essential part of the disruption of cyber threats (Back et al., 2019; García-Segura, 2020; Jalkanen, 2019). This assumes that the activities of end-users in cyberspace and online social environments substantially influence the nature of cybercrime victimization (Lee & Downing, 2019; Lee et al., 2019). It also assumes cybercrime is like traditional crime in its dependence on social and situational factors (Bossler & Berenblum, 2019; Reyns et al., 2018; Rodriguez et al., 2017). However, a debate exists as to the congruence between cybercrime and traditional crime due to the perceived anti-temporal and anti-spatial properties of cyberspace (Sternberg, 2012; Taylor et al., 2019; Yar, 2005). Theorists and analysts of cyberspace suggest that concepts of proximity, distance, finite location, and time cannot be easily applied to cyberspace (Kitchin, 2009; Lattimer, 2013; Mitchell, 1995; Pratt et al., 2010; Smith & Stamatakis, 2020; Taylor et al., 2019; Yar, 2005). Hence, theories that are dependent on the concept of spatiotemporal alignment to define opportunities for the successful occurrence of a crime event such as the Routine Activities Theory (RAT) should be limited in their usability with cybercrime.

The RAT suggests that crime events result from social and situational conditions that lead to spatiotemporal convergence of a suitable target and a motivated offender in the absence of guardianship. The application of the RAT in research has shown promising results in some cases for the transferability of the RAT to cybercrime (Choi, 2008, 2018; Leukfeldt & Yar, 2016; Marcum et al., 2010; Smith & Stamatakis, 2021), while others have found little or no empirical support for the applicability of RAT to cybercrime (Bossler & Holt, 2009; Holt & Bossler, 2009; Ngo & Paternoster, 2011; Rodriguez et al., 2017). Therefore, the outcome is not definitive in its assessment of the congruence of cybercrime and traditional crime. Rather, it is suggestive of a mixed level of congruence dependent on the type of cybercrime. This is possibly further confounded by the fact that some cybercrimes employ various threat vectors that can limit the actual dependence on network technology. Hence a better approach would be to test whether the temporal and spatial differences/similarities between cyberspace and terrestrial environment and the associated effects are constant between different forms of cybercrime (Ilievski, 2016; Reyns, 2013; Yar, 2005). The results of the aforementioned comparison will form a base from which the congruence of cybercrime to traditional crime (which may be inconsistent across cybercrimes) can be assessed.

The difference in the temporal and spatial effects cannot be measured directly but since they are a direct result of the crime existing in cyberspace, the level of involvement/dependence on technology may be a suitable proxy. Since the level of involvement/dependence on technology is not a finite measure, for simplicity, the difference may be observed by classifying the crime into levels of technology involvement/dependence and subsequent comparison of the two levels.

The cybercrimes examined in this study were separated into pairs using a binary classification for cybercrimes, that is, cyber-dependent (techno-centric/computer-based) and cyber-enabled (people-centric/person-based) based on McGuire and Dowling (2013). Cyber-dependent crimes ('pure' cybercrimes) are offenses that can only be committed using a computer or networked technology (Furnell et al., 2015; Kirwan & Power, 2012). They are activities primarily directed against computers and network resources, which are often associated with secondary criminal outcomes. Cyber-enabled crimes are traditional crimes that can be exacerbated in scale or reach by using computers and information communication technology (ICT). Unlike cyber-dependent crimes, they can be committed without the use of ICT (McGuire & Dowling, 2013). If cybercrime is not congruent to traditional crime, that is, if the properties of cyberspace make it unique, the ability of the RAT should be limited because the properties of cyberspace do not align with the theory's core assumptions (Bock et al., 2017; Kringen & Felson, 2014; Yar, 2005). Therefore, the usability of the RAT should be markedly different between the two classes of cybercrime but similar within the classes. However, an inconsistent or unpaired outcome would suggest that other factors exist. Further, unexpected pairing may suggest that cybercrime may be better classified along the lines of another unidentified variable rather than technology, in which case the properties of cyberspace do not have a marked effect on the nature of the crime.

This study specifically focused on the relationship between the usability of the RAT with cybercrime victimization and dependence of the crime on technology, and the threat vector employed in the execution of the crimes. It assumes that a generalizable theory that has been broadly applied to traditional crime and is heavily dependent on spatial and temporal factors usability will be proportional to the congruence between the two categories. Therefore, this study seeks to examine the consistency of the effect of these properties on the usability of traditional theory and by extension congruency to traditional crime. However, while the nature of the relationship has been theoretically demonstrated previously, this study takes an empirical approach. This is achieved through a novel model-comparison approach to assessing cybercrimes along lines of technology dependence using Bayesian statistics. In their work, Prins and Kingdom (2018) demonstrate the model-comparison approach that allows the assessment of research hypotheses by comparing the model-performance between or among different outcome variables to predictors combinations through a series of examples and referenced studies. The approach taken provides insight into the potential proportionality of the effect of cyberspace on the congruence of cybercrime to traditional crime as constrained by spatial and temporal properties. This research aligns with one of the key gaps in cybercrime research, which is a need for the examination of behaviour and victimization in cyberspace from a multidisciplinary lens (Jaishankar, 2010, 2011; Leukfeldt, 2017; Ngo & Jaishankar, 2017). Specifically, Ngo and Jaishankar (2017) suggest that, in order to develop the sub-field of cyber criminology, research must be done to determine which theoretical framework best explains specific substantive forms of cybercrime. For example, RAT may be an effective predictive mechanism but only for a specific subset of cybercrimes.

Conceptual Framework

The Routine Activities Theory

The RAT has been applied to a range of victimization experiences (Argun, & Daglar, 2016; Leukfeldt & Yar, 2016; Louderback & Roy, 2018). The approach of RAT is based on two basic premises: first, crime occurs when motivated offenders are in proximity to targets with insufficient guardianship; second, on the probability of spatiotemporal convergence leading to the criminal act being affected by an individual's 'routine activities' (de Jong et al., 2019; Cohen & Felson, 2003; DeGarmo, 2011; Howell et al., 2019). Routine activities can be defined as generalized temporal and spatial patterns of recurrent and prevalent social activities which provide for basic population and individual needs, whatever their biological or cultural origins and generalized patterns of social activities in society (Cohen & Felson, 1979; DeGarmo, 2011; Wikstrom, 2018). However, Yar (2005) opined that RAT's approach to crime causation is dependent on the ability to define convergence of target, offender, and absence of a capable guardian in time and space. However, unlike the ontology of terrestrial space, which affords relations of proximity and distance between offenders and their prospective targets, the virtual environment appears to be configured very differently. Mitchell (1995, p. 8) states that cyberspace is "anti-spatial" because all points in the virtual world are equidistant, which makes the discussion of convergence or divergence between offenders and targets problematic. Yar (2005) also stated that time in cyberspace is not necessarily a supplementary variable that can account for simultaneity or temporal alignment (Leukfeldt & Yar, 2016). Therefore, the properties of cyberspace itself create a constraint on the RAT due to its limitations in accounting for spatio-temporal convergence.

Thus far, cybercrime victimization studies show an inconsistency regarding the efficacy of RAT in explaining cybercrimes. Some have shown promising results for the transferability of the RAT to cybercrime (Choi, 2008, 2018; Leukfeldt & Yar, 2016; Marcum et al., 2010), while others have found little or no empirical support for the applicability of RAT to cybercrime (Bossler & Holt, 2009; Holt & Bossler, 2009; Ngo & Paternoster, 2011; Rodriguez et al., 2017). This may be at least partially attributable to the variations in the type of cybercrime examined and the focus on select aspects of RAT. However, based on the results of their empirical study, Rodriguez et al. (2017) speculated that the predictors of cyber-victimisation may be dependent on the nature of the cybercrime it seeks to predict. For example, cyber-dependent crimes may have a stronger correlation to technological variables such as antivirus software and cyber-enabled protocols, which are computer focused, and cyber-enabled crimes would be dependent on social interaction variables (such as level of use of social networks). This study seeks to test this hypothesis put forward by Rodriguez et al. (2017) and other scholars by formally comparing cybercrimes based on the binary classification framework of cyber-dependent crimes ('true cybercrime') and cyber-enabled crimes. This binary classification that was developed by McGuire and Dowling (2013) separates cybercrime based on the crime's dependence on cybertechnology (or cybertechnology's role in its execution or as a target of the crime).

Methodology

The study assesses four types of cybercrime victimization experiences separated into pairs using a binary classification for cybercrimes: cyber-dependent (techno-centric/computer-based) and cyber-enabled (people-centric/person-based). The two classifications of cybercrime are operationally defined as follows:

- Cyber-enabled crimes are crimes which can be committed without the use of cyber-technology. They are not dependent on cyber-technology and can be executed independently in the physical world, and the end target of the crime is generally a human (for example, stalking and bullying).
- Cyber-dependent crimes are those which can only be committed using cyber-technology such as computers or computer networks. Further, in this form of cybercrime, the networked technology itself can be the target of the crime (for example, hacking and malware infection).

The cybercrime victimization/experiences examined were hacking, malware infection, online harassment and receiving unsolicited nude photos/explicit content. These cybercrimes were chosen to provide at least two types of cybercrime for each category. However, the selection was limited by the types of cybercrime found to be occurring in the jurisdiction of interest (Trinidad and Tobago) based on examination of the literature (Inter-American Development Bank & Organization of American States, 2016; Jessop, 2019). Unfortunately, since there are no dedicated cybercrime laws, only international sources could be examined as local law enforcement does not have an established recording system for cybercrimes.

The effects of online exposure, accessibility, digital and personal guardianship that relate to the tenets of RAT's factors of target suitability (Exposure/Visibility and Accessibility) and capable guardianship were examined as the primary independent variables using the Naïve Bayes classification algorithm. The performance of the models derived using Naïve Bayes was used as a measure of the ability of the RAT to explain that cybercrime. This means that the performance metrics of each model should correlate to the usability of the RAT and give it a numeric value. This was used to identify if a difference in the effect of the properties of cyberspace on different categories of cybercrime existed. If a difference was found to exist, further examination would determine if there was a pattern or proportionality to the level of effect.

Hypotheses

The present study aimed to investigate the extent to which the RAT (general criminological theory) are applicable across different types of cybercrime. Specifically, the hypotheses were as follows:

- Hypothesis 1 (H1): The utility of the RAT to explain victimization is not consistent across different cybercrimes.

- Hypothesis 2 (H2): The greater dependence on technology to achieve the criminal outcome the smaller the usability of RAT, that is, the RAT will have greater utility in explaining cyber-dependent crime than cyber-enabled crime

The Sample

The sampling method used in this research is exponential non-discriminative snowball sampling, which is a non-probability sampling method (Dudovskiy, 2018). The initial seed consisted of 450 Facebook users of varying ages (over 18) and geographic locations within the Republic of Trinidad and Tobago. These persons were provided with a link to a webpage describing the study from which the self-administered survey could be accessed and shared via instant messaging or social media.

Data collection was done in two phases with the initial distribution getting 98 responses. Bayesian sequential analysis with robustness check (Figure 1) suggested that the sample was statistically equivalent to the population of interest (age). However, given the low response, a reminder was sent one month after the initial distribution, which resulted in a further 102 responses. Given that snowball sampling was used, the absolute response rate could not be identified. However, of persons visiting the survey link, only 18% answered the survey. Analyses performed in this study were done on the resulting sample of 200 respondents.

Measures

Dependent Variables

Dependent variables in this study were "hacking", "malware", "online harassment" and "unsolicited content". These variables were coded as dichotomous (No = 0, Yes = 1). Hacking is defined as the subversion of a computer, system, or network for malicious purposes to access content or control the device without the owner's permission; in other words, subversion of computer security for malicious purposes. Malware is software that is installed on the user's computer without his/her consent to obtain control of the computer partially or fully for malicious purposes. Online harassment includes any activity occurring online that intends to humiliate and/or terrorize a victim. Unsolicited content (nude photographs/explicit material) victimization is defined as the receipt of unsolicited nude photographs or other sexually explicit images while online or directly through a messaging service. Respondents were asked to self-identify as victims based on a description of experiences associated with each type of crime. The questions were: "Has anyone ever gained unauthorized access to your email accounts or computer files?"; "Have you experienced any security problems over the Internet or on your computer involving malware such as virus, ransomware, or spam?"; "Has anyone sought to humiliate or terrorize you online (cyberbullying, revenge porn, online harassment)?"; "Have you received unsolicited nude photos or other such explicit content on your computer or mobile device?"

Independent Variables

The independent variables observed in this study are derived from the RAT and include capable guardianship and target suitability (target exposure and target accessibility). The assumption is made in this study that the presence of a motivated offender is a constant (Yar, 2005) and it is the suitability of the target and capable guardianship (absence or presence) that predict victimization. The operationalization of the observed variables used as proxies to these constructs and their respective operational definitions are as follows:

- *Target accessibility*. This variable assesses the placement of the individual, which increases, the potential risk of the intended attack. This was measured by the variables of Internet browsing hours, online shopping hours, social media hours, watching adult content, and downloading music or videos represent the time spent per day by the respondent engaging in these online activities. These activities create opportunities for online victimization through accessibility or online visibility.
- *Target exposure.* This variable assesses the visibility of a user, which solidifies the suitability of the target. This was measured by three items assessing the degree of self-disclosure of personal information in various online settings. The survey items measured the number of times a user posts their picture, personal information, and location online per week.

- *Capable guardianship.* This aspect of the RAT is subdivided into the two variables of technical and personal guardianship for this study. Technical guardianship refers to technological solutions such as antivirus, firewall, and anti-spyware, which are developed to guard/protect computer systems and networks from offenders. Personal guardianship was measured using three items that assessed the potential target's skill level and knowledgeability with networked technology and awareness of cybercrime victimization risk.

Control Variables/Demographics

Research has shown that the demographic characteristics of the user are connected to victimization (Holtfreter et al., 2006; Reyns, 2013; Yucedal, 2010). To control for potentially contributing factors of demographic data on respondents' sex and age were collected.

Development of Naïve Bayes Classification model with B-*Course*

As with any Bayesian method the foundation is the choice of a prior. Although no simple approach can be found to determine a non-informative prior, B-Course takes an approach that produces a prior that is both computationally practical to implement and is not too sensitive for variable transformation (Berger et al., 2015; Myllymaki et al., 2002). The formula employed in B-Course use an equivalent sample size (ESS), 'N' which is close to Jeffrey's prior so that if an empty network is calculated, 'N' is equal to the Jeffreys' prior (Myllymaki et al., 2002). B-Course deals with missing data using a method termed "ignoring", where it disregards only the missing parts of the data row rather than the entire row, which is suitable in cases where the data is Missing

Completely at Random (MCAR) (Myllymaki et al., 2002; Scheffer, 2002). Since missing values were less than 5% and by observation were random in respect to the questions that had missing data, the researcher decided that the default method would be suitable in this study.

Model selection was executed using a combination of greedy and stochastic heuristics, which incrementally alter inconsistent value assignments to all the variables as they move to 'more complete' through a randomized search (Barták, 1998; Myllymaki et al., 2002).). For networks with 5 -15 nodes, 100 to 1000 data vectors are sufficient to construct an optimal network regardless of its complexity using this algorithm (Myllymaki et al., 2002). The predictive accuracy of each model considered is assessed by leave-one-out (LOO) cross-validation. In LOO a single data vector is removed iteratively from the data matrix leaving a N-1 vector for each iteration (Myllymaki et al., 2002). The resulting model from the N-1 is used to determine the class identity of the data vector which was removed. B-Course does not simply use the prediction but also the certainty on the prediction provided by the model to rate the classifier. If the classifier finds the prediction "positive" (78%), "negative" (22%), and the actual value is positive, the classifier gains 0.78 points. The aforementioned process is repeated for each data vector and the geometric average predictive accuracy is calculated and used as the estimated accuracy of the classifier. Since technically the classifier produced in each iteration will be different, it is the aspect that remains constant between all classifiers derived in the set of predictors utilized. The final model is the set of predictors that performed best in LOO and was subsequently trained with all N classified data vectors. The key aspect of the utilization of this method is that the choice of the final model is the product of cross-validation that is the best model with the most accurate representation of the model in the given confusion matrix.

Comparative Evaluation of Binary Classifiers: Goodness-of-Fit

Classification models in criminology are used in risk prediction and aid in identifying at-risk groups by assessing known characteristics. They are also valuable tools for developing strategic plans and mitigation programs (Latessa & Lovins, 2014; Oswald et al., 2018; Tollenaar & van der Heijden, 2019). In this study, the relative performance of the RAT is based on a comparison of model performance between the two classifications of cybercrime. As such, the identification of a suitable and acceptable metric of performance upon which to base the comparison is required. For skewed classes, scalar values tend to provide an overoptimistic rating of the classifier's performance on detecting the majority class (Akosa, 2017; Bekkar et al., 2013). In real-world scenarios, especially in studies relating to crime and health, it is unlikely to have an even distribution of classes (Song et al., 2016; Wang & Yao, 2013). Further, the scalar measures do not assess the need for a balance between precision and recall, which is required for real-world applications. The limitations of scalar measures can be overcome by using combination metrics such as F_1-Score and Matthews Correlation Coefficient (MCC), which provide a more realistic evaluation of model performance in real-world conditions.

The F_1-Score (Sørensen–Dice coefficient) provides for the measurement of recall and precision simultaneously and as such the two important factors can be balanced (Dey et al., 2019). The MCC is a measure of the correlation between the observed and anticipated classes using the data contained in the confusion matrix (Song et al., 2019; Chicco & Jurman, 2020). Unlike the F_1-Score, to obtain a high MCC, the binary classifier must predict the majority of both the positive and negative classes accurately (Chicco, 2017; Delgado & Xavier-Andoni, 2019).

Results

The first section of the results (Bayesian Networks) provides the results of model development using the Naïve Bayes classifier. While this section primarily focuses on assessing H1, it also addresses aspects of H2. Regarding H1, it identifies the variables correlated to each cybercrime so that trends can be identified as it relates to which variables are statistically significant for the different cybercrimes., Further, it identifies the number of variables identified by the RAT that are relevant in explaining each type of victimization. While both factors can give insight into the inconsistency of the utility of the RAT's explanatory power across different cybercrimes (H1), the latter can also be used as a de facto measure of the utility of the RAT (H2). The second section of the results (Performance Metrics) seeks to add greater depth to the analysis, which is necessary to assess the veracity of H2 by providing empirical evidence of the difference in the utility of the RAT across cybercrime by assessing the differences in model performance.

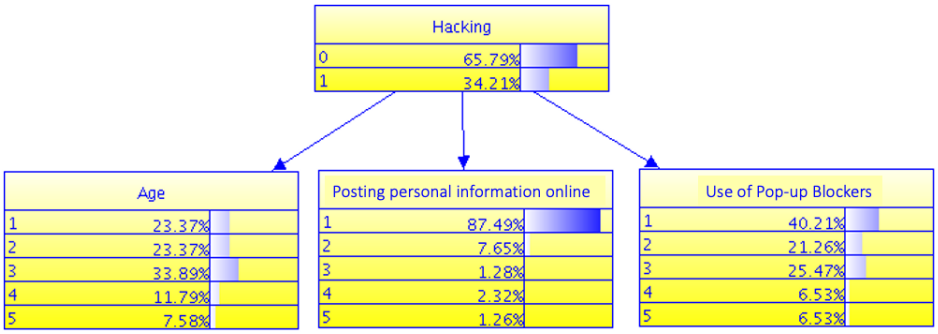
Bayesian Networks

Hacking Victimization

The Naïve Bayes analysis of the dataset indicates that the Hacking is best predicted by three variables, which are shown in Figure 1. The directed acyclic graph provided for the model also shows the distribution of responses for each class of the variables included in the model. The use of pop-up blockers (PhysG3) acts as a protective factor while posting personal information online (TS2) increases the risk of victimization. Further analysis of the conditional probability tables indicates that hacking victimization risk increases with age. The estimated classification accuracy of the best model found is 78.72%. Further, the output of the analysis indicates that the model will accurately identify 95.2% of the data instances of a given dataset that belongs to the negative class (Sensitivity). Similarly, the model accurately identifies 46.9% of instances belonging to the positive class (Specificity).

Figure 1

Naive Bayes Directed Acyclic Graphical Model for Hacking Victimization with Variable Class Distribution

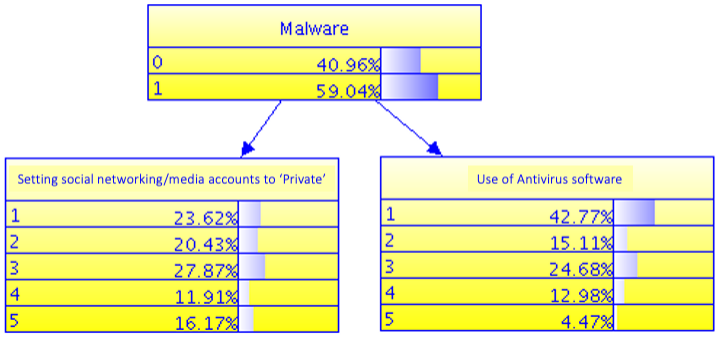


Malware Victimization

The relationship between malware victimization and the predictors identified in the analysis is represented in Figure 2. The use of antivirus software (PhysG1) decreases the risk of victimization while the model suggests that overall, knowledge of victimization risk/awareness of cybercrime (PerG3) increases the risk of victimization. The estimated overall classification accuracy of the selected model is 66.67%. The sensitivity was found to be 85.5% and the specificity was 39.5%.

Figure 2

Naive Bayes Directed Acyclic Graphical Model for Malware Victimization with Variable Class Distribution



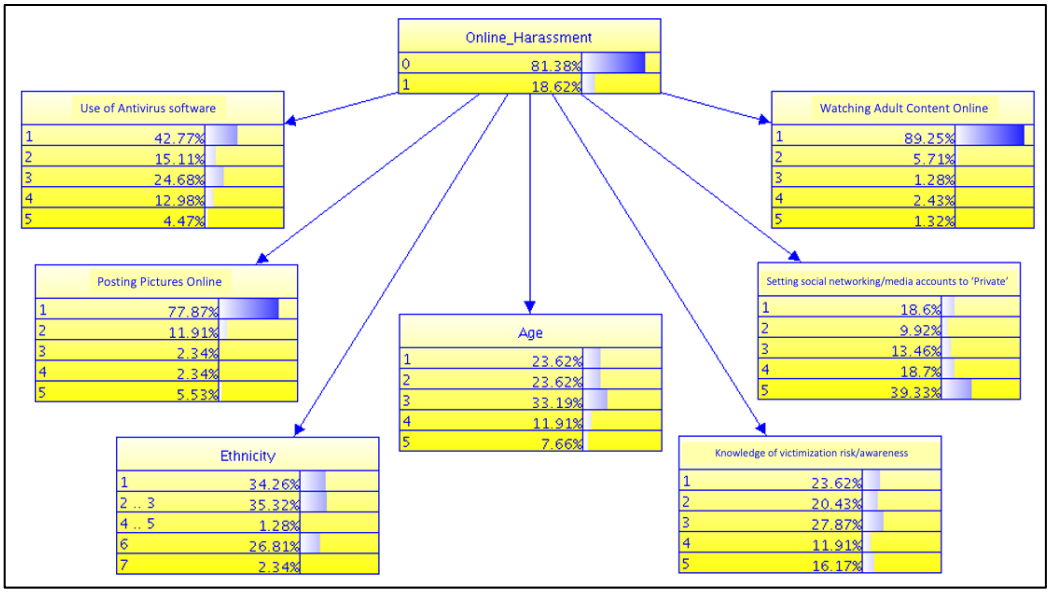
Online Harassment Victimization

Online harassment is best predicted by the subset of variables shown in Figure 3. The analysis indicates that the predictors of posting pictures online (TS1), watching adult content (E4), and knowledge of victimization risk/awareness of cybercrime (PerG3) increase the risk of online harassment victimization. Setting social media accounts to 'private' (PerG1) and use of antivirus software (PhysG1) were found to be protective factors as they resulted in decreased risk of online harassment victimization. The risk of online harassment was found to decrease as age increased.

The age group of 15-24 showed the highest probability of online harassment victimization. The estimated classification accuracy of the best model found is 89.25%; while the sensitivity was found to be 47.1% and the specificity was 98.7%.

Figure 3

Naive Bayes Classification Model for Online Harassment Victimization with Variable Class Distribution

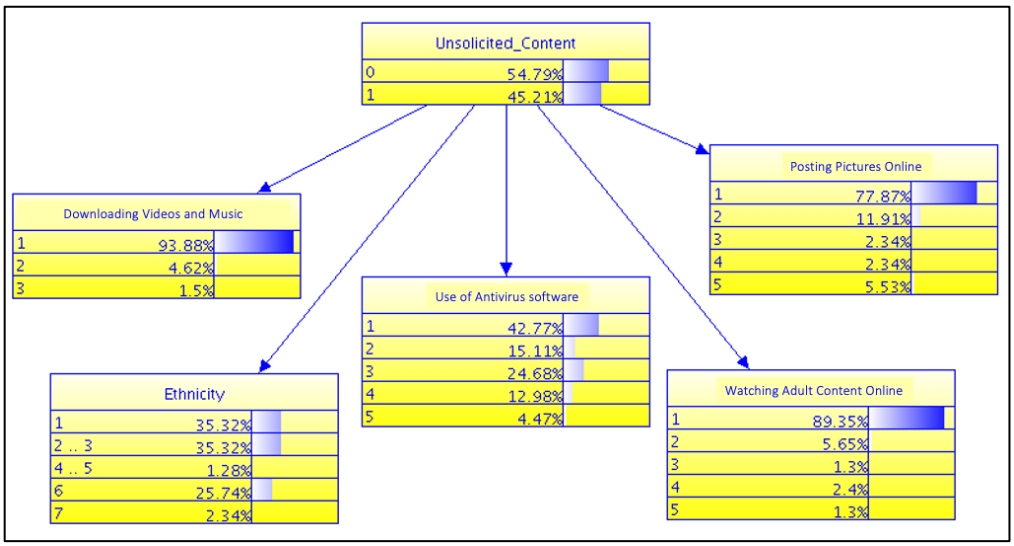


Unsolicited Content Victimization

The statistically significant predictors of unsolicited content victimization determined by the Naïve Bayes algorithm are shown in Figure 4. The unsolicited content victimization predictors of posting pictures online (TS1), watching adult content (E4), and downloading videos and music (E5) were found to increase the risk of victimization. However, as expected, the variable of use of antivirus software (PhysG1) was found to be a protective factor. The estimated overall classification accuracy of the selected model is 72.04%. The sensitivity of the model was found to be 69.0% while the specificity was 74.5%.

Figure 4

Naive Bayes Classification Model for Unsolicited Content Victimization with Variable Class Distribution



Performance Metrics

This section assesses the comparative performance of the models developed for each form of cybercrime. Combination performance metrics are used to account for data imbalances in hacking and online harassment, that is, normalization of the comparative statistic. Both F_1-Score and MCC are presented because the former does not account for the importance of true negatives. Further, the F_1-Score only focuses on the balance of precision and recall, while the latter is holistic in its assessment of all four outputs of the confusion matrix. It should be noted that these statistics are at a fixed threshold and do not account for possible optimization through adjusting the threshold for the application of the model, which is beyond the scope of this project. Table 1 provides the F_1-Scores and MCC values for the four Naïve Bayes models. The numeric values for the F_1-Score showed hacking (0.60) with the lowest score, followed by online harassment (0.62) then unsolicited content (0.69) and finally malware (0.75) with the highest score. The hacking and online harassment models both have scores between 0.45-0.62 and are classified as Poor, while malware and unsolicited content with scores between 0.63-0.72 is classified as Fair.

Arranging the models in ascending order of MCC value gives malware (0.28) followed by unsolicited content (0.44) followed by hacking (0.51) and online harassment (0.60). This is similar to the order of models given by the F_1-Scores but in reverse. The elementary metrics of accuracy, precision and specificity follow the same pattern as that for the MCC. The MCC scores suggest that the models increase in Goodness of Fit from Poor (malware) to Fair (unsolicited content) to Good (hacking and online harassment).

Table 1

Comparison of Model Performance between Cyber-dependent and Cyber-enabled Crimes Using Elementary Metrics and Combination

Performance Metric	Cyber-dependent (%)		Cyber-enabled (%)	
	Malware	Hacking	Unsolicited Content	Online Harassment
Accuracy	0.67	0.79	0.72	0.89
Precision	0.67	0.83	0.69	0.89
Recall	0.85	0.47	0.69	0.47
Specificity	0.39	0.95	0.75	0.99
F_1-Score	0.75	0.60	0.69	0.62
MCC	0.28	0.51	0.44	0.60

Discussion

Statistical comparison of cyber-dependent crimes showed that there is inconsistency in the ability of the RAT to explain cybercrime victimization. This inconsistency relates to the variables which are statistically significant to predicting victimization, the effect size and nature (positive or negative) of the variables on the probability of victimization and the predictive power of the classification model that can be derived using the RAT. Further, the empirical evidence suggests that the separation of cybercrimes as it relates to the usability of the RAT and by extension the role of cyberspace is not binary.

Initial examination of the cybercrime models showed that the Cyber-enabled crimes shared some predictors while the Cyber-dependent crimes did not share any. At the macro-level (three main tenets of the RAT), cyber-dependent crimes only shared capable guardianship, while the cyber-enabled crimes were represented in all categories. Therefore, while there may be similarities

between the two classes of cybercrimes, they remain individualistic relative to predictive factors. Further, the latter observation suggests that victimization by cyber-enabled crimes is more closely related to routine activities than cyber-dependent crimes. Ordering of the cybercrimes by the number of predictors gives malware (2), hacking (3), unsolicited content (5) and online harassment (7). This supports the hypothesis (H1) presented by the study proposed that the utility of the RAT to explain victimization is not consistent across different cybercrimes. The ordering implies that there exists a level of proportionality between cyber-dependence and the importance of routine activities on victimization. This supports the hypothesis (H2) tested in this study that cyber-dependent crimes would be less effectively described by the RAT than cyber-enabled crimes.

The empirical evidence supports the hypothesis (H2) that since the RAT is based on presuppositions that are spatiotemporally rooted in terrestrial space, that if cyberspace is anti-spatial as proposed, the usability of the RAT should be proportional to the level of involvement of cyberspace in cybercrime (Lee et al., 2017; Vakhitova et al., 2016;). However, this study also showed that while a link exists between the level of involvement of cyberspace and the explanatory power of the RAT, it does not follow a binary typology. This outcome is also possibly linked to the situational nature of the theory different crimes require different conditions for suitable opportunities to be successful; suggesting that different routine activities will be relevant in the prediction of different crimes. However, this still can align with a general pattern between groups of crimes and the measures of the RAT that are relevant.

Assessment based on the performance of the models produced using the RAT and the Naïve Bayes classifier shows a slightly different perspective compared to simply evaluating the number of predictors. While the MCC scores presented in Table 1 show that malware and online harassment remain at separate ends of the performance spectrum, the cybercrime of hacking and unsolicited content victimization are switched. Hacking (MCC = 0.51) showed a slightly higher Goodness of Fit than unsolicited content (MCC = 0.44). The literature indicates that both hacking and unsolicited content victimization can employ techno-centric and human-centric approaches to victim acquisition. This means that cyber-enabled crimes may have threat vectors that simply require connecting to the Internet and cyber-dependent crimes can have human-centric threat vectors (for example, methods that rely on social engineering).

Hacking can be heavily dependent on social engineering and not necessarily only based on targeting exploitable weaknesses in technology. Social engineering is defined as an attack vector that targets human psychology and relies on human interaction to gain unauthorized access to networks and systems (Anderson, 2008; Trevizo, 2019). In hacking, threat actors can utilize social engineering techniques to influence, manipulate or trick users into divulging critical or restricted information, or even unwittingly grant them access to network resources (Anderson, 2008; Hurych, 2019; Jain et al., 2016). The nature of the opportunity required by the RAT for a crime event switches from technical exploitation to be the result of routine activities. Therefore, the

opportunistic nature of crime remains in line with the RAT. However, the situation from which the opportunity arises is dependent on a motivated offender finding an exploitable weakness in technology in one instance, while in the other instance, the offender seeks exploitable weaknesses arising out of human behaviour/routines/characteristics.

Similarly, while receipt of unsolicited content can occur through interpersonal interaction, it can also be somewhat automated through spam emails and pop-ups. For example, spam is defined as unsolicited commercial e-mails, often from someone trying to sell something, which may include pornographic content. Persons may relinquish their email to obtain free services or gain access to pirated content and as a result expose themselves to spam (Madigan et al., 2018). Further, unwanted exposure can also come from sexually explicit images or videos in pop-up windows as well as on websites (Lagorio-Chafkin, 2019; Madigan et al., 2018). The fact that the variables of 'watching adult content' and 'downloading music and video' were found to be significant factors link to the aforementioned point since they are known to be associated with explicit pop-ups.

While studies on unwanted sexual material on the Internet have focused on adolescents, they highlight the fact that persons can be exposed to such content by simply being online (Madigan et al., 2018; Wolak et al., 2007). The RAT's conceptualization of the occurrence of a crime event requiring convergence between a motivated offender and a suitable target in the absence of capable guardianship still holds. However, the manifestation and hence the measures of the RAT concepts will not be the same as in a target-specific approach. In this 'passive' approach, the accessibility of the target may remain relevant as the user's activities will be indicative of the probability of converging with the 'trap' set by the offender. However, target exposure may not be relevant as the offender is not actively searching for a specific target or target suitability indicators.

Similarly, for unsolicited content, the RAT can account for the human aspect of the spatiotemporal alignment of factors leading to victimization being limited by the confounder of threat vector, which may only require the victim to be online and engaging in regular activities such as web searches or downloading. Hence, an identifiable limitation of the RAT or its implementation is that it focuses only on the routine activities of the victim, simplifies the offender's role to only being motivated to commit the crime, and ties the offender to direct contact or targeting of a specific victim. However, the offender's choice of threat vectors is potentially linked to which routine activities will increase or decrease the risk of victimization as these define the mechanism or pathway to the target and the specificity of the mode of attack. The aforementioned redefines the classification, or rather the boundaries between forms of cybercrime, as it relates to the efficiency of the RAT.

In addition, a closer look at the individual metrics and overall model performance suggest that the RAT indeed is limited in its explanatory power in relation to cybercrime in general. Comparison of the F1-Scores and MCC values for the cybercrime models in Table 1 shows the majority being

fair. This is suggestive of the limited explanatory power of the RAT when applied to cybercrime. Further, looking at the TPR and TNR paying for the various cybercrimes shows a large skewing for hacking, malware and online harassment. In all three cases, one is generally twice the size of the other. For both hacking and online harassment, the TNR is twice the TPR, suggesting that the identified predictors are associated with a decreased risk rather than predictive of opportunities for a successful crime event. Conversely, the TNR for malware is half of its TPR suggesting that engaging in the specified activities do increase the risk of victimization. However, not engaging in them does not necessarily decrease the risk of victimization.

This finding aligns well with previous speculations of the intrinsic criminal environment created by cyberspace, which in some cases only require that the user go online or use networked technology to become a victim (Brady et al., 2016; Reyns et al., 2018). Specifically, Holt and Bossler (2014) suggest that the Internet has a normalizing effect, where users are all equally at risk if all other factors are constant. In such cases, cybercrime would not be compatible with an opportunity framework to explain victimization (Bossler & Holt, 2009; Ngo & Paternoster, 2011; Reyns, 2015; Yucedal, 2010). The best performance of the RAT was seen with unsolicited content with high F1-Score, TPR and TNR, which were suggestive of at least good performance. This aligns with the speculations of Rodriguez et al. (2017) that predictors of cyber-victimization are dependent on the nature of the cybercrime it seeks to predict, with cyber-enabled crimes dependent on social interaction variables. Therefore, it would logically follow that RAT would be best suited to assess the temporal and spatial patterns of recurrent and prevalent social activities (DeGarmo, 2011; Wikstrom, 2018).

Conclusion

The results of this study suggest that the relationship between cybercrime and traditional crime may vary depending on the specific type of cybercrime and its dependence on technology, where the dependence is defined by both its intrinsic need for technology to exist, and the nature of the threat vector employed in its execution. Additionally, the empirical evidence suggests that different cybercrimes have different levels of similarity to traditional crimes.

A cybercrime's dependence on cyberspace as defined by its ability to exist in the absence of technology is not always equivalent to its dependence based on how it is executed. Rather, it is potentially a duality of dependence on what is the target and threat vector employed. The threat vector may be considered a proxy for a motivated offender as it potentially defines the asymmetric nature or reach of the attack and its target; that is, the action targets a specific user or randomly target online users en masse. In the latter 'en masse' approach, the choice of victim is technically outside the control of the offender. In other words, the offender does not have an unambiguous profile of desired victims. Furthermore, the disparity between the number of predictors in the models for hacking and unsolicited content and their respective model performance shows that the

two are not necessarily correlated; thus, the number of predictors does not necessarily define the model performance.

The existence of this duality of threat vectors can potentially blur the line between cyber-dependent and cyber-enabled crimes from a victimization standpoint. It suggests that while there are indeed extremes based on the role of cyberspace as seen with malware and online harassment, there is also a nexus. A nexus exists between the dichotomy of cyber-dependent and cyber-enabled crimes due to the mixing of threat vectors. Therefore, the effect of cyberspace on cybercrime is not fully defined by the crime's dependence on technology or the target (human or technological). Rather it is also dependent on the threat vector employed as suggested by Kranenbarg (2018) whose work showed that there was natural clustering of cybercrimes by modus operandi. In which case, when the threat vector is not accounted for in model development, it can act as a confounder. Further, singular crimes such as hacking with two pathways may be better described by two models or a merged model which is artificially balanced to account for both threat vectors.

Several limitations should be considered in the review of the outcomes of this study. First, while the data collected using chain referral were suitable for the exploratory nature of this study, future work should be made generalizable by using a large representative (randomly selected) population sample. Second, like most cybercrime research, this study utilized self-reported behaviour rather than actual behaviour (van't Hoff-de Goede et al., 2021). However, this potentially lends to errors in participant recall, acquiescence or social desirability bias, and errors associated with assuming that intent is equivalent to actual behaviour. In addition, it is somewhat difficult to ascertain whether a person has been a victim of cybercrime; particularly the 'true' cybercrimes such as hacking and malware as people may become victims and be unaware. In addition, they may not possess the basic knowledge to discern the difference between computer and mobile devices, issues associated with faulty hardware, software or Internet service provider (ISP) problems from the effects of hacking or malware. Hence, there is the possibility of a skewed interpretation of outcomes given the underlying issue of occurrence versus detection. Lastly, a better insight into the difference in the utility of the RAT in explaining cybercrime victimization and any associated typological patterns would require a study into a larger variety of crime types, such as online consumer fraud, romance scams, and ransomware. Additionally, extended observation of participants to collect longitudinal data for a multitude of crimes would contribute to a deeper understanding of the consistency/stability of the relationship between routine activities and victimization. This would be very important in establishing the importance of threat vector in the trajectory of cybercrime victimization.

The main policy implications associated with the results of this study are associated with the design/focus of cybercrime victimization initiatives. The study's results demonstrate that vulnerability to risk depends on crime-specific, routine activity aspects of computer and Internet use. This suggest that the formation of an unambiguous risk profile for cybercrime is improbable.

Therefore, strategies/policies should focus on educating users on 'cyberhygiene' and protective online behaviours such as target hardening and target removal. For example, the general effectiveness of physical guardianship across the various cybercrimes in the study suggest that cybercrime campaigns should encourage users to install antivirus software and pop-up blockers as 'physical' barriers to victimization. Further, the policies deployed by the public and private sectors should include plans to educate and develop the technical skills of users to help them understand the need to use protective measures and methods of utilization. Resources should be invested into determining factors that motivate users to engage in protective behaviours and that may aide or stymie the transition from intent to actual behaviour. Additionally, criminologist and technocrats must work in tandem to develop holistic strategies for cybersecurity/online safety as there is not only a need to create and implement digital barriers, but to also ensure their use and effectiveness.

Future research can improve and build on this study in a variety of ways. First, future research needs to focus more on specific incident characteristics to disentangle the relationship between routine activities and victimization. This would require specific data collection efforts to identify the source of the offending (or specific offender) and greater specificity with the cybercrimes. For the latter, broad categories such as online harassment should be sub-divided into cyberbullying, cyber-sexual harassment, online stalking and similarly related behaviours. Second, though self-reporting is a useful and insightful method of collecting data, it offers little in understanding the mechanism by which these behaviours are linked to victimization. This is especially important as the same behaviour does not always result in victimization. Any improvement to the existing knowledge on the mechanism of victimization would require longitudinal qualitative data collection and the observation of actual behaviour (Crossler et al., 2013). The latter can be achieved in an ethical, cost-effective way in the offline world by confronting respondents with realistic scenarios that request personal information (for example, asking them to share their password or requiring them to select the most appropriate password from a list of options).

References

Akosa, J. (2017, April 2-5). *Predictive accuracy: A misleading performance measure for highly imbalanced data*. [conference session]. SAS Global Forum 2017 Conference, Orlando, FL. https://support.sas.com/resources/papers/proceedings17/0942-2017.pdf

Anderson, R. (2008). *Security engineering: A guide to building dependable distributed systems* (2nd ed.). Wiley Publishing Inc.

Argun, U., & Daglar, M. (2016). Examination of routine activities theory by the property crime. *International Journal of Human Sciences, 13*(1), 1188-1198. doi: 10.14687/ijhs.v13i1.3665

Back, S., La Prade, J., Shehadeh, L., & Kim, M. (2019). Youth hackers and adult hackers in South Korea: An application of cybercriminal profiling. *IEEE European Symposium on Security and Privacy Workshops*, 410–413. doi: 10.1109/EuroSPW.2019.00052

Bartak, L. (1998). Book reviews. *Autism*, *2*(3), https://doi.org/10.1177/1362361398023015

Bekkar, M., Djemaa, H., & Alitouche, T. (2013). Evaluation measures for models assessment over imbalanced data sets. *Journal of Information Engineering and Applications*, *3*(10), 27–38.

Berger, J. O., Bernardo, J. M., & Sun, D. (2015). Overall objective priors. *Bayesian Analysis*, *10*(1), 189–221.

Bock, K., Shannon, S., Movahedi, Y., & Cukier, M. (2017). Application of routine activity theory to cyber intrusion location and time. *13th European Dependable Computing Conference (EDCC)*, 139-146, doi: 10.1109/EDCC.2017.24.

Bossler, A., & Berenblum, T. (2019). Introduction: New direction in cybercrime research. *Journal of Crime and Justice*. https://doi.org/10.1080/0735648X.2019.1692426

Bossler, A., & Holt, T. (2009). On-line activities, guardianship, and malware infection: An examination of routine activities theory. *International Journal of Cyber Criminology*, *3*(1), 400-420.

Brady, P., Randa, R., & Reyns, B. (2016). From WWII to the World Wide Web: A research note on social changes, online "places," and a new online activity ratio for routine activity theory. *Journal of Contemporary Criminal Justice*, *32*(2), 129–147. https://doi.org/10.1177/1043986215621377

Chicco, D. (2017). Ten quick tips for machine learning in computational biology. *BioData Mining*, *10*(1), 1–17.

Chicco, D., & Jurman, G. (2020). The advantages of the Matthews Correlation Coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, *21*(1), 1–13.

Choi, J. (2018). *A Lifestyle-Routine Activity Theory (LRAT) Approach to Cybercrime Victimization: Empirical Assessment of SNS Lifestyle Exposure Activities* [Seoul National University]. https://s-space.snu.ac.kr/bitstream/10371/141256/1/000000150664.pdf

Choi, K. (2008). Computer crime victimization and integrated theory: An empirical assessment. *International Journal of Cyber Criminology*, *2*(1), 308–333.

Cohen, L. E., & Felson, M. (1979). Social change and crime rate trends: A routine activity approach. *American Sociological Review*, *44*(4), 588–608.

Cohen, L., & Felson, M. (2003). Routine Activity Theory. In *Criminological Theory: Past to Present* (2nd ed.) (F. T. Cullen & R. Agnew, Eds.). Los Angeles, CA: Roxberry Publishing Company.

Crossler, R. E., Johnston, A. C., Lowry, P. B., Hu, Q., Warkentin, M., & Baskerville, R. (2013). Future directions for behavioral information security research. *Computers & Security*, *32*, 90–101. https://doi.org/10.1016/j.cose.2012.09.010

de Jong, E., Bernasco, W., & Lammers, M. (2019). Situational correlates of adolescent substance use: An improved test of the routine activity theory of deviant behavior. *Journal of Quantitative Criminology*. https://doi.org/10.1007/s10940-019-09433-w

DeGarmo, M. (2011). Understanding the comparisons of routine activities and contagious distributions of victimization: Forming a mixed model of confluence and transmission. *International Journal of Criminology*, *4*(1), 584–603.

Delgado, R., & Xavier-Andoni, T. (2019). Why Cohen's Kappa should be avoided as performance measure in classification. *PLoS ONE, 14*(9).

Dey, I., Heitman, E., & Vewrano, J. (2019). Evaluating classification models. https://towardsdatascience.com/hackcvilleds-4636c6c1ba53

Dudovskiy, J. (2018). Snowball sampling. Retrieved 28 July 2019, from Research-Methodology website: https://research-methodology.net/sampling-in-primary-data-collection/snowball-sampling/

Furnell, S., Emm, D., & Papadaki, M. (2015). The challenge of measuring cyber-dependent crimes. *Computer Fraud & Security*, *10*, 5–12. https://doi.org/10.1016/S1361-3723(15)30093-2

García-Segura, L. (2020). European cybersecurity: Future challenges from a human rights perspective. In J. Ramírez & J. Biziewski (Eds.), *Advanced sciences and technologies for security applications*. Springer.

Holt, T, & Bossler, A. (2014). An assessment of the current state of cybercrime scholarship. *Deviant Behavior*, *35*(1), 20–40. https://doi.org/10.1080/01639625.2013.822209

Holt, Thomas, & Bossler, A. M. (2009). Examining the applicability of lifestyle-routine activities theory for cybercrime victimization. *Deviant Behavior*, *30*(1), 1–25. http://dx.doi.org.ezproxy.uky.edu/10.1080/01639620701876577

Holtfreter, K., Reisig, M., & Blomberg, T. (2006). Consumer fraud victimization in Florida: An empirical study. *St. Thomas Law Review*, *18*, 761–879.

Howell, C. J., Burruss, G. W., Maimon, D., & Sahani, S. (2019). Website defacement and routine activities: Considering the importance of hackers' valuations of potential targets. *Journal of Crime and Justice*, *42*(5), 536–550. https://doi.org/10.1080/0735648X.2019.1691859

Hurych, L. (2019, June 30). Hacking and social engineering with a 70% success rate. *Medium* https://medium.com/swlh/hacking-and-social-engineering-with-a-70-success-rate-17bcddd06e99

Ilievski, A. (2016). An explanation of the cybercrime victimisation: Self-control and lifestyle/routine activity theory. *Innovative Issues and Approaches in Social Sciences*, *9*(1), 30–47. https://doi.org/10.12959/issn.1855-0541

Inter-American Development Bank & Organization of American States. (2016). *Cybersecurity: Are we ready in Latin America and the Caribbean? 2016 Cybersecurity Report.* Washington, D.C.: Inter-American Development Bank. https://publications.iadb.org/publications/english/document/Cybersecurity-Are-We-Ready-in-Latin-America-and-the-Caribbean.pdf

Jain, A., Tailang, H., Goswami, H., Dutta, S., Sankhla, M., & Kumar, R. (2016). Social engineering: Hacking a human being through technology. *Journal of Computer Engineering*, *18*(5), 94-100.

Jaishankar, K. (2010). The future of cyber criminology: Challenges and opportunities. *International Journal of Cyber Criminology*, *4*(1 & 2), 26–31.

Jaishankar, K. (Ed.). (2011). *Cyber criminology: Exploring Internet crimes and criminal behavior*. Routledge. https://doi.org/10.1201/b10718

Jalkanen, J. (2019). *Is human the weakest link in information security?: A systematic literature review* [Masters thesis, University of Jyvaskyla]. https://jyx.jyu.fi/handle/123456789/64186

Jessop, D. (2019). The Caribbean's cybersecurity response must evolve. https://www.caribbean-council.org/caribbeans-cybersecurity-response-must-evolve/

Kirwan, G., & Power, A. (2012). *The psychology of cybercrime: concepts and principles*. IGI Global. https://doi.org/10.4018/978-1-61350-350-8

Kitchin, R. (2009). Space II. In R. Kitchin & N. Thrift (Eds.), *International Encyclopedia in Human Geography* (pp. 268-275). Elsevier Ltd. doi: 10.1016/B978-008044910-4.01126-3

Lagorio-Chafkin, C. (2019). Dating apps are rife with unwanted sexual images. bumble's new a.i.-enhanced 'private detector' might change that. Retrieved 19 March 2020, from Inc website: https://www.inc.com/christine-lagorio/bumble-whitney-wolfe-herd-private-detector.html

Latessa, E.J., Lovins, B. (2014). Risk assessment, classification, and prediction. In: G. Bruinsma, & D. Weisburd, (Eds.), *Encyclopedia of Criminology and Criminal Justice* (pp 4457–4466). Springer. https://doi.org/10.1007/978-1-4614-5690-2_26

Lattimer, C. (2013). The Future of Geospatial Technologies in Securing Cyberspace. Retrieved January 18, 2020, from E-International Relations website: https://www.e-ir.info/2013/08/03/the-future-of-geospatial-technologies-in-securing-cyberspace/

Lee, J., & Downing, S. (2019). An exploratory perception analysis of consensual and nonconsensual image sharing. *Journal of Cybersecurity Intelligence & Cybercrime*, *2*(2), 23–43.

Lee, J. M., Hong, J. S., Yoon, J., Peguero, A. A., & Seok, H. J. (2017). Correlates of adolescent cyberbullying in South Korea in multiple contexts: A review of the literature and implications for research and school practice. *Deviant Behavior*, *10*, 1–16.

Lee, S.-S., Choi, K., Choi, S., & Englander, E. (2019). A test of structural model for fear of crime in social networking sites. *International Journal of Cybersecurity Intelligence & Cybercrime*, *2*(2), 5–22.

Leukfeldt, E. R., & Holt, T. J. (2019). *The Human Factor of Cybercrime* (E. R. Leukfeldt & T. J. Holt, Eds.). New York, NY: Routledge.

Leukfeldt, E., & Yar, M. (2016). Applying routine activity theory to cybercrime: A Theoretical and empirical analysis. *Deviant Behavior*, *37*(3), 263-280. https://doi.org/10.1080/01639625.2015.1012409

Louderback, E. R., & Roy, S. S. (2018). Integrating social disorganization and routine activity theories and testing the effectiveness of neighbourhood crime watch programs: Case study of Miami-Dade County, 2007-15. *British Journal of Criminology*, *58*(4), 968–992.

Madigan, S., Villani, V., Azzopardi, C., Laut, D., Smith, T., Temple, J. R., … Dimitropoulos, G. (2018). The prevalence of unwanted online sexual exposure and solicitation among youth: A meta-analysis. *Journal of Adolescent Health*, *63*(2), 133–141. https://doi.org/10.1016/j.jadohealth.2018.03.012

Marcum, C., Ricketts, M., & Higgins, G. (2010). Assessing sex experiences of online victimization: An examination of adolescent online behaviors using routine activity theory. *Criminal Justice Review*, *35*(4), 412–437.

McGuire, M., & Dowling, S. (2013). *Cyber crime: A review of the evidence*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/246749/horr75-summary.pdf

Mitchell, W. J. (1995). *City of bits: Space, lace and the Infobahn*. MIT Press. http://cumincad.architexturez.net//doc/und/oai-cumincadworks-id-4c7e

Myllymaki, P., Silander, T., Tirri, H., & Uronen, P. (2002). B-Course: A web-based tool for Bayesian and causal data analysis. *International Journal of Artificial Intelligence Tools*, *11*(3), 369–388.

Ngo, F., & Jaishankar, K. (2017). Commemorating a decade in existence of the international journal of cyber criminology: A research agenda to advance the scholarship on cyber crime. *International Journal of Cyber Criminology*. https://doi.org/10.5281/zenodo.495762

Ngo, F., & Paternoster, R. (2011). Cybercrime victimization: An examination of individual and situational level factors. *International Journal of Cyber Criminology*, *5*(1), 773–793.

Oswald, M., Grace, J., Urwin, S., & Barnes, G. C. (2018). Algorithmic risk assessment policing models: Lessons from the Durham HART model and experimental proportionality. *Information & Communications Technology Law*, *27*(2), 223–250.

Pratt, T., Holtfreter, K., & Reisig, M. (2010). Routine online activity and Internet fraud targeting: extending the generality of routine activity theory. *Journal of Research in Crime and Delinquency*, *47*(3), 267–296. https://doi.org/10.1177/0022427810365903

Prins, N., & Kingdom, F. (2018). Applying the model-comparison approach to test specific research hypotheses in psychophysical research using the Palamedes Toolbox. *Frontiers in Psychology*, *9*. http://www.qgso.qld.gov.au/about-statistics/survey-methods/index.php

Reyns, B. (2013). Online routines and identity theft victimization: Further expanding routine activity theory beyond direct-contact offenses. *Journal of Research in Crime and Delinquency*, *50*(2), 216–238.

Reyns, B. W. (2015). A routine activity perspective on online victimisation. *Journal of Financial Crime*, *22*(4), 396–411. https://doi.org/10.1108/JFC-06-2014-0030

Reyns, B. W., Fisher, B. S., Bossler, A. M., & Holt, T. J. (2018). *Opportunity and self-control: Do they predict multiple forms of online victimization?* https://doi.org/10.1007/s12103-018-9447-5

Reyns, B. W., Henson, B., & Fisher, B. S. (2011). Being pursued online: Applying cyberlifestyle–routine activities theory to cyberstalking victimization. *Criminal Justice and Behavior*, *38*(11), 1149–1169. https://doi.org/10.1177/0093854811421448

Rodriguez, J. A., Oduber, J., & Mora, E. (2017). Routine activities and cybervictimization in Venezuela. *Latin American Journal of Safety Studies*, *20*, 63–79.

Song, H., Lynch, M. J., & Cochran, J. K. (2016). A macro-social exploratory analysis of the rate of interstate cyber-victimization. *American Journal of Criminal Justice*, *41*(3), 583–601. https://doi.org/10.1007/s12103-015-9308-4

Song, Q., Guo, Y., & Shepperd, M. (2019). A comprehensive investigation of the role of imbalanced learning for software defect prediction. *IEEE Transactions on Software Engineering*, *45*(12), 1253–1269.

Smith, T., & Stamatakis, N. (2020). Defining cybercrime in terms of routine activity and spatial distribution: issues and concerns. *International of Cyber Criminology*, *14*(2), 433–459. https://dx.doi.org/10.5281/zenodo.4769989

Smith, T., & Stamatakis, N. (2021). Cyber-victimization trends in Trinidad & Tobago: The results of an empirical research. *The International Journal of Cybersecurity Intelligence and Cybercrime*, *4*(1), 46–63. https://doi.org/10.52306/04010421JINE3509

Sternberg, J. (2012). *Misbehavior in cyber places: The regulation of online conduct in virtual communities on the*. University Press of America.

Taylor, R. W., Fritsch, E. J., & Liederbach, J. (2019). Cybercrime and cyberterrorism (Fourth Edition). Hoboken, New Jersey: Pearson.

Tollenaar, N., & van der Heijden, P. G. M. (2019). Optimizing predictive performance of criminal recidivism models using registration data with binary and survival outcomes. *PLoS ONE*, *14*(3).

Trevizo, T. (2019). The three components of social engineering attacks. Retrieved 19 March 2020, from Hackernoon website: https://hackernoon.com/how-hackers-use-social-engineering-to-target-companies-zu1738k0

Vakhitova, Z., Reynald, D., & Townsley, M. (2016). Toward the adaptation of routine activity and lifestyle exposure theories to account for cyber abuse victimization. *Journal of Contemporary Criminal Justice*, *32*(2), 1–20. https://doi.org/10.1177/1043986215621379

van't Hoff-de Goede, M. S., Leukfeldt, E. R., van der Kleij, R., & van de Weijer, S. G. A. (2021). The online behaviour and victimization study: The development of an experimental research instrument for measuring and explaining online behaviour and cybercrime victimization. In M. Weulen Kranenbarg & R. Leukfeldt (Eds.), *Cybercrime in Context: Vol. I* (pp. 21–41). Springer International Publishing. https://doi.org/10.1007/978-3-030-60527-8_3

van Wilsem, J. (2013). Hacking and harassment - Do they have something in common? Comparing risk factors for online victimization. *Journal of Contemporary Criminal Justice*, *29*(4), 437–453. https://doi.org/10.1177/1043986213507402

Wang, S., & Yao, X. (2013). Using class imbalance learning for software defect prediction. *IEEE Transactions on Reliability*, *62*(2), 434–443.

Weulen Kranenbarg, M. (2018). *Cyber-offenders versus traditional offenders: An empirical comparison*. http://dare.ubvu.vu.nl/handle/1871/55530

Wikstrom, R. (2018). *The evolution of technology*. https://www.oxfordbibliographies.com/view/document/obo-9780195396607/obo-9780195396607-0010.xml

Wolak, J., Mitchell, K. J., & Finkelhor, D. (2007). Unwanted and wanted exposure to online pornography in a national sample of youth Internet users. *Official Journal of the American Academy of Pediatrics*, *119*(2), 247–257. https://doi.org/10.1542/peds.2006-1891

Yar, M. (2005). The novelty of 'cybercrime': An assessment in light of routine activity theory. *European Journal of Criminology*, *2*(4), 414. https://doi.org/10.1177/147737080556056

Yucedal, B. (2010). *Victimization in cyberspace: An application of routine activity and lifestyle exposure theories*. Kent State University. http://rave.ohiolink.edu/etdc/view?acc_num=kent127929098