User Modeling and User-Adapted Interaction manuscript No.
(will be inserted by the editor)

Exploring User Behavioural Data For Adaptive Cybersecurity

Joyce H. Addae · Xu Sun · Dave Towey · Milena Radenkovic

Received: date / Accepted: date

Abstract This paper describes an exploratory investigation into the feasibility of predictive analytics of user behavioural data as a possible aid in developing effective user models for adaptive cybersecurity. Partial Least Squares Structural Equation Modelling (PLS-SEM) is applied to the domain of cybersecurity by collecting data on users' attitude towards digital security, and analysing how that influences their adoption and usage of technological security controls. Bayesian-network modeling is then applied to integrate the behavioural variables with simulated sensory data from the web browser and other empirical data gathered to support personalized adaptive cybersecurity decision-making. Results from the empirical study show that predictive analytics is feasible in the context of behavioural cybersecurity, and can aid in the generation of useful heuristics for the design and development of adaptive cybersecurity mechanisms. Predictive analytics can also aid in encoding digital security behavioural knowledge that can support the adaptation and/or automation of operations in the domain of cybersecurity. The experimental results demonstrate the effectiveness of the techniques applied to extract input data for the Bayesian-based models for personalized adaptive cybersecurity assistance.

Keywords Cybersecurity · Behavioural analytics · Predictive modelling · Bayesian-inference · Adaptive assistance

J.H. Addae, X. Sun and D. Towey
Faculty of Science and Engineering, University of Nottingham Ningbo China, Ningbo, China.
E-mail: Xu.Sun@nottingham.edu.cn

M. Radenkovic
School of Computer Science, University of Nottingham, UK.

1 Introduction

The need to understand users within any human-computer system has long been identified as a critical design principle by HCI researchers and professionals. In recent years, there has been an increasing interest in the role users play in maintaining security within the digital economy. The adoption and appropriate usage of security mechanisms by home computer users (hereinafter referred to as users or HCUs) in particular have become a central concern for the usable security research community. Howe et al. (2012) described HCUs users as people who have not received any formal training to use computers but use them to support various tasks in non-work environments. Despite advances in cybersecurity technological solutions, most HCUs are still unable to effectively access them for the protection of their digital assets. As HCUs are increasingly targeted in security breaches (Crossler and Bélanger, 2014), there is a consensus among both cybersecurity researchers and key industry players about the urgent need to understand their cybersecurity behaviours and how best to enhance them.

To the same degree that efforts are being geared towards the security of cyberspaces, the need exist to equally make cybersecurity mechanisms accessible to the average user. People need to improve their security practices regularly which means they must be willing to learn and adopt the best security policies, and the mechanisms to ensure those policies. The National Institute of Standards and Technology (NIST) suggests that the best way of involving everybody is to create incentives that can motivate everybody within the cyber economy (Schwartz, 2011). Several usability studies on different types of security controls (e.g. firewalls, anti-virus) have illustrated how usability issues prevent end users from effectively leveraging them for their protection against security attacks (Cheung et al., 2001; Wong, 2008). Furnell and Clarke (2012) touched on anti-virus software usability and pointed out that users are faced with more complex interfaces due to the new trend of integrated internet security suits. Thus the consequent burden of understanding the full set of security functionalities provided through the surrounding options in web browsers has increased.

A reasonable assumption is that improved usability of cybersecurity mechanisms can serve as a major incentive for users to adopt better security controls and behaviour online. However, adoption of security systems remains problematic partially as a result of security researchers focusing less on the usability of systems within their social context (Church, 2008). It is becoming increasingly difficult to ignore the impact of individual differences and other socio-cultural variables when applying usable security design heuristics. Adaptive and/or personalised user interaction design have been proposed as possible ways of addressing usability and acceptability issues related to different user domain and contexts (Akiki et al., 2015; Mezhoudi et al., 2015; Bunt et al., 2004; Jason et al., 2010). Liu et al. (2016) for instance operationalised Personalized Privacy Assistant (PPA) and found improvement in the acceptability and usability of more suitable permissions recommended for mobile applications users. The

concept of a Personalised Adaptive Cybersecurity (PAC) here implies security and/or privacy functions for online applications would have to adapt not just to contextual changes but individual user preferences or needs as well. Thus individual differences can influence not just the perceived usability, but also the perceived risk, attitudes and acceptability of how a specified cybersecurity mechanism is designed (Dillon, 2001; Holden and Rada, 2011). There is the need to further understand the factors that affect users' perceived benefits of security control as well as the dimensions that wholly describe their attitude towards cybersecurity to better support the provision of PAC.

Acquiring knowledge about users and their perceptions is therefore a critical step in the process of improving the usability of cybersecurity mechanisms. Previous studies have identified useful insights into users' security behaviour by focusing on one or two influential factors from existing cognitive theories such as the Theory of Reasoned Action (TRA) (Lu et al., 2005), Theory of Planned Behaviour (TPB) (Ng and Rahim, 2005), Diffusion of Innovation theory (Conklin, 2006) and the Protection Motivation Theory (PMT) (LaRose et al., 2005; Milne et al., 2009). Our research model explores a wider variety of these dimensions by integrating TAM with PMT to explain and predict individuals' security behaviours. The model is further augmented by introducing *attitude to personal data* as part of the key determinants of intention to practice cybersecurity. As part of understanding users' attitude towards cybersecurity, this research focused on the inherent vulnerabilities of web browsers and how users interact with their built-in cybersecurity features (e.g. malware prevention, content filtering, private browsing, password manager, etc.) for security online.

The study combines and applies behavioural science and machine learning (ML) techniques to better support user modelling in personalized adaptive cybersecurity applications. An integrated model of cybersecurity adoption is developed and tested to determine influential factors which will impact on people's attitude to web browser security. Partial Least Squares Structural Equation Modelling (PLS-SEM) is applied to analyse empirical data collected using an online questionnaire-based survey. The empirical data and findings from the PLS-SEM model then serve as input for building the Bayesian-Network (BN) models for personalized adaptive cybersecurity (PAC). Thus the empirical experimentation with PLS-SEM assisted in determining which variables should be considered to support the personalization capability of the BN. The resulting components and structure of the Bayesian-network-based model illustrate how cybersecurity assistance can be intelligently provided.

2 ISSUES RELATED TO THE ADOPTION OF CYBERSECURITY CONTROLS

The cyberspace as an interconnection of web technology makes the sharing of digital information, products and services available to a broader range of participants. Cybersecurity is concerned with the protection of devices, appli-

cations, and data that connects to these web technologies through the internet from unauthorised access and usage. As more and more things are being attached to networks and connected to the internet (the era of Internet of Things (IoT)), it is becoming quite impossible to separate security on stand-alone computers from cybersecurity. Canongia and Mandarino Jr (2013) defined cybersecurity as:

"The art of ensuring the existence and continuity of the information society of a nation, guaranteeing and protecting, in Cyberspace, its information, assets and critical infrastructure."

This definition broadly views cybersecurity from a national perspective with no reference to personal safety or privacy within cyberspace. Cavelty (2014) made a distinction between national security and human security. He indicated that the former entails actions that affect social functions relying on IT and other critical infrastructures while the latter involves actions affecting acquired values like anonymity, privacy and other personal freedoms. Craigen et al. (2014) draw attention to the fact that, most definitions on cybersecurity miss the interdisciplinary nature of the field and tends to focus on the technical perspective. They posited the following definition after reviewing the literature and engaging with a multidisciplinary group of cybersecurity practitioners from varying backgrounds:

"Cybersecurity is the organization and collection of resources, processes, and structures used to protect cyberspace and cyberspace-enabled systems from occurrences that misalign de jure from de facto property rights."

Accordingly, their proposed definition is aimed at capturing the multi-dimensions of cybersecurity to promote more interdisciplinary approach in addressing emerging complex security challenges in cyberspace. Most often than not, cybersecurity strategies tend to be targeted at protecting national and/or organisational security. Adopting a top-down approach by focusing on the higher level especially the nation and big corporations have only led to individuals' security needs being undermined. There is, therefore, the need to systematically balance national and individual security. This research seeks to provide a holistic understanding of the effect of individuals' cybersecurity perceptions, attitudes and/or behaviours. This holistic view of human online security is not just relevant in determining appropriate policies but also in improving the usability of cybersecurity controls and increasing their acceptability among non-expert users.

Ross and Johnson (2010) classify security controls into three categories of management, operational and technical countermeasures that are applied to protect the confidentiality, integrity, and availability (CIA) of systems and the information they handle. Operational and managerial controls focus on security risks and incidents that are monitored and managed by people (e.g.usage policies, business continuity planning, training, etc.). Technical controls are mechanisms that use technology-based set-ups such as such as firewalls, antiviruses, user authentication, encryption technologies, Intrusion Detection Systems (IDS), etc. as system protection measures. As more and more people are able to gather, process, transfer or store sensitive commercial and personal

data over the internet, cybersecurity threats are also rapidly evolving. Achieving the aforementioned security goals of CIA are therefore as vital to the data protection needs of domestic internet users as to corporate and government networks. People generally want to be assured that, nobody will tamper with their information without their consent. People also want their data to be readily available and accessible at any point in time. Unfortunately, any form of data, be it corporate or personal, that is exposed to the internet are at risk of being compromised. Internet users, therefore, need to be able to easily adopt and correctly use available cybersecurity mechanisms in minimising such risks.

However, most non-security expert users find it quite challenging to understand and correctly configure available security mechanisms to avoid system breaches and cyber-attacks. Usability of security mechanisms have longed been identified by computer security researchers as critical to ensuring the protection of information systems (Whitten and Tygar, 1999; Zurko and Simon, 1996). This is because humans are a key component of any security system yet they are largely considered to be the weakest link of security. Mitnick and Simon (2011) pointed out that no matter how technically robust a security technology is, an attacker can breakthrough by exploiting the human element. A cybersecurity mechanism thus can lose its value if users are unwilling to adopt it or cannot use it due to poor usability hence impact negatively on the usability of internet based applications (Cambazoglu and Thota, 2013).

There has been little success with incorporating usability guidelines and standards into security-related interfaces. Security-related interfaces in the context of this research refer to the programs that allow users to manipulate security mechanisms on a system as well as control the effects of the users' manipulation and how security status is indicated. Although several consumer software are now successfully designed to be usable, security applications still seem to be lacking in their user-friendliness. A number of usable security studies (e.g. (Hof, 2015; Kainda et al., 2010)) have made a distinction between usability of security software and non-security software and argued that usable security design strategies should essentially consider and address inherent properties that make the security domain quite challenging. Accordingly, different interface design techniques are required for effective security-related interfaces and a special case exist when adopting prevailing general usability standards for security mechanisms.

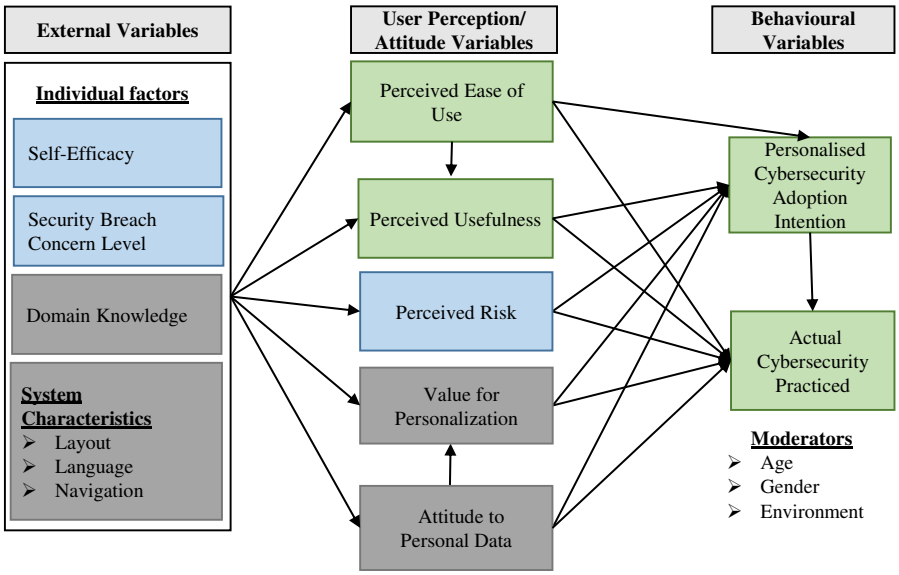
Although usability evaluation is critical in determining the proper implementation of security tools, it cannot fully explain and predict actual adoption and usage. Usability which is part of the overall system acceptability focuses on the extent to which the system can be used while acceptability is concerned with how well the system supports the needs and requirements of all stakeholders (Nielsen, 1994; Bordo, 2010). An acceptance model is thus required to explain and predict the acceptability of cybersecurity designs and implementation. Studies exploring acceptability in the field tend to focus on the factors influencing the acceptance of security policies and solutions within organizational context. Topa and Karyda (2015) recently reviewed the literature on employee security behaviour and classified the factors influencing them into

individual, organizational and technical. Accordingly, organizations aiming to improve security policy compliance are recommended to adopt a holistic approach that addresses issues related to all three category of factors. HCUs may, however, not be able to access such support that may enable them to improve their information security behaviour.

Recently, researchers have shown an increased interest in understanding users' security behaviour not only in the context of an organization, but also within non-corporate settings. Coventry et al. (2014) describes several possible scenarios affecting decision making within the context of cybersecurity differently than in other behavioural contexts. Omidosu and Ophoff (2016) highlighted the need for more studies into the security behaviours of non-corporate computer users based on their systematic review of the extant literature on information security behaviour in both organization and home contexts. Accordingly, a considerable knowledge gap exists where the security behaviour of individual cyber citizens operating within non-corporate context is concerned. The findings reported in this paper fills some of that gap by incorporating empirical evidence for actual cybersecurity related attitudes and behaviours into the development of user models for personalized adaptive cybersecurity. It is very critical to assess and ensure the usefulness as well as user friendliness of security tools developed for security inexpert users. In non-corporate environments, technical factors influencing security behaviour includes quality, performance and usability of the technological controls. Consequently, it is becoming increasingly important to focus on making the use of computer security tools effortless. The user model proposed and evaluated in this study for personalized adaptive cybersecurity is geared towards this goal of effortlessness.

3 THEORETICAL FRAMEWORK FOR PROPOSITIONS

Factors affecting the acceptance of various computer technologies have been a central research focus underlying the implementation of computer systems. Davis et al. (1989) determined that resistance to computer technologies aimed at increasing performance can be assessed and addressed with predictive behavioural models. This has led to the development of various models aimed at verifying the effect of identified factors on the acceptance of different kinds of technologies. These factors can be broadly categorised as individual, contextual and system characteristics. Two prominent models designed to predict specific security behaviour are the Technology Acceptance Model (TAM) and the Protection Motivation Theory (PMT) (Howe et al., 2012). Our research model, as shown in Fig. 1, integrates components from both these models, and includes other factors found to be possible determinants such as value for personalization and attitude to personal data. The model consists of three main components (External Variables, User Perceptions/Attitudes and Cybersecurity Behaviours), and explores how the identified external variables may influence perceived ease of use (PEOU), perceived usefulness (PU), perceived risk



Note: Components in green and blue boxes represent TAM and PMT components respectively.

Fig. 1 Predictive model for user cybersecurity behavioural Intentions

(PR), value for personalization (VFP), and attitude to personal data (APD); and how these can then predict an individual's cybersecurity intentions (BI) and actual cybersecurity practised and/or behaviour (ACB).

The TAM introduced by Davis et al. (1989) has since been adopted in studying and predicting user acceptance of various forms of technology (e.g.(Lee, 2009; Mun and Hwang, 2003; Abdullah et al., 2016)). This has led to a substantial amount of theoretical and empirical support being accumulated in its favour and is particularly regarded as being the most robust framework in explaining the adoption behaviours of information technologies (e.g. (Venkatesh and Davis, 2000)). In our work, TAM described the relationships between the users' acceptance, perceptions, and external variables . As shown in Fig. 1, user acceptance is examined by two cybersecurity behaviours — *intention to use* and *actual usage*. TAM identifies two considerations in an individual's decision to adopt an information system: Perceived Usefulness (PU) and Perceived Ease of Use (PEOU). Through these, TAM provides a theoretical framework for exploring the effect of external variables on beliefs that are internalised, and their subsequent impact on intentions and actual behaviour. According to TAM, PU and PEOU are the primary determinants of the intention to use and subsequent usage behaviour. PMT on the other hand measures the components of a fear appeal in determining the variables that impact on protection motivation in the form of behavioural intentions. Our study took TAM as a core theoretical foundation and extends it with PMT's cognitive mediation processes of threat and coping appraisal to develop a predictive model. The model is further augmented with two additional user insights related to

personalized digital security as primary determinants to empirically assess and predict the user's cybersecurity behaviour. These are Value for Personalization (VFP) and Attitude to Personal Data (APD). The ensuing paragraphs provide justification for the inclusion of these determinants in the research model, and related propositions.

3.1 Proposition Set 1: User Perceptions

Beliefs that users have about the usefulness of systems and their ease of use affect their intention to use and usage of the actual system. These perceptions have been extensively explored in previous technology acceptance research, and provide support for the following propositions with regards to web browser security controls (WBSC).

3.1.1 ***Perceived Usefulness (PU)***

- *H1: PU of WBSC is positively related to cybersecurity behaviour*

In the TAM, perceived usefulness refers to an individual's intrinsic belief about job related benefits, such as productivity, effectiveness and performance, associated with using a new technology. In the context of this research, PU refers to the degree to which a person believes web browser security settings would improve their protection against cyber-attacks. This definition captures both PU in the TAM model and response efficacy in the PMT model. Perceived usefulness has been reported to have a positive impact on the adoption and usage of information systems (Davis, 1989; Igbaria et al., 1997; Woon et al., 2005). Woon et al. (2005) found response efficacy (similar to perceived usefulness) significantly impacted home computer users' decision to protect their wireless network. Jeyaraj et al. (2006) reviewed and analysed empirical studies conducted on IT innovation adoption in the past decade and found perceived usefulness to be the best predictor for behavioural intention. The proposition here is that users are more likely to adopt security measures if they believe the security mechanism provided (in this case web browser security settings) are effective in making them cyber-secured.

3.1.2 ***Perceived Ease of Use (PEOU)***

- *H2: PEOU of WBSC is positively related to cybersecurity behaviours*
- *H3: PEOU of WBSC is positively related to PU*

PEOU refers to an individual's perception of the cost in terms of time and effort (mental and physical) involved in using a system (Davis, 1989). In previous studies, PEOU has been found to have both a direct and indirect effect on behaviour through its impact on PU of the technology being investigated. Suh and Han (2003) also discovered that both security concerns and usability dimensions have significant direct and interaction effects on the

adoption of smartphones for internet banking. Thus PEOU can influence users' attitudes towards a system application as well as their perception about the application's usefulness during use, therefore impacting on behaviour both explicitly and implicitly (Alharbi and Drew, 2014; Davis, 1989; Venkatesh and Davis, 2000). In the context of digital security, Ellis (2009, p. 41) noted that *"if security systems are burdensome, people may avoid using them, preferring convenience and functionality to security"*. There is also empirical support for response cost (similar to PEOU) having a significant negative impact on intention to enable security settings on a wireless network (Woon et al., 2005). It is therefore posited that WBSC that are difficult to use and require a lot of effort to configure will most likely be ignored and/or undervalued by users.

3.1.3 ***Perceived Risk (PR)***

– *H4: PR about WBSC is negatively related to cybersecurity behaviour*

Threat appraisal is a key aspect of the PMT, and refers to the beliefs that individuals form about perceived risk when they become aware of security threats. Their perceived risk is then evaluated against the effectiveness of the coping mechanisms that are made available. PMT includes rewards, severity and vulnerability to explain how threats are perceived. In our model, we consider rewards to be like PU and PR as the degree to which a user feels the uncertainties and negative effects of configuring some web browser security settings in areas of functional, time, information, physical and social risks (Lu et al., 2005). Perceived risk is considered to be a multi-dimensional construct in the literature consisting of different types of risk (e.g. physical, functional, social, etc.) (Jacoby and Kaplan, 1972; Kaplan et al., 1974; Lu et al., 2005). This study examined only five types of risk that are considered to be most relevant in the context of security technology adoption. Functional or performance risk describes the potential ineffectiveness of a security mechanism, hence failure to achieve the desired security goals. Time risk refers to the perceived time lost that may occur due to difficulty in configuring some security settings correctly. Information risk is the likelihood that instructions regarding the correct use of the security mechanism is inadequate/unreliable (risk associated with information failure). Physical risk means the extent to which an individual believes adopting the security technology can protect them against some form of loss, such as data, privacy or any component of the computer system (e.g. hard disk). Social risk describes the possibility that an individual may be worried about losing their reputation in a social group due to the adoption of a security control or technology.

Perceived risk has received considerable attention as a key predictor of consumer behaviour within the marketing literature (e.g.(Dai et al., 2014; Forsythe et al., 2006; Forsythe and Shi, 2003)). The construct has also been integrated into various predictive models and has been found to have significant impact on technology adoption behaviour (e.g. (Bélanger and Carter, 2008; Featherman and Pavlou, 2003; Lee, 2009; Özkan et al., 2010)). However,

far too little attention has been paid to it as a possible predictor of cybersecurity behaviour. Lu et al. (2005) is one of the few studies that examined and found that perceived risk impacted on intention to adopt an Online Anti-Virus through PU and Attitude towards use. More recently, Chang (2010) proposed an extended TAM model that includes risk-related factors for the prediction of managerial attitude towards the adoption of security technologies within an organisation. Based on findings of significant effects of PR in previous technology adoption studies, we propose that computer users perceiving high risk associated with WBSC will have a negative attitude towards cybersecurity in general.

3.1.4 ***Value for Personalization (VFP)***

– *H5: High VFP will positively affect intention to adopt personalized adaptive cybersecurity*

Personalization is the adaptation of services or products to the needs and/or preferences of a user. Whereas adaptive systems can be built to suit a categorized group of users, personalization takes it further to a more individual level. A number of online vendors now provide personalized products and services through online profiles of their consumers (e.g. eBay, Dell, Amazon etc.). Different ML techniques are adopted in constructing these consumer profiles to facilitate the provision of personalized products and services (Izquierdo-Yusta et al., 2015; Kim et al., 2001; Raghu et al., 2001). In marketing/e-commerce, personalization has been recognized as a significant influential factor in various consumer behavioural models (e.g.(Kim et al., 2001; Xu, 2006)). User-specific profiles allow online vendors to relate to their customers on individual basis, leading to improved customer satisfaction and loyalty. From the online users' point of view, however, the overall benefit of creating an online profile is the convenience of having different parts of their browsing experience personalized. Personalization can contribute to the effectiveness of technical security controls through improvement of user interactions and experience with the system. The nature of personalization may however differ for different types of user experience based on the context within which user profiles are defined and techniques used to create them. VFP in this study refers to the level of appreciation that a user has for all types of personalization possibilities within cyberspace. Because we recognise personalization as an important determinant of user experience and usage, assessing its significance within the structural model of a comprehensive set of other possible determinants of cybersecurity behaviours is imperative. The assumption here is that users who generally have positive attitudes towards the different types of personalized products and services available online are more likely to accept and use personalized adaptive cybersecurity.

3.1.5 ***Attitude to Personal Data (APD)***

– *H6: APD is positively related to cybersecurity behaviours.*

The construct of personal data (PD) and how it is perceived by individuals are identified in our research as critical components in explaining and predicting individuals' attitudes towards cybersecurity. Security in the digital world is often argued to be concerned with three main goals: confidentiality, integrity and availability. The confidentiality aspect of security is a basic privacy goal, and is concerned with the prevention of unauthorised access to sensitive data (Schneier, 2011). Because personal data is a common factor underlying the constructs of both security and privacy (Pearson, 2013), we have theorised that personal data, and how it is perceived by individuals, influences security related behaviour (Addae et al., 2016). APD here refers to the value people place on their data, and their tendency to adopt measures to protect it. It appears that many people now recognize and accept that an increasing part of life in the digital age involves disclosure of personal data. This does not, however, void the concerns that people may have about the actual use of the provided data (EU, 2011). Haddadi et al. (2015) highlighted the complex nature of personal data as a construct and how users' preferences and concerns differ based on context and sociological factors. To aid the inclusion of APD in cybersecurity behavioural research models such as ours, we conducted a study that explored APD dimensions towards the development of a personal data attitudes measurement scale (Addae et al., 2017). Based on findings from this study, we hypothesise that users who are generally protective towards their personal data are more likely to adopt cybersecurity measures.

3.2 Proposition Set 2: Moderating effects of external factors

Moderators are variables that modify the direction or strength of relationships between independent and dependent variables in a predictive model. Moderating variables alter relationships through interaction with either endogenous or exogenous variables, or by reallocating the error terms. Moderating factors have been shown to be very significant in various technology acceptance models as they can potentially improve the predictive validity of a model under investigation (Chin et al., 2003; Venkatesh et al., 2003). Moderators may also account for inconsistent factor findings in various user technology acceptance models (Sun and Zhang, 2006). Sun and Zhang (2006) examined the moderating effects in technology acceptance models and concluded that the exclusion of important moderators reflecting individual and contextual differences may account for lower explanatory power (predictive validity) and factor inconsistencies in previous findings. Accordingly, models that are extended with moderators such as gender, experience and cultural background, are more able to capture the intricacy of complex contexts. Prior empirical studies have identified several moderating factors involving differences in individual, organisational, cultural, context and system characteristics. In this study external variables reflecting both individual contextual differences and system characteristics are examined.

3.2.1 ***Individual Differences***

The acceptance and adoption of cybersecurity technologies may vary from one individual to another depending on differences in their characteristics. Individuals differ in terms of personality, level of experience, cognitive characteristics, background, and other demographics. Various aspects of individual differences have been examined in previous research (see below). Most studies have only considered a limited number of the variables pertaining to individual differences. A need for a holistic approach to cybersecurity user modelling that examines the relations between various aspects of individual differences and cybersecurity related factors thus remains. This study explores a wider variety of these individual characteristics and examines their impact on the perceived risk, usefulness, ease of use and attitude to personal data within the context of cybersecurity. As observed already, TAM is based on the fundamental principle that user perceptions mediate the influence of all other external factors that may influence technology acceptance and usage. The taxonomy of individual difference variables from previous research (e.g.(Alavi and Joachimsthaler, 1992; Bostrom et al., 1990)) was considered in identifying individual variables of interest that can be reliably measured alongside the cybersecurity behavioural variables in our predictive model. Consequently, individual difference variables in the model do not only cover the categories of demographics (age, gender, and environment) but also examine the descriptive characteristics of domain knowledge (DK), self-efficacy (SE) and users' security breach concern levels (SBCL) as external variables impacting on behavioural intentions towards cybersecurity.

Demographic Variables: Age has been found to moderate various factors in technology adoption and usage in the workplace (Morris and Venkatesh, 2000). In the area of cybersecurity, netizens between the ages of 18 and 25 were found to be more susceptible to phishing than other age groups (Kumaraguru et al., 2009; Sheng et al., 2010). The existence of gender differences in perception attributes has also been confirmed with a variety of IS diffusion models including TAM (Venkatesh and Morris, 2000). Shin (2009) also examined and found significant moderating effects of demographics variables, including income, on the interactions among attitudes and behavioural factors in their Unified theory of acceptance and use of technology (UTAUT) model for mobile payment. More recently, Anwar et al. (2017) observed gender differences in perceived computer security aptitudes and found that among employees from different organizations, men scored higher on self-reported cybersecurity behaviour than women. The usefulness and usability of a computer technology has also be found to be dependent on several contextual factors including the technical, organisational and physical environment within which it is adopted and used (Parsons et al., 2010; Maguire, 2001). (Gratian et al., 2018) for instance examined the influence of personality traits on cybersecurity behaviour intentions highlighting the mediating effects of environmental factors on individual differences in making security decisions. Consequently, we included three main demographic moderators (age, gender, and environment) in the

study analysis to examine the moderating effects of internet users' demographics on cybersecurity behaviour. The environment in our model refers to the physical location where participants in the study most often use their laptop/ desktop computers to access the internet.

– *H7: User demographic of age, gender and environment will moderate the relationship amongst the constructs of the proposed predictive model for cybersecurity behavioural intentions.*

Descriptive Characteristics: Security Breach Concern Level (SBCL) and Self-Efficacy (SE) are PMT constructs adapted to examine the mediating effects of a participant's protection motivation on cybersecurity behaviour. In PMT, a person's protection motivation is derived from two cognitive appraisal processes — threat appraisal and coping mechanisms. Apart from PR, fear arousal (the level of concern invoked by the threat) also captures threat appraisal within PMT models. Threat susceptibility has been found to predict security intentions in a number of PMT based models used to study safety behaviours (Tsai et al., 2016). An individual's assessment of the probability and consequences of a security threat is externalized as a security concern in this study. SBCL therefore refers to the degree of security threat an individual feels exists towards their personal safety online. The more convinced a user is about cybersecurity threats posing a significant damage to their personal digital assets, the more concerned they will be, resulting in a more positive attitude towards protection mechanisms. Hence we can assume that:

– *H8: High SBCL will positively influence attitude towards cybersecurity.*

Several studies have examined self-efficacy by integrating it with TAM (e.g.((Amin, 2007; Hasan, 2006; Hong et al., 2002; Ramayah, 2006)). Chau (2001) for instance, incorporated computer attitude and self-efficacy into the original TAM as external variables affecting perceived usefulness and ease of use. Related research into security behaviours finds support for the prediction that high self-efficacy positively influences attitude towards security countermeasures (Herath and Rao, 2009; LaRose et al., 2008; Milne et al., 2009; Woon et al., 2005). Self-efficacy has also been shown to influence adoption and usage of IT (LaRose et al., 2008; Compeau et al., 1999). In this study, cyber-citizens' self-efficacy influencing and/or predicting attitude towards cybersecurity behaviour is examined. The expectation is that individuals with high self-efficacy about their ability to optimise web browser security settings will have a more positive attitude towards cybersecurity than those with low self-efficacy. Therefore:

– *H9: High SE about WBSC will positively influence attitude towards cybersecurity.*

3.2.2 ***System Characteristics***

System Characteristics such as quality, interface design, speed/reaction time, etc., are some of the external factors proposed to have an indirect effect on

the acceptance and usage of information systems (IS) through user perceptions (Davis et al., 1989; Lin and Lu, 2000). For instance, Pituch and Lee (2006) included system characteristics as part of the external variables influencing e-learning use through perceived ease of use and usefulness. To do this, they solicited user ratings on three different aspects of e-learning systems — functionality, interactivity and response time. System characteristics especially functionality and interactivity were found to have the strongest total effect on the dependent variables of their model. The role of system characteristics in predicting technology acceptance through user perceptions has been explored in different contexts with a variety of system-specific features. According to Calisir et al. (2014), system characteristics such as security, reliability and speed, as a measure of system quality, influence expectation of the user experience level, hence increasing users' perceived ease of use. In one of the earliest studies conducted to measure user acceptance of information technology, the functional and interface characteristics of an electronic mail and a text editor were found to have significant direct effect on attitude towards usage (Davis, 1993). We identified three interface characteristics (layout, terminology and navigation) as critical for user interaction with WBSC in our study. Thus we argue that usability features such as clear, consistent layout and easy navigation will impact on a users' perception of WBSC, and hence the decision to accept or reject usage.

– *H10: The quality of WBSC interface design will positively influence attitude towards cybersecurity.*

4 THE EMPIRICAL STUDY

4.1 Research Design

The main research objective is to investigate influential factors which will impact on people's security behavioural intentions towards predictive analysis of a user's acceptance of personalized adaptive cybersecurity (PAC) for web browsers. A quantitative data collection and analysis approach similar to those employed by (Lee and Kozar, 2008; Lin, 2012; Venkatesh et al., 2003; Xu et al., 2008) in predicting behavioural intention was adopted. A field survey consisting of an online measurement instrument designed to collect data regarding factors influencing cybersecurity attitude and behaviours was conducted. The survey instrument was developed and administered using Qualtrics, an online survey tool. The measures were mostly adapted from previous studies that have explored various types of determinants of technology usage and specific computer security practices. For instance, the original measurement scales of TAM were adapted and modified to fit the context of WBSC usage. All construct measures were assessed with a 5-point Likert type scale ranging from "strongly agree" to "strongly disagree", except for the demographics and questions related to user preferences and/or experiences. Both positively and negatively worded items were included on the scales. Negatively worded items

were reverse-coded during the data analysis to ensure that a higher numbered response on the Likert scale would represent a higher positive attitude score, and vice versa.

The measurement instrument developed for the cybersecurity behavioural model has four main conceptual/ theoretical components consisting of Individual differences, user perceptions/ attitudes, behavioural variables and cybersecurity personalization components. The individual differences section consists of four exogenous driver constructs (i.e., IC, DN, SE and SBCL) as well as basic demographics such as age, gender, and environment. Thus the section measures participants' experience with web browser security (DK), self-efficacy (SE), personal preferences in terms of browser types and their respective user interfaces (IC) and their levels of concerns for security breach (SBCL). The second part of the instrument assessed participants' general attitudes towards cybersecurity from five main user perceptions: Ease of Use, Usefulness, Risk, Personalization and Personal Data. Hence TAM and PMT items (PU, PEOU, and PR) together with value for personalization (VFP) and attitude to personal data (APD) items represent key determinants of the endogenous target constructs.

To minimize respondent fatigue, the APD scale adopted from (Addae et al., 2017) was simplified by selecting only eight items based on overall cluster membership predictor importance of the APD factors as well as the reliability score of the measured items. Consequently, questions on Personal Data Awareness (PDA), Personal Data Protection (PDP) and Privacy Concerns (PC) measured reflectively, captured the major facets of the APD as a Type II second-order construct. This was to allow us to fully assess participant's attitude to personal data in relation to cybersecurity intention and usage behaviour. The third section (Behavioural variables) consists of measures for the target constructs of interest (i.e., BI and ACB) and asked whether the respondents had ever used or attempted to use web browser security functionalities as well as intentions toward personalized web browser security assistance. In the final part, items adapted from Xu et al. (2008) were used to collect participants' ratings on the personalization dimensions identified for the purposes of building a Bayesian-based network model for adaptive cybersecurity. All measured items included in the survey instruments are described with references to where they were adapted from in Appendix A. Items are grouped into the factors represented on the research model (Fig. 1) to ensure that a complete dataset is collected for hypothesis testing and data analysis.

4.2 Study Overview

The protocol analysis methodology is combined with observation and the system usability survey (SUS) in a SUT set-up to evaluate the usability of three commonly used web browser's security settings (GC, IE and FF). The primary goal is to identify underlying usability issues as well as merits of specific interface attributes prefered by users allowing us to propose design recom-

mendations for future web browser security interface and user interactions. In reviewing existing work on usable security, it has become very clear that several security labs and research studies have yielded valuable insights into user's security behaviour over the past decade. A major gap, however, is lack of studies that reflect users' actual security behaviour (e.g. have they optimised the security settings of their own personal computer?) within specific contexts.

To better understand users' web browsing security behaviour, it was deemed necessary to inspect study participants' actual browser security settings. Gathering real web browser security settings dataset on users' personal computers/laptops could help in measuring the impact of actual security behaviours exhibited by users on the security state of their personal computers. To this effect, in addition to the SUT adopted for this study, physical inspection of participants' browser security settings was carried out to better compare users' behavioural intentions and actual security behaviours. During the inspection, participants were interviewed on their motivation for choosing specific security configurations after accessing whether or not the said settings adequately meet their security/privacy goals.

4.3 Data Collection

A pilot test was first conducted with a mix of 50 university students and lecturers to ensure the survey instrument is comprehensible and valid. Feedback from the pilot was used to revise the final version. Convenience sampling was adopted to collect data with the questionnaire on two main university campuses in China and UK via emails. The questionnaire was also distributed online using various social media platforms including Facebook, Twitter, WeChat and LinkedIn. A total of 421 participants took part in the survey however, 37 incomplete and invalid responses had to be removed resulting in 384 usable responses. Alluding to the "ten times" rule of thumb on minimum sample size, the 384 valid responses meets the requirement for a PLS-SEM analysis. Accordingly, the 384 sample size is more than ten times the largest number of structural paths (six) directed at the most targeted construct in the model (ACB) and also more than ten times the number of indicators (six) used to measure the most complex construct in the model (APD) (Hair et al., 2011). The raw data were imported from Qualtrics and coded into the IBM SPSS statistic program for a descriptive analysis of respondent profiles.

4.4 Data Analysis

The settings and goals of this research favours the use of PLS-SEM based on the criterion identified by Hair et al. (2011). Using the SmartPLS 3 software, the Structural Equation Modelling (SEM) technique of Partial Least Squares (PLS) was employed to assess the theoretical model (Ringle et al., 2015).

PLS-SEM has proven to be a very valuable approach to developing and testing models in behavioural research. The approach is particularly versatile for extending models and running complementary analysis such as nonlinear relationships and moderation alongside hierarchical component models allowing for more complex model relationships to be tested. The PLS-SEM technique also deals with data related threats such as sample size, unobserved heterogeneity and normality in the dataset, to the validity of standard predictive analytics. PLS-SEM computes parameter estimates from least square estimation hence minimizing the demands on required assumptions about the dataset including the measurement scale for the data collection, sample size and residual distributions (Henseler et al., 2016). The PLS-SEM approach also allows for formative and multi-level constructs making it favourable for exploring possible causal relationships while avoiding parameter estimation biases typical of regression analysis. With reference to the two-step analytical process described in Hair et al. (2011), the measurement model was first evaluated for reliability and validity as the first step. The structural theory is then verified to determine the significant levels of the hypothesized relationships at the second step. The 2-step approach ensures inferences drawn from the structural relationship are based on validated measurement scales.

5 RESULTS

5.1 Sample Characteristics

Table 1 summarizes the characteristic and demographic distribution of participants. 51.3% of respondents were female and 48.7% males. The majority of respondents were students (70.3%) and fall within the age group of 18-24 (62.0%). A total of 99% of the respondents were educated well above 12^{th} grade and 72.7% earned an income of 1,000 to 8,000 US Dollars per month and 27.3% earned less than $1,000.

5.2 Reliability and Validity of the Measurement Model

The outer measurement model was examined for reliability and convergent validity with the same PLS software. All variance inflation factor (VIF) values are below 5.0 which suggests multicollinearity is unlikely to be a problem in our data analysis. Following guidelines in Hair Jr et al. (2016), VIF was further checked to determine if the first-order factors of APD were three distinct constructs. The VIF values of all constructs were below the conventional estimate of 5.0 with the highest being 3.195. Convergent validity for items in this study was assessed through their factor loadings in order to support the theory that sufficient convergent validity is achieved when the item measures the target latent construct. All the indicator items had significant path loadings at an alpha level of 0.01 and had high loading (> 0.5) on their respective parent constructs (Hair Jr et al., 2016; Urbach and Ahlemann, 2010).

Table 1 Respondent Profile

Demographic Variables		**Freq. N=384**	(%)	Σ %
Age	18 - 24 years	238	62.0	62.0
	25 - 34 years	93	24.2	86.2
	35 - 44 years	42	10.9	97.1
	< 45 years	11	2.9	100.0
Education	12th grade or less	4	1.0	1.0
	High school diploma	118	30.7	31.8
	Some college (no degree)	61	15.9	47.7
	Associate degree	9	2.3	50.0
	Bachelor's degree	86	22.4	72.4
	Graduate/ postgraduate	106	27.6	100.0
Employment	Employed for wages	74	19.3	19.3
	Self-employed	13	3.4	22.7
	Unemployed	22	5.8	26.8
	A homemaker	2	0.5	28.9
	A student	270	70.3	99.2
	Retired	3	0.8	100.0
Ethnicity	Asian/ Pacific Islander	29	7.6	7.6
	African/ Black	52	13.5	21.1
	Caucasian/ White	67	17.4	38.6
	Chinese	193	50.3	88.9
	Hispanic/ Latino	14	3.7	92.6
	Other	29	7.5	100.0
Environment	Home	205	53.4	53.4
	Corporate	111	28.9	82.3
	Public	68	17.7	100.0
Gender	Male	187	48.7	48.7
	Female	197	51.3	100.0
Income per month	Less than $1,000	105	27.3	27.3
	$1,000 to $5,000	165	43.0	70.3
	$5,000 to $8,000	66	17.2	87.5
	$8,000 or more	48	12.5	100.0

All of the outer loadings in the measurement model were above the minimum recommended level of 0.708 with the exceptions of ACB_4 (0.622) and PU_3 (0.651). We retained these two items in the measurement model because they were very close to 0.70 and the criteria for reliability and convergent validity were met (Hair Jr et al., 2016). For the higher order construct (HOC) APD, all paths from the three exogenous driver constructs were meaningful (PDA=0.20, PDP=0.68 and PC=0.21). All the values of composite reliability (CR) and average variance extracted (AVE) were well within the recommended threshold (Hair et al., 2010; Urbach and Ahlemann, 2010), with CR ranging from 0.81 to 0.95 and AVE from 0.62 to 0.86 (Table 2). The square root values of all the AVEs shown in bold and placed diagonally in Table 3 show that discriminant validity is well established. The distinctiveness of the contents captured by the three individual first-order factors of APD is demonstrated by their correlations which are well below the 0.80 boundary for establishing discriminant validity. In summary, the results of the statistical analysis support the

reliability, convergent and discriminant validity of the scales in our research model.

Table 2: Constructs Reliability and Validity

Latent Variables	**Scale Items**	**Loadings**	**CR**	**AVE**
Behaviour	ACB_1	0.90	0.90	0.69
	ACB_2	0.90		
	ACB_3	0.88		
	ACB_4	0.61		
Experience	DK_1	0.87	0.90	0.82
	DK_2	0.94		
Intention	BI_1	0.88	0.93	0.81
	BI_2	0.92		
	BI_3	0.91		
Interface	BI_1	0.82	0.90	0.70
	IC_2	0.84		
	IC_3	0.85		
	IC_4	0.82		
PD Awareness	PDA_1	0.79	0.87	0.69
	PDA_2	0.88		
	PDA_3	0.82		
PD Protection	PDP_1	0.79	0.83	0.62
	PDP_2	0.76		
	PDP_3	0.81		
Privacy Concern	PRI_1	0.93	0.92	0.86
	PRI_2	0.92		
Perceived Risk	PR_1	0.92	0.93	0.81
	PR_2	0.91		
	PR_3	0.88		
Personalization	VFP_1	0.93	0.95	0.86
	VFP_2	0.95		
	VFP_3	0.91		
Security Concerns	SBCL_1	0.82	0.93	0.78
	SBCL_2	0.90		
	SBCL_3	0.91		
	SBCL_4	0.91		
Self-Efficacy	SE_1	0.71	0.85	0.65
	SE_2	0.83		
	SE_3	0.87		
Usability	PEOU_1	0.80	0.81	0.68
	PEOU_2	0.85		
Usefulness	PU_1	0.85	0.84	0.64
	PU_2	0.87		
	PU_3	0.65		

Table 3 Inter-construct correlations and Fornell-Larcker Criterion Analysis

	ACB	DK	BI	IC	PDA	PDP	PR	VFP	PC	SBCL	SE	PEOU	PU
ACB	**0.83**												
DK	0.67	**0.90**											
BI	0.29	0.05	**0.90**										
IC	0.07	-0.13	0.67	**0.84**									
PDA	0.59	0.52	0.21	0.13	**0.83**								
PDP	0.78	0.73	0.13	-0.04	0.66	**0.79**							
PC	-0.14	-0.12	0.07	0.13	-0.01	-0.14	**0.90**						
VFP	0.65	0.56	0.14	-0.05	0.68	0.62	-0.01	**0.93**					
PC	0.68	0.82	0.05	-0.20	0.51	0.76	-0.11	0.61	**0.93**				
SBCL	0.34	0.41	-0.06	-0.21	0.49	0.45	0.26	0.56	0.47	**0.88**			
SE	0.53	0.55	0.02	-0.16	0.70	0.66	-0.06	0.64	0.59	0.65	**0.81**		
PEOU	0.49	0.27	0.76	0.53	0.21	0.33	-0.08	0.14	0.24	-0.08	0.09	**0.83**	
PU	0.54	0.47	0.12	0.00	0.86	0.64	0.01	0.66	0.52	0.62	0.78	0.13	**0.80**

5.3 Structural Model

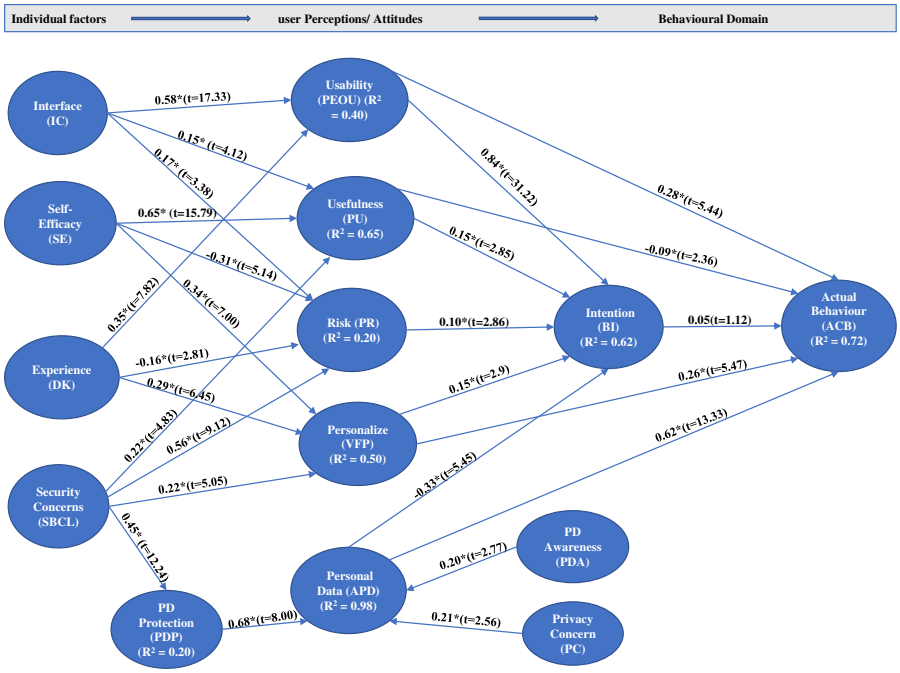


Fig. 2 Path Model and PLS-SEM estimates

Results of the structural model analysis are displayed in Fig 2. Paths in a PLS structural model can be interpreted similarly to standardized regression betas hence the overall predictive strength of the model is assessed by the explained variance in the endogenous variables. Tests of significance of all paths were performed following the bootstrap resampling procedure outlined in Garson (2012). In the model, R^2 value indicates the total variance explained by the endogenous latent variables. R^2 values of 0.19, 0.33, or 0.67 for endogenous

variables in the path model are described as weak, substantial or moderate respectively. A bootstrapping resampling procedure (5000 samples) was used to determine the significance of the path coefficients. Here, a multistage approach is adopted to facilitate the assessment of the APD impact on the two main endogenous variables in our extended-TAM model. The first model consisted of only the TAM and PMT Latent Variables (LV) as mediators and explained 59% and 49% of the variances in the two target constructs BI and ACB respectively. The value for personalization (VFP) factor was included in the second stage which increased the variance explained in ACB to 64%. The effect size (f^2) was assessed with the following equation:

$$\mathrm{f}^2 = \frac{\mathrm{R}^2_{\mathrm{included}} - \mathrm{R}^2_{\mathrm{excluded}}}{1 - \mathrm{R}^2_{\mathrm{included}}} \tag{1}$$

Where $\mathrm{R}^2{}_{included}$ and $\mathrm{R}^2{}_{excluded}$ are the R^2 values of the dependent LV when specific independent LV are included or excluded from the model. Values $\geq$0.02, $\leq$0.15, and $\leq$0.35 for f^2 respectively, represent small, medium and large effects of the exogenous LV (Hair Jr et al., 2016). The effect size of VFP on the endogenous construct ACB was large (0.40) and significant (p<0.001). Subsequently, the APD LV was added to the model and this second-order factor increased the R^2 of BI from 59% to 63%, and that of ACB from 64% to 74%. The effect size f^2 is large (0.47) and significant (p<0.001) for the predictive value of APD on ACB. There is also a small effect size (0.10) of APD on PCAI, which is significant at (p<0.005). Fig.2 provides the R^2values for each endogenous variable in the full PLS model along with path coefficients and associated t-values of the paths. To simplify the structural model and make it more legible, only paths that have significant relationships (indicated with asterisk on the path coefficient) are included in Fig 2. However, the insignificant path from BI to ACB is included since they are the two main output variables that mainly require further discussion.

The results (Table 4) show all five behavioural attitude determinants PEOU, PU, PR, VFP and APD, have significant effects on the behavioural intention to accept adaptive personalized cybersecurity. The five constructs together explain 63% of the variance in behavioural intention (BI). However only three of them were found to predict actual previous adoption of cybersecurity tools as the hypothesized path from BI was not statistically significant. The relationship between PU and ACB was significant ($\beta = -0.09, p < 0.05$), but not in the predicted direction. In this study, PEOU had the highest of the five path coefficients and a significant positive relationship with BI ($\beta = 0.84, t = 51.5, p < 0.001$) while APD appears to be the most important variable in the model predicting ACB ($\beta = 0.64, t = 12.87, p < 0.001$). Value for Personalization was also found to have a significant effect on BI and ACB hence justifying its importance in influencing users behavioural intention and attitude towards adaptive cybersecurity in a personal context.

In addition to evaluating the magnitude of the R^2 values as a criterion of predictive accuracy, the model's out-of-sample predictive power (Q^2) values

Table 4 Summary of Findings

Hypothesized Paths	Path Coefficients	t-values	f^2	Supported?
Experience (DK) –> Perceived Risk (PR)	-0.16	2.81**	0.02	Yes
Experience (DK) –> Personalization (VFP)	0.29	6.49***	0.12**	Yes
Experience (DK) –> Usability (PEOU)	0.35	7.75***	0.20**	Yes
Intention (BI) –> Actual Behaviour (ACB)	-0.05	1.09	0	No
Interface (IC) –> Perceived Risk (PR)	0.17	3.37**	0.04	Yes
Interface (IC) –> Usability (PEOU)	0.58	17.38***	0.54***	Yes
Interface (IC) –> Usefulness (PU)	0.15	4.09***	0.06**	Yes
Personal Data (APD) –> Actual Behaviour (ACB)	0.62	12.87***	0.47***	Yes
Personal Data (APD) –> Intention (BI)	-0.33	5.51***	0.10**	Yes
Perceived Risk (PR) –> Intention (BI)	0.10	2.90**	0.03	Yes
Personalization (VFP) –> Actual Behaviour (ACB)	0.26	5.38***	0.12**	Yes
Personalization (VFP) –> Intention (BI)	0.15	2.92**	0.03	Yes
Security Concerns (SBCL) –> PD-Protection (PDP)	0.45	11.97***	0.25***	Yes
Security Concerns (SBCL) –> Perceived Risk (PR)	0.56	9.08***	0.22***	Yes
Security Concerns (SBCL) –> Personalization (VFP)	0.22	5.00***	0.05**	Yes
Security Concerns (SBCL) –> Usefulness (PU)	0.22	4.89***	0.08**	Yes
Self-Efficacy (SE) –> Perceived Risk (PR)	-0.31	5.09***	0.06**	Yes
Self-Efficacy (SE) –> Personalization (VFP)	0.34	6.95***	0.11***	Yes
Self-Efficacy (SE) –> Usefulness (PU)	0.65	15.77***	0.70***	Yes
Usability (PEOU) –> Actual Behaviour (ACB)	0.28	5.24***	0.10**	Yes
Usability (PEOU) –> Intention (BI)	0.84	31.50***	1.59***	Yes
Usefulness (PU) –> Actual Behaviour (ACB)	-0.09	2.38**	0.01	No
Usefulness (PU) –> Intention (BI)	0.15	2.90**	0.03	Yes

Note: *p < 0.05. **p < 0.01. ***p < .001

were also examined. Here a sample re-use technique called blindfolding that omits part of the data matrix and uses the model estimates to predict the omitted part is applied to obtain the Q^2 values for the endogenous constructs (Hair Jr et al., 2016; Tenenhaus et al., 2005). Q^2 values greater than zero for specific reflective endogenous LV indicate the predictive relevance of the path model for that particular construct. Relatively values of 0.02, 0.15 and 0.35 indicate that the model respectively has a small, medium or large predictive relevance for the specified endogenous construct. Table 5 shows that all Q^2 values are considerably greater than zero, thus providing support for the cybersecurity behavioural model's predictive relevance for all the endogenous constructs especially having large predictive relevance ($Q^2 > 0.35$) for both of our two main target constructs (BI and ACB).

5.4 Moderating Effects

Further analysis was conducted to examine the moderating effects of demographic variables (Age, Gender) as well as the moderating influence of context of use (Home vs Corporate vs Public environments) on the hypothesized relationships in our model. When included in the model as control variables, age ($\beta = 0.17, t = 2.74, p < 0.05$), gender ($\beta = 0.08, t = 2.00, p < 0.05$) and environment ($\beta = -0.23, t = 3.71, p < 0.001$) were significantly associated with BI but none of them were significantly associated with ACB. Context of use (Environment) was negatively associated with BI, and it seems that users who more often access the web in public places are less interested in personalized adaptive cybersecurity. Although no specific hypothesis were declared for demographic variables of income education and ethnicity, their moderating effects were also explored in the analysis. However, since their effects were not

Table 5 Predictive accuracy R^2 and out of sample predictive power Q^2 values

Endogenous LV	R^2 Value	Q^2 Value
Behaviour (ACB)	0.74	0.48
Intention (BI)	0.63	0.48
PD Attitude (APD)	0.98	0.48
PD Protection (PDP)	0.20	0.12
Perceived Risk (PR)	0.20	0.15
Personalization (VFP)	0.50	0.40
Usability (PEOU)	0.40	0.26
Usefulness (PU)	0.65	0.39

statistically significant, they were not included in the results presented here for further analysis. PLS-SEM multi-group analysis (PLS-MGA) was conducted to determine whether significant differences are present between coefficients for the observed heterogeneity (age, gender and environment). PLS-MGA is used for comparing PLS model estimates across groups of data when the groups pre-exist (Hair Jr et al., 2016; Sarstedt et al., 2011).

To explore the moderating influence of gender, the data was split into Male (n=184) and Female (n=200) subgroups and separate analyses were computed for each group with the full model. Three subgroups were created for age 18-34 (n=169), 35-44 (n=139) and <44 (n=76), as well as for environment and/or context of use Corporate (n=111), Home (n=205) and Public (n=68). As the maximum number of arrows pointing to an endogenous variable in our model is five, a minimum of 5*10=50 observations per group is required according to the 10-times rule. The group-specific sample sizes for the three moderating variables can therefore be considered to be sufficient for the PLS-MGA. Since more than two groups are being compared in the case of age and environment, the Omnibus test of group differences (OTG) approach was applied as a first step to assess whether the path coefficients are equal across the three age and three environment groups.

The analysis (Table 6) yields F_R values ranging from 493.35 to 23289.54 for paths between the mediating variables and the two target variables for the environment groups. F_R values ranging from 686.06 to 10143.30 were yielded for the age group differences on direct paths to our target variables. The null hypothesis that the path coefficients across the three groups of age and that of the environment are the same can, therefore, be rejected. Thus the test rendered all differences among the groups significant at $p \leq 0.01$ suggesting at least one path coefficient differs from the remaining two across the three groups both in the case of age and environment.

Table 7 shows the differences in the path coefficient estimates of the group comparisons with respect to all the direct paths to the two DVs in the model, and provides the results of multigroup comparisons based on PLS-MGA and Welch-Satterthwait (W-S) Test. While the PLS-MGA is a non-parametric test for difference of group-specific results based on PLS-SEM bootstrapping results, the W-S is a parametric test that assumes unequal variances across

Table 6 Results of OTG for Age and Environment

Relationship	Group	B	SS-Between	SS-Within	F_R	p
Intention -> Behaviour	Environment 3	5000	206.64	0.10	2078.45	0.00
PD Attitude -> Behaviour	Environment 3	5000	614.86	0.17	3632.59	0.00
PD Attitude -> Intention	Environment 3	5000	544.58	0.06	8713.34	0.00
Perceived Risk -> Intention	Environment 3	5000	13.94	0.01	1277.02	0.00
Personalization -> Behaviour	Environment 3	5000	92.18	0.19	493.35	0.00
Personalization -> Intention	Environment 3	5000	11.93	0.02	669.01	0.00
Usability -> Behaviour	Environment 3	5000	253.81	0.12	2204.58	0.00
Usability -> Intention	Environment 3	5000	287.35	0.01	23289.54	0.00
Usefulness -> Behaviour	Environment 3	5000	419.16	0.05	8141.08	0.00
Usefulness -> Intention	Environment 3	5000	488.88	0.03	16709.36	0.00
Intention -> Behaviour	Age Group 3	5000	99.51	0.03	3627.87	0.00
PD Attitude -> Behaviour	Age Group 3	5000	20.39	0.02	1209.78	0.00
PD Attitude -> Intention	Age Group 3	5000	32.16	0.05	686.06	0.00
Perceived Risk -> Intention	Age Group 3	5000	28.84	0.02	1568.55	0.00
Personalization -> Behaviour	Age Group 3	5000	149.00	0.02	7804.65	0.00
Personalization -> Intention	Age Group 3	5000	87.93	0.03	2860.52	0.00
Usability -> Behaviour	Age Group 3	5000	59.70	0.04	1662.00	0.00
Usability -> Intention	Age Group 3	5000	500.49	0.05	9669.07	0.00
Usefulness -> Behaviour	Age Group 3	5000	32.01	0.01	3194.20	0.00
Usefulness -> Intention	Age Group 3	5000	265.61	0.03	10143.30	0.00

groups to determine the significance difference of group-specific PLS-SEM. As a one-tailed test, a typical cut-off level of significance for PLS-MGA results is >0.95 or <0.05, but the cut-off level can be set to >0.90 or <0.10 for smaller sample sizes. Slight differences between the PLS-MGA and W-S with respect to the significance of some of the group differences for specific relationships were observed. For instance, in the comparison of the Home and Public subsamples, the test rendered the relationship between Usefulness and Behaviour significant ($p \leq 0.10$) for PLS-MGA whereas this was insignificant in the W-S test ($p = 0.15$).

Table 8 summarizes the PLS-MGA results into a matrix to give a more simplified visual interpretation on determining significant effects based on demographics/ moderators. The findings support the assumption that the effects of the attitudinal variables on the two target constructs may be dependent on moderating variables. The results revealed significant differences in the group specific PLS path coefficients for the influences of the five mediating variables on ACB as well as BI on ACB. With regard to the age groups, there were significant differences between the groups for the relationship from BI to ACB, VFP to ACB, PEOU to BI, and PU to BI. In terms of Gender, the relationship between BI and ACB was negative and significant ($\beta = -0.25, t = 3.04, p < 0.05$) for males while non-significant for the females. This suggest that the unexpected negative relationship between BI and ACB that was found in the full sample results (Fig. 2) seems to be largely based on the male respondents. Two other significant differences between males and females subgroups are the relationships from PEOU to BI and from PU to BI. Although the relationship between PEOU and BI is positive and highly significant ($p < 0.001$) for both groups, the MGA results shows that usability is somewhat more important in determining BI among females than males. Meanwhile, the relationship be-

Table 7 Multigroup Comparison Test Results

Paths/Relationships	Comparison	PLS-MGA		Welch-Satterthwait Test	
		Path Coefficients - diff	p-Value	t-Value	p-Value
Intention -> Behaviour	Male vs Female	0.27	0.98	2.07	0.04
Usability -> Intention	Male vs Female	0.22	1.00	3.15	0.00
Usefulness -> Intention	Male vs Female	0.21	0.04	1.90	0.06
PD Attitude -> Intention	Home vs Corporate	0.44	0.94	2.15	0.03
Usability -> Intention	Home vs Corporate	0.15	0.02	2.02	0.05
Usefulness -> Intention	Home vs Corporate	0.20	0.05	1.75	0.08
Usefulness -> Behaviour	Corporate vs Public	0.38	0.04	1.75	0.08
Usefulness -> Intention	Corporate vs Public	0.43	0.99	2.59	0.01
Usability -> Intention	Home vs Public	0.27	0.00	3.12	0.00
Usefulness -> Behaviour	Home vs Public	0.26	0.07	1.47	0.15
Usefulness -> Intention	Home vs Public	0.22	0.95	1.59	0.12
Personalization -> Behaviour	Age <44 vs Age >34	0.21	0.97	1.88	0.06
Usability -> Intention	Age <44 vs Age >34	0.28	0.97	1.25	0.21
Usefulness -> Intention	Age <44 vs Age >34	0.23	0.04	1.71	0.09
Personalization -> Behaviour	Age 35-44 vs Age >34	0.15	0.92	1.43	0.16
Personalization -> Intention	Age 35-44 vs Age >34	0.13	0.91	1.30	0.19
Usefulness -> Intention	Age 35-44 vs Age >34	0.21	0.02	2.08	0.04
Intention -> Behaviour	Age 35-44 vs Age <45	0.19	0.07	1.44	0.15
Usability -> Intention	Age 35-44 vs Age <48	0.25	0.10	0.97	0.33

Note: significance levels <0.10 are highlighted in green and blue highlights indicates significance levels determined at >0.90.

tween PU and BI was positive and significant ($\beta = 0.22, t = 2.16, p < 0.05$) for males while insignificant for females.

For the Environment subgroups, there were significant differences for relationships from APD to BI, PEOU to BI, PU to ACB and PU to BI. Interestingly, the path from PU to ACB was negative and moderately significant ($\beta = -0.31, t = 1.83, p < 0.10$) for the public user group but insignificant for the corporate environment group. Thus usefulness in not important in predicting cybersecurity usage behaviour for those who mostly assess the internet within a corporate environment while most home and especially public users do not adopt cybersecurity tools though they may think they are useful. The differences in the environment groups for the relationship from PU to BI is also worth noting. Here PU seems to be more important in predicting positive BI of the public (β = .25, t=1.92, p<0.10) and home (β = 0.470.03, t=1.63, p>0.10) user groups than for the corporate group (β = -.0.17, t=1.73, p<0.10). We speculate that, due to the availability of professional IT services in corporate environments, these user group feel more secured when assessing the internet, and hence may not see the need for an easier to use cybersecurity mechanism. Whereas, those who mostly assess the internet from non-corporate environments may have no access to cybersecurity experts, and may thus perceive personalized adaptive cybersecurity as an easier way of ensuring their security and privacy online. It should also be noted that the influence of at-

Table 8 Multigroup Analysis Matrix

Paths/ Relationships	**Age**	**Gender**	**Environment**
Intention -> Behavior		Male** vs Female	
PD Attitude -> Intention			Home vs Corporate*
Personalization -> Behavior	<44 vs >34** 35-44 vs >34*		
Usability -> Intention	<44 vs >34** 35-44* vs <44	Male vs Female***	Home** vs Corporate Home*** vs Public
Usefullness -> Behavior			Corporate** vs Public Home vs Public*
Usefullness -> Intention	35-44** vs >34	Male** vs Female	Home*** vs Corporate Corporate vs Public**

Notes: Significant levels are associated with the subgroups with the highest PLS path coefficients where *p < 0.10., **p < 0.05., ***p < 0.001

titude to personal data was relatively consistent across the different groups, except in the case of the home subgroup where APD did not seem to be influential in determining their BI, although it is important in predicting their actual cybersecurity usage ($\beta = 0.47$ t=7.57, p<0.001). Thus attitude towards personal data appears to have a strong influence on cybersecurity behaviour and intentions across different user age and gender groups, and for both corporate and non-corporate users.

6 FRAMEWORK FOR PERSONALIZED ADAPTIVE CYBERSECURITY

Technology users differ in various ways in terms of goals, attitudes, and a host of individual characteristics and preferences that tends to influence their user experience. Design of user interaction for security and privacy technologies needs to accommodate different user goals and preferences. In the context of personal computing, web browsers provide a good platform to demonstrate the provision of adaptive and personalised cybersecurity configurations. Most current versions of web browsers allow users to sign in and synchronise their custom configurations across devices. This provides an opportunity to personalise default browser security settings as well as the presentation of alerts to improve their acceptance rate and reduce cognitive loads associated with digital security on a personal level. User model development is fundamental in an adaptive architecture for personalising user preferences. A user model consists of essential information and assumptions about users that can then be used to adapt the interaction of an application to specific individual users' needs. Building user models for adaptation and personalization often consists of two different approaches: one for the general user model and one for the personalised model. The general user model requires research and user experimentation to identify domain based generalization and classification of user interaction behaviours into specific user profiles. The personal model on the other hand will adapt new interactions based on observed data from an

individual user session. An individualised profile for adaptive cybersecurity, for instance, will include background information on an identified user, goals, preferences as well as information on the target device and web application. Thus, the amalgamation of the user and personal model enables adaptation to be personalized through the classification of users based on demographic information and several other contextual and individual characteristics.

Research has shown that the cybersecurity field requires a multidisciplinary approach to identifying and translating the salient factors influencing specific privacy and security decisions into more effective user models. While findings from PLS-SEM are useful to determine these salient factors and their dependencies, a lot of uncertainty remains in the attempt to recognize a user's goals from observations of behaviour. A powerful modelling technique developed by the artificial intelligence and ML community for effective reasoning in conditions of uncertainty in a sound mathematical manner is Bayesian Networks (BNs) (Nadkarni and Shenoy, 2004). BNs, also known as Belief Networks, provide a consistent way of replicating the essential features of plausible reasoning and have been successfully applied in the fields of medicine (e.g.(Sakellaropoulos and Nikiforidis, 2000)), marketing (Ahn and Ezawa, 1997) and business management. BNs are known to be particularly useful in handling uncertainties in user modelling for different kinds of application domains. They are typically used in situations where variables characterise the existence or absence of a quantifiable outcome.

In our study, BNs serve as an important tool to complement the user modelling process for adaptive cybersecurity. This is because the relationships between the many factors influencing a user's digital security decisions are mostly unclear. The empirical study conducted has allowed us to identify these influential factors and determine the directionality of their interactions. This makes directed edges in BNs more appropriate for our model than undirected edges in Markov Random Fields (Koller et al., 2007). In the context of cybersecurity where access control policies and privacy breach regulation are major concerns, accessing real-life behavioural data for research is always a challenge. Complementing the PLS-SEM used to derive additional domain knowledge with a Bayesian-based modelling technique is therefore an efficient way to deal with sparse and/or incomplete data. BNs allow us to intuitively infer the hidden states of the influential factors from the PLS-SEM through observation of their interrelating effects. With Bayes' rule, the inference problem can then be formulated as a case of resolving the probability of an unknown variable from values of variables observed in the empirical study. Apart from being able to describe uncertainty with BNs, there is the added advantage of being able to integrate different types of variables and related data within a single framework, and the flexibility of updating the models with new information at any given time.

The components of the framework (Fig. 4) were extracted from the empirical study described in section 4. Following the validation of the behavioural research model, the statistical analysis of data on the personalization dimensions proposed in Fig. 3 is used to support the construction of the Bayesian network

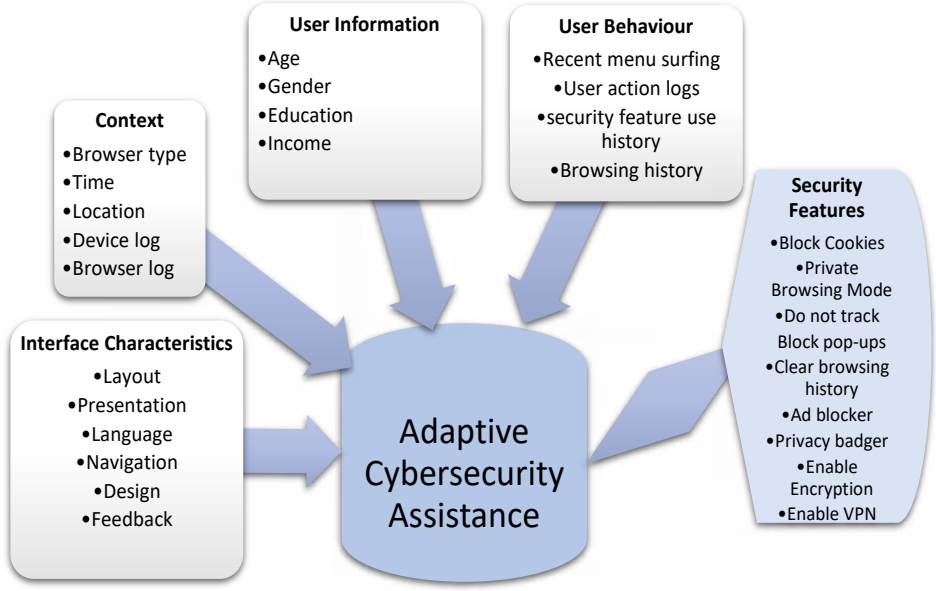


Fig. 3 Proposed dimensions of personal adaptive cybersecurity assistance (PACA)

model in our study. Nielsen and Jensen (2009) described Bayesian networks as a directed acyclic graph (DAG) consisting of a set of variables and a set of directed edges between variables. The structure is mathematically referred to as a DAG whereby variables represents events and a link from event A to B represents a causal relation whereby A is a parent of B and B is a child of A. Each variable B with parents $A_1, \ldots A_n$ has the potential table $P(B|A_1, \ldots A_n)$ which holds conditional probability distributions. Consequently, the proposed Bayesian networks will yield both quantitative measures in the form of conditional probability distributions as well as qualitative relationships between the components of personalized cybersecurity. The network of relationships in the BNs highlight how the various components interact with each other to influence the decision making process. Analysing the personalization components of cybersecurity with a Bayesian network can help in the characterization of various interactions between user context, profile, preferences and cybersecurity behaviour. To summarize, a user profile constituting personal information and observed behaviour, system characteristic variables (e.g. browser type, security settings etc.), and context of use are the factors being considered for personalized or adaptive cybersecurity within web browsers.

6.1 Structuring the Bayesian-Network-Based Model

Given the results from the empirical studies, we decided to build and assess Bayesian models that can determine a user's security/privacy needs and likelihood to adopt available cybersecurity solutions. Defining appropriate variables

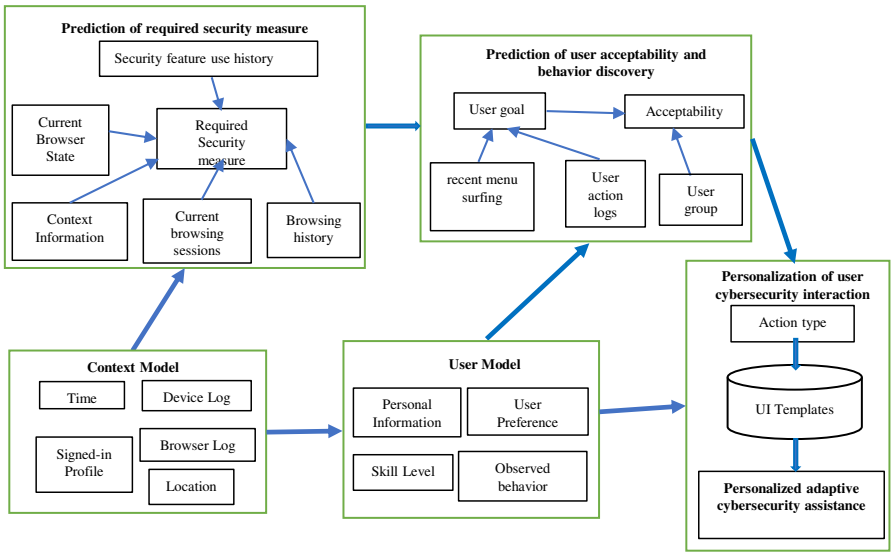


Fig. 4 Bayesian Network framework to infer and provide personalized adaptive cybersecurity assistance

and states of the identified variables are the building blocks of an effective user model. We wanted to achieve quality inferences from the models by incorporating contextual information, user's actions including queries (both current and previous), as well as the user's background and personal preferences. It is important to define the states of the variables included in the model clearly so users can be monitored and the conditional probabilities assessed. To establish a database for the BN model, the impacts of attributes related to web browser security features are analysed together with individual characteristics and context of use factors. Information from the survey instrument is used to produce a table values for the personalization component variables and used to calculate the prior probabilities of the model. To simplify the analysis, the levels within most of the variables were reduced. For instance the variable "location" was reclassified into three categories: home, public and corporate instead of the seven different locations measured with our survey scale (Home, School, Office, Public Transport, Cafes, Lecture rooms and Friend's house). Time of use was also set to peak and non-peak where peak time denotes periods where the user may normally be involved with official use of the internet for work or business related goals, and non-peak for pleasure or non-business related goals. Using a BN for analysis of responses to the cybersecurity personalization survey data can uncover and characterize the interaction of the personalization components and user's cybersecurity behaviour. Consequently, the output of a BN will reveal both the qualitative relationships between the attributes of personalized adaptive cybersecurity as well as the quantitative

measures in the form of conditional probability distributions of the factors' dependencies and interactions.

BNs can be modelled based on priori domain knowledge and/or training datasets (Heckerman et al., 1995). Since it is not easy to acquire cybersecurity related datasets on HCUs, we complemented the available dataset we gathered from the survey with domain knowledge to obtain the best combination of nodes for the BNs. Simulated datasets may not guarantee findings that fully reflect real-world data problems but they are widely adopted to garner deep insights and train machine learning models for various application domains (e.g. (Judson et al., 2008; Tsanas and Xifara, 2012)). Simulating aspects of network systems for instance has allowed researchers to overcome challenges of using data mining and ML for cyber analytics and to incorporate their intuition into building training models for intrusion detection (Buczak and Guven, 2016). In this context, the cybersecurity personalization factors extracted from the data analysis along with knowledge about web browser security features are used to develop the initial BN models for security-related tasks and subtasks. A complete model can then be obtained by combining several partial models developed from domain knowledge and simulated data focusing on representative nodes. For instance, if we know a relation exists between a user's security/privacy perceptions and expertise, these nodes can be connected by amending their conditional probability table (CPT) bounds of states accordingly. Thus conditional probability distributions (CPDs) of the form — the probability of B given A (p(B|A)), are used to encode the relationships between variables in the BN. For each node B, the likelihood that the variable will be in each possible state given its parents' node A states will be dependent on domain knowledge acquired from the empirical study as well as the frequency observed in both the measured variables and the simulated dataset (see Fig. 6). This approach ensures a prior distribution is estimated for the model parameters and used alongside those learned from data. This will help minimize incorrectly assigned probabilities if possible combinations are not observed in the training data (Gelman et al., 2014).

As an example, we considered a simple scenario of inferring the likelihood that a user will welcome the automatic blocking of a third party cookie. Considering observation of recent actions taken by the user on the web browser, example assumptions and reasoning that can be made here are that there might be a 50% chance of a random user accepting to block 3^{rd} party cookies if the user is completing an online form requiring sensitive information, but if the user is on a university campus, that probability will become 62% based on observations of user behaviour in similar context. Moreover, in considering the user's profile information, if the user was female the likelihood might decrease to 43%. Prior probability can also be indicated for a user based on age and frequency of using specific security features of the web browser. Qualitative inputs in terms of the variables and their dependencies are generated by domain knowledge and expert opinions. Quantitative data are subsequently generated using data analysis and model simulation.

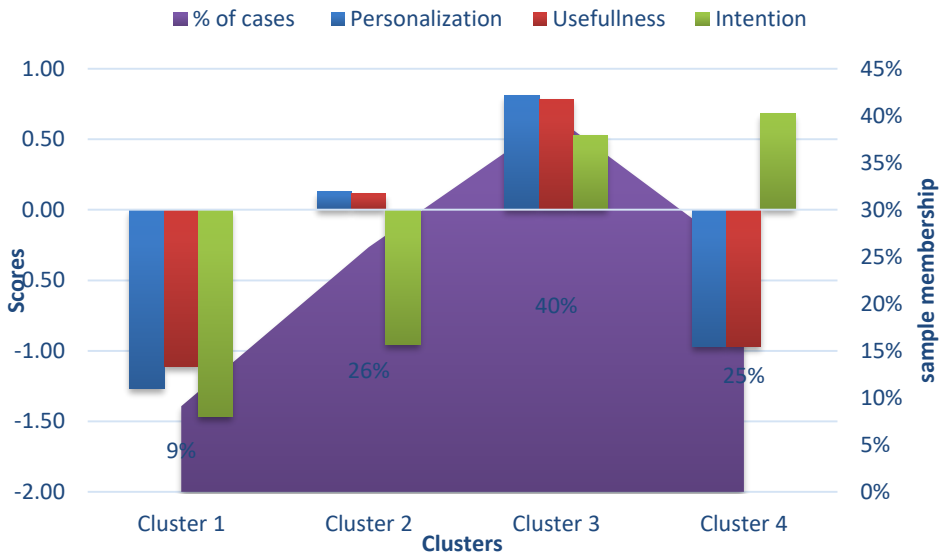
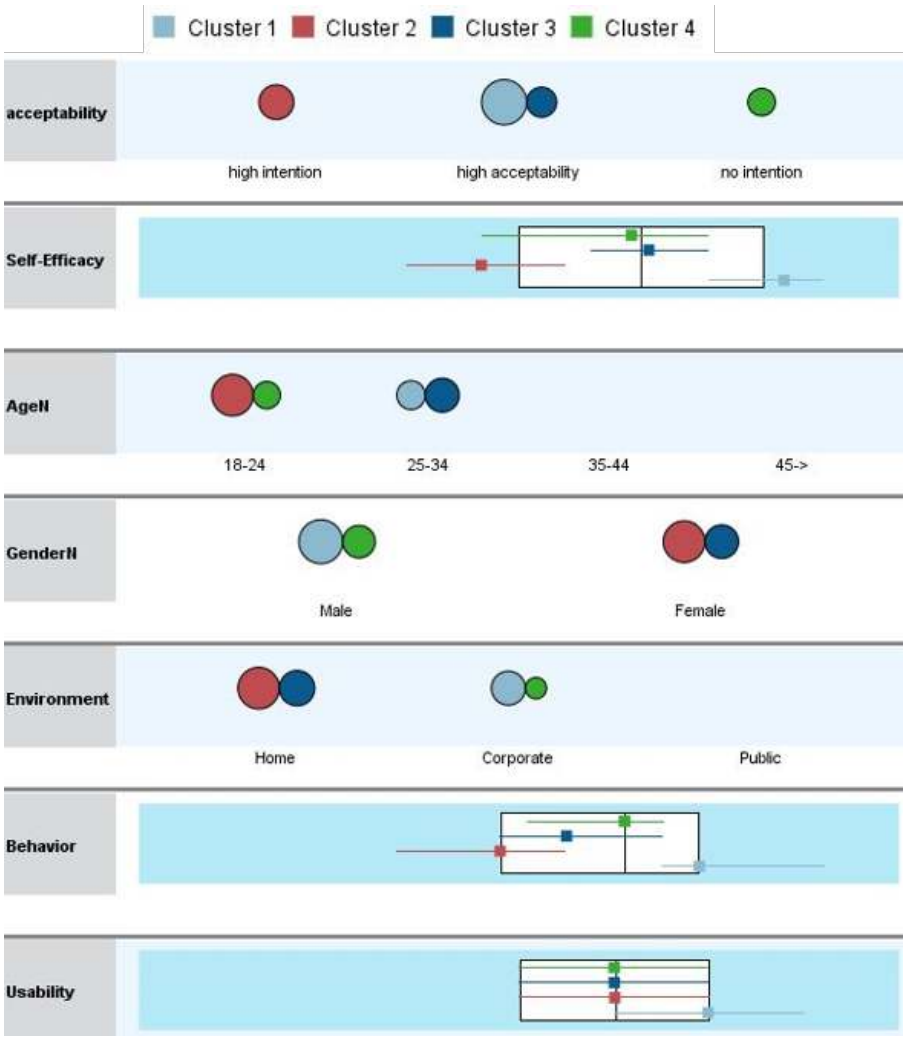


Fig. 5 Cluster groups based on acceptability factors

Table 9 Cluster distribution of respondents showing cluster centres sorted by overall cluster membership predictor importance

Cluster	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Description	Highest cluster group has high acceptability of PAC	2nd highest cluster group has high intention to adopt PAC	3rd highest cluster has moderate acceptability of PAC	The smallest cluster group has low acceptability of PAC
Size	31.8% (122)	26.8% (103)	23.4% (90)	18.0 (69)
Inputs	Acceptability 100%	Acceptability Intention(69.9%)	Acceptability 60%	Acceptability No intention (69%)
	Self-Efficacy μ=0.91	Self-Efficacy μ =-0.9	Self-Efficacy μ =0.02	Self-Efficacy μ =-0.16
	Age 25-34 (40.2%)	Age 18-24 (100%)	Age 25-34 (76.7%)	Age 18-24 (63.8%)
	Gender Male (94.3%)	Gender Female (100%)	Gender Female (75.6%)	Gender Male (91.3)
	Environment Corporate (56.6%)	Environment Home (100%)	Environment Home (83.3%)	Environment Corporate (39.1%)
Evaluation fields	ACB μ =0.84	ACB μ = -0.84	ACB μ = -0.24	ACB μ =0.07
	PEOU μ =0.45	PEOU μ = -0.16	PEOU μ =-0.26	PEOU μ =-0.23

To identify homogenous groups in the data set, a Two-Step clustering that is able to automatically determine the optimal number of clusters in a data set was adopted. Respondents were first clustered based on their factor scores on three acceptability variables determined from the PLS-SEM model (VFP, PU and BI) with k-means clustering. The results show that the majority of our participants have favourable consideration for PAC (Fig.5). The acceptabil-



a *Cluster 1 – high acceptability (100%) of PAC, score highest on self-efficacy and mostly access the web using corporate (56.5%) and public (43.3%) networks and more likely to have previously adopted a cybersecurity solution and found it user friendly.*

b *Cluster 2 – high intention to adopt PAC (69.9%) but scored the lowest on self-efficacy, mostly access the web using home network and less likely to have previously adopted a cybersecurity solution.*

c *Cluster 3 – Moderate acceptability (60%) with about 25% likelihood of rejection and 15% intention to adopt PAC. Moderate score on self-efficacy, mostly access the web with a home (83.3%) and sometimes corporate (16.7%) network and less likely to have previously adopted cybersecurity solutions.*

d *Cluster 4 – Low acceptability of PAC as 65.2% of these respondent group have no intention to adopt PAC and only 34.8% indicated high intention to adopt PAC. Low score on self-efficacy and access the web with all the three types of networks with about 39.1% likelihood for corporate, 33% likelihood for home and 27.9%. They are likely to have previously adopted a cybersecurity solution and not found it user friendly.*

Fig. 6 Visualization of cluster comparison

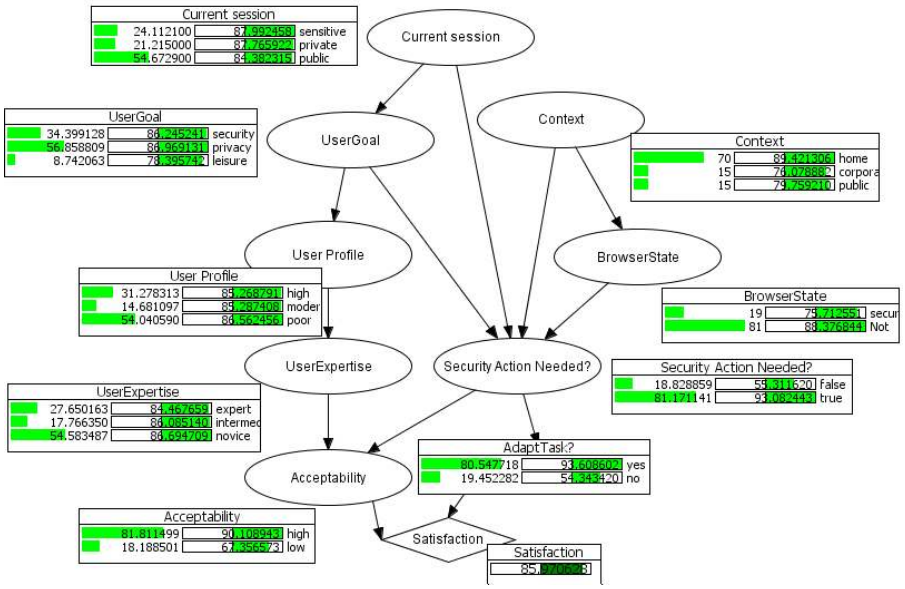


Fig. 7 The qualitative representation of the LMID used for decision making in PAC with priors based on data analysis

ity cluster membership was then combined with other adaptive cybersecurity personalization variables (such as, context/environment, gender, age etc.) for the Two-Step clustering and evaluated on self-reported previous use of cybersecurity tool (ACB) and PEOU. The results are summarized in Table 9 and visualised with Fig 6.

The joint probabilities are then used to specify the CPTs. To make a prediction from the BN, the model propagates the information at any given instance based on its structure and prior/conditional probabilities and provides the post-probabilities associated with the acceptability status (high or low) for a particular cybersecurity task to be adapted to the user's preference. Consequently, the BN-based decision engine will take output probabilities from both the context and user models as causal factors, together with the web browser configuration log and security task models to make a prediction. A decision status (e.g. block cookies, send alert or not) with an associated probability is arrived at after information is propagated in the BN. If the "acceptability" and "security need" probabilities are higher than a preset threshold, an automated security assistance in this scenario (auto block 3^{rd} party cookies or a preferred form of user alert) is provided for the user (see Fig. 7). Based on the evaluation of the level of satisfaction with the automated assistance provided, the user preference model is updated accordingly. Fig. 7 illustrates a personalized cybersecurity adaptive task limited memory influence diagram (LMID) built using domain knowledge with records from the survey data analysis.

Evaluation starts with the BN built based on the proposed LMID which, we refer to as the base BN. Next, data analysis is used to populate the CPT

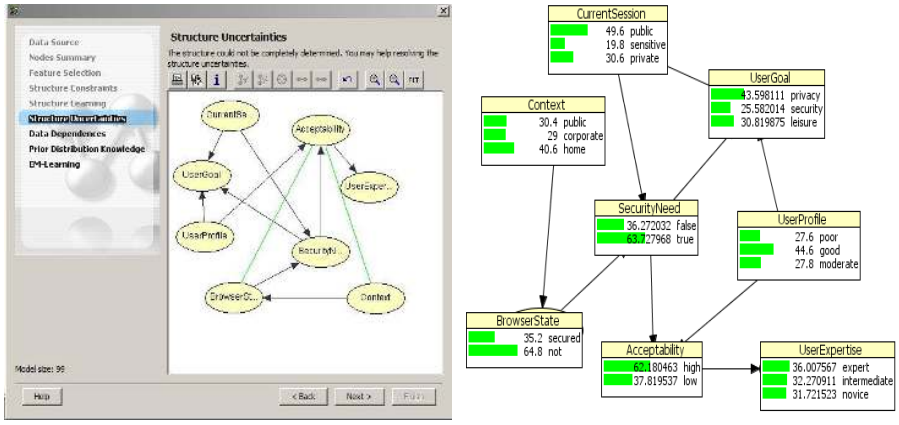


Fig. 8 The intermediate structure and CPT estimates for the Learned BN

of the base BN which is then used to generate a simulated data set. The Learning wizard in the Hugin Software (Madsen et al., 2005)[1] is then used to automatically discover a new network called intermediate BN from the simulated dataset (Fig. 8). Prior domain knowledge is then applied to resolve any uncertainties that may be present in the intermediate BN structure. With the discovered network and the generated database, parameter learning is carried out to specify a new CPT for the ensuing network called learned BN (Fig. 8). Finally, the performance results for the originally proposed BN structure are compared with corresponding BNs automatically discovered from both the survey and simulated data sets. The comparison evaluates the model's ability to produce applicable explanations in which relationships reflects adaptive cybersecurity as a domain from which the data were generated (Shaughnessy and Livingston, 2005). For prediction accuracy, we consider real usage scenarios in determining whether or not the levels of acceptability predicted are plausible. A receiver operating characteristic (ROC) curve is a fundamental measure of a model's performance for predicting specific states and the area under the ROC curve (AUC) allows the quality of the model to be expressed using a single value (Fig. 9). The analysis shows how well the predictions of the built BNs match the cases in the dataset. All in all, the probability changes among specified scenarios for the proposed BN parameters, were similar to those obtained by the learned BN.

7 DISCUSSION AND CONCLUSIONS

This research has two goals. One is to conduct an empirical study using a behavioural science approach to determine the factors influencing users' cybersecurity behavioural decisions. The second is to illustrate how Bayesian

[1] http://www.hugin.com

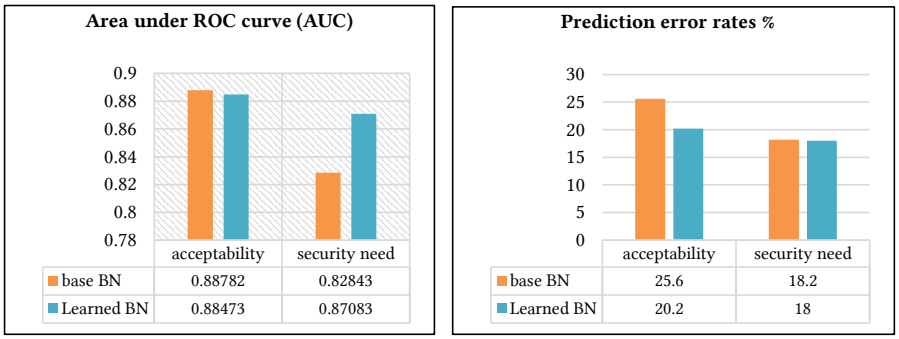


Fig. 9 Comparison of performance measures results for the base and learned BN structure

networks can be built by integrating findings from empirical studies into the ML approach of user and system modelling. To this end, a cybersecurity behavioural model was first introduced and empirically tested in this paper. The effects of five attitudinal constructs on cybersecurity behavioural intentions and behaviour were examined and in doing so, we (1) augmented the original TAM model with additional dimensions – Perceived Risk, Value for Personalization and Attitude towards Personal Data, and (2) evaluated the influence of three sample demographic variables on cybersecurity behavioural intentions. Although not all the hypothesized paths were found to be statistically significant, some interesting findings resulted from this study. The results suggest that both security-related perceptions and general external factors contribute to individual cybersecurity adoptive behaviour. The results also provide some evidence that these factors are moderated by the user's gender, age and the environment within which the internet is mostly accessed. Following the testing and verification of the behavioural model, those empirical findings were combined with the ML technique of Bayesian-network modelling for the development of a personalized adaptive cybersecurity framework. The research illustrated the model framework for personalized adaptive cybersecurity assistance.

The proposed behavioural model successfully explained most of the variance in the dataset. Similar to earlier studies (Alharbi and Drew, 2014; Venkatesh and Davis, 2000), TAM proved to be a useful theoretical framework to explore and explain factors influencing individuals' behavioural intentions towards technological innovations. Although the study confirmed the direct and indirect effects of some of the TAM constructs on cybersecurity behaviour, some of our results are inconsistent with prior research findings,and warrant further discussion. The results support prior empirical work that found a relationship between perceived ease of use, usefulness and behavioural intentions towards technological innovations (e.g.(Lee, 2009; Yiu et al., 2007)). However, contrary to suggestions from most prior studies that perceived usefulness is the main determinant of usage intentions in other IS research contexts (e.g. (Davis, 1989; Gefen et al., 2003; Jeyaraj et al., 2006)), our results show per-

ceived ease of use has a greater influence in predicting behavioural intentions in the context of cybersecurity.

Our results are however consistent with some previous studies that applied the TAM to some online applications, finding a strong effect of perceived ease of use on usage intentions and behaviour (e.g. (Castaneda et al., 2009; Gefen and Straub, 2000; Mun and Hwang, 2003; Özkan et al., 2010). The original TAM theorize PU have direct effect on behavioural intention while PEOU indirectly influences the intention through PU, hence depicting PEOU as a weak predictor of usage intentions. Our model, however, supports a direct effect of PEOU on behavioural intentions and usage of cybersecurity, and points to a greater significance of the ease of use factor in the context of digital security. A possible explanation of this finding could be attributed to the assertion that the effect of PEOU is dependent upon whether the type of use is intrinsic or extrinsic to the technology (Gefen and Straub, 2000). Thus, as our PEOU measured how easy the participant found it to learn and configure the security settings of their preferred web browser, the types of tasks involved here are intrinsic in that cybersecurity itself is an integrated component of the web browser with an interface that delivers the desired security and privacy control. Although our model did not support influence of PEOU on PU as theorized in the original TAM, PU did have a substantial impact on behavioural intention, which is consistent with extant findings in the TAM literature. The results confirms the direct relationship between PU and behavioural intention, though PEOU did not have a significant effect on PU and the proportion of the BI variance accounted for by PEOU far outweighed that of PU in our cybersecurity behavioural model. Also, PEOU is a significant determinant of self-reported actual cybersecurity practised in this study, while PU is a non-significant determinant. PEOU therefore provides a considerable explanatory power in the context of cybersecurity usage among home computer users.

Another major conclusion from this study that differs from the classical TAM-related studies is the role of behavioural intention. Based on findings from previous behavioural models, we had originally hypothesised that behavioural intentions will predict actual self-reported adoption of cybersecurity mechanisms. However, contrary to what the extant literature suggests, our dataset did not support this hypothesis, and upon further review, we realised this finding is reasonable in our specific research context. This is because our behavioural intention construct focused on personalized adaptive cybersecurity (PAC) rather than general cybersecurity, and hence participants may not yet have been exposed to it. Moreover, in the context of cybersecurity it is generally logical to expect the inherent inexplicability of security to impede actual usage though users may have intended to adopt available countermeasures. Thus factors such as complexity, inexperience and the secondary nature of security configuration to web browsing in general tend to deter adoption and usage of cybersecurity tools. Our findings however highlight the moderating role of gender as the effect of BI on actual self-reported usage was significant for males but not for females although the relationship was negative. More-

over, the effect of PEOU on BI was much stronger for the female subgroup, indicating that female netizens may be more hesitant to adopt difficult-to use cybersecurity controls.

The results also suggest that the strongest predictor of self-reported actual usage of cybersecurity controls is the second order construct of attitude towards personal data. Thus, participants who showed higher concern for the collection and use of their personal data were more likely to have attempted to, or actually adopted a cybersecurity countermeasure to ensure their privacy/security online. Interestingly, the relationship between the APD construct and BI to adopt personalized adaptive cybersecurity was negative, indicating that users who are very privacy conscious are less likely to adopt cybersecurity mechanisms that rely on their personal data to provide adaptivity. The relevance of the proposed BN framework is clearly supported by these findings. The BN-based models complements available machine-generated data (e.g. location, time, web logs, etc.) with domain knowledge data for the design of an intelligent cybersecurity mechanism. This minimizes the need to actively mine personal data to support prediction of acceptance of intended security task to be automated. The BN can also learn from real usage experience data to automatically update the probabilities when the inherent adaptability function is executed in practice. Users will be more satisfied if automated cybersecurity assistance provided is relevant to their primary cyber goals and delivered in a manner acceptable to them based on appropriate factors influencing their personal preferences. This requires a complex decision-making process involving predictive analysis of system and usage behaviour with a host of uncertainties. Building the predictive model with a BN which has the inherent facility to handle uncertainties will ensure a more effective provision of automated assistance that meets differing users' preferences compared to random automation of security tasks.

7.1 Implications for theory and practice

This study has implications for both researchers and practitioners of cybersecurity. From a research perspective, the extension of the TAM explained a significant amount of the variance in behavioural intention and adoption of web browser security controls. The study validates the significant role of user perceptions of ease of use, usefulness, risk, and personalization in predicting individual's intention to adopt PAC to achieve their security and privacy goals while accessing resources in the cyberworld with their web browsers. As discussed, the ease of use factor which is known to have a weaker influence in the classic TAM literature, takes on a much more significant role when it comes to cybersecurity control usage and intentions. This implies that individuals who normally disregard cybersecurity countermeasures may have the intention of adopting PAC if they realize that it will be useful and easy to do so. The study introduced additional constructs from protection motivation theory and personal data research that better reflect the complex context of

cybersecurity which encompasses digital security and privacy in its entirety. The findings from the PLS-SEM generally support the importance of the additional constructs, especially attitude towards personal data in predicting adoption behaviour in the domain of cybersecurity. Consequently the findings from the empirical behavioural study provide theoretical contributions in the area of cybersecurity acceptance and usage. This is with respects to both re-validation and extension of past theoretical framework as applied to the new context of security behaviour modelling. The findings from this research therefore add substantially to our understanding of cybersecurity behavioural intentions and personalization dimensions.

The findings also have implications for practice and design as it can inform several aspects of improving the usability of cybersecurity mechanisms. This study suggests that, cybersecurity mechanisms targeted at HCUs need to be very usable with minimal demand on cognitive resources. The study also endorses the value of incorporating data and privacy protection into system design right from the onset, which are the underlining principles of recent *privacy-by-design* projects. For instances both the new EU GDPR and PRIPARE projects (Notario et al., 2015; Huth, 2017) highlight the need for *privacy-by-design*. However, almost no direct comprehensive studies exist on non-expert users' privacy preferences towards adaptive cybersecurity in non-corporate environments. Our proposed predictive model for providing personalization takes on individual's disposition to their personal data into account. This provides a framework for incorporating data privacy controls from the design stage. In so doing, personalization is provided at the preferred level for each individual. Thus, our design framework will facilitate the process of determining and limiting access to such data that a user might consider too sensitive in providing adaptive cybersecurity.

In summary, the contributions of the research presented in this paper are both novel and significant paving the way for further empirical study on personalised adaptive cybersecurity in the public domain. As research exploring the provision of PAC for HCUs is still in its infancy, the issues discussed in this paper fill a fundamental gap in the current literature. The empirical approach of PLS-SEM has been used to explore the statistical relationship between various cybersecurity behavioural input variables to predict two output variables (BI and ACB). This provided essential insights into the specific issue of predicting user behavioural intentions toward the provision of PAC assistance. An example BN-based framework is developed to illustrate how these insights can be incorporated into building PAC user models. The BN is thus built using a range of diverse input variables including behavioural, context and simulated web browser features to demonstrate the provision of PAC in web browsers.

7.2 Limitations and future research

It is important to highlight the limitations of the studies presented in this paper. Notably, generalization will need to be done with caution as the uni-

versity students and staff were used as a convenience sample. The data set has however been successfully used to provide empirical evidence for the usefulness of predictive analysis of users' behavioural data to the design of adaptive cybersecurity. Although we had some measured data sets (such as self-efficacy), observation data such as the actual level of user's cybersecurity expertise and security state of the browser were not available during the development of the BN-based models. The data used for the BNs were thus obtained by simulating aspects of the proposed framework. Therefore, the possibility that the result may have some bias can not be overlooked. Nevertheless, the simulated data provided a good indication of the likely percentage change to determine the underlying trend that may be present in real-world data scenarios. Moreover, the primary goal of this work is to demonstrate the incorporation of insights gained from behavioural empirical studies into training machine learning models that can better support prediction and decision-making in the domain of cybersecurity. This work represents a first-step towards the design and development of a user friendly adaptive cybersecurity which adheres to the concept of *privacy-by-design*. We also recognize that future research is needed to fully evaluate the proposed BN-based models. This will require additional dataset and further optimisation and testing before implementation. Although the preliminary results using simulated data are promising, no real trial data was available for a full validation. However, since our goal is to illustrate the feasibility of the approach rather than validate, we sought to evaluate the model on prediction accuracy. Thus considering real usage scenarios, we are able to determine the underlying trends in predicting acceptability and usability with the set of parameters identified.

Continuing with our combined approach of empirical studies and modelling technique, we determined three future research directions. First, more broader samples are required to replicate the behavioural model and validate inferences that can be made based on either a PLS or Covariance-based SEM results. Secondly, more factors that will influence cybersecurity personalization need to be considered and their appropriate measure determined so they can be incorporated into the Bayesian network system. Third, the Bayesian-based models need to be implemented in a prototype web browser for practical evaluation of the function and further optimization with real sensory data.

Acknowledgment

The authors acknowledge the financial support from the International Doctoral Innovation Centre (IDIC), Ningbo Education Bureau, Ningbo Science and Technology Bureau, China's MoST and The University of Nottingham. This work was also supported by the Horizon Digital Economy Research, UK.

References

Abdullah, F., Ward, R., and Ahmed, E. (2016). Investigating the influence of the most commonly used external variables of tam on students' perceived ease of use (peou) and perceived usefulness (pu) of e-portfolios. *Computers in Human Behavior*, 63:75–90.

Addae, J., Radenkovic, M., Sun, X., and Towey, D. (2016). An augmented cybersecurity behavioral research model. In *Computer Software and Applications Conference (COMPSAC), 2016 IEEE 40th Annual*, pages 602–603. IEEE.

Addae, J. H., Brown, M., Sun, X., Towey, D., and Radenkovic, M. (2017). Measuring attitude towards personal data for adaptive cybersecurity. *Information & Computer Security*, 25(5):560–579.

Ahn, J.-H. and Ezawa, K. J. (1997). Decision support for real-time telemarketing operations through bayesian network learning. *Decision Support Systems*, 21(1):17–27.

Akiki, P. A., Bandara, A. K., and Yu, Y. (2015). Adaptive model-driven user interface development systems. *ACM Computing Surveys*, 47(1).

Alavi, M. and Joachimsthaler, E. A. (1992). Revisiting dss implementation research: A meta-analysis of the literature and suggestions for researchers. *Mis Quarterly*, pages 95–116.

Alharbi, S. and Drew, S. (2014). Using the technology acceptance model in understanding academics' behavioural intention to use learning management systems. *International Journal of Advanced Computer Science and Applications*, 5(1):143–155.

Amin, H. (2007). Internet banking adoption among young intellectuals. *Journal of Internet Banking and Commerce*, 12(3):1–13.

Anwar, M., He, W., Ash, I., Yuan, X., Li, L., and Xu, L. (2017). Gender difference and employees' cybersecurity behaviors. *Computers in Human Behavior*, 69:437–443.

Bélanger, F. and Carter, L. (2008). Trust and risk in e-government adoption. *The Journal of Strategic Information Systems*, 17(2):165–176.

Bordo, V. (2010). Overview of User Acceptance Testing (UAT) for Business Analysts (BAs).

Bostrom, R. P., Olfman, L., and Sein, M. K. (1990). The importance of learning style in end-user training. *MIS Quarterly*, pages 101–119.

Buczak, A. L. and Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications Surveys & Tutorials*, 18(2):1153–1176.

Bunt, A., Conati, C., and McGrenere, J. (2004). What role can adaptive support play in an adaptable system? In *Proceedings of the 9th international conference on Intelligent user interfaces*, pages 117–124. ACM.

Calisir, F., Altin Gumussoy, C., Bayraktaroglu, A. E., and Karaali, D. (2014). Predicting the intention to use a web-based learning system: Perceived content quality, anxiety, perceived system quality, image, and the technology acceptance model. *Human Factors and Ergonomics in Manufacturing &*

Service Industries, 24(5):515–531.

Cambazoglu, V. and Thota, N. (2013). Computer science students' perception of computer network security. In *Learning and Teaching in Computing and Engineering (LaTiCE)*, pages 204–207. IEEE.

Canongia, C. and Mandarino Jr, R. (2013). Cybersecurity: The new challenge of the information society. *Crisis Management: Concepts, Methodologies, Tools, and Applications*, page 60.

Castaneda, J. A., Frías, D. M., and Rodríguez, M. A. (2009). Antecedents of internet acceptance and use as an information source by tourists. *Online Information Review*, 33(3):548–567.

Cavelty, M. D. (2014). Breaking the cyber-security dilemma: Aligning security needs and removing vulnerabilities. *Science and engineering ethics*, 20(3):701–715.

Chang, A. J.-T. (2010). Roles of perceived risk and usefulness in information system security adoption. In *Management of Innovation and Technology (ICMIT), 2010 IEEE International Conference on*, pages 1264–1269. IEEE.

Chang, P. V.-C. (2004). The validity of an extended technology acceptance model (tam) for predicting intranet/portal usage.

Chau, P. Y. (2001). Influence of computer attitude and self-efficacy on it usage behavior. *Journal of organizational and end user computing*, 13(1):26.

Chellappa, R. K. and Sin, R. G. (2005). Personalization versus privacy: An empirical examination of the online consumer's dilemma. *Information technology and management*, 6(2):181–202.

Cheung, J., Li, S., Totolici, A., and Zheng, P. (2001). Usability analysis of sophos antivirus.

Chin, W. W., Marcolin, B. L., and Newsted, P. R. (2003). A partial least squares latent variable modeling approach for measuring interaction effects: Results from a monte carlo simulation study and an electronic-mail emotion/adoption study. *Information systems research*, 14(2):189–217.

Church, L. (2008). End user security: The democratisation of security usability. *Security and Human Behaviour*.

Compeau, D., Higgins, C. A., and Huff, S. (1999). Social cognitive theory and individual reactions to computing technology: A longitudinal study. *MIS quarterly*, pages 145–158.

Conklin, W. (2006). *Computer security behaviors of home pc users: a diffusion of innovation approach*. The University of Texas at San Antonio.

Coventry, L., Briggs, P., Blythe, J., and Tran, M. (2014). Using behavioural insights to improve the public's use of cyber security best practices. *Gov. UK report*.

Craigen, D., Diakun-Thibault, N., and Purse, R. (2014). Defining cybersecurity. *Technology Innovation Management Review*, 4(10).

Crossler, R. and Bélanger, F. (2014). An extended perspective on individual security behaviors: Protection motivation theory and a unified security practices (usp) instrument. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, 45(4):51–71.

Dai, B., Forsythe, S., and Kwon, W.-S. (2014). The impact of online shopping experience on risk perceptions and online purchase intentions: does product category matter? *Journal of Electronic Commerce Research*, 15(1):13.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, pages 319–340.

Davis, F. D. (1993). User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. *International journal of man-machine studies*, 38(3):475–487.

Davis, F. D., Bagozzi, R. P., and Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, 35(8):982–1003.

Dillon, A. (2001). User acceptance of information technology.

Ellis, G. (2009). NAE grand challenges for engineering. *IEEE Engineering Management Review*, 1(37):3.

EU (2011). Attitudes on data protection and electronic identity in the european union. *Eurobarometer Special Surveys*, 359.

Featherman, M. S. and Pavlou, P. A. (2003). Predicting e-services adoption: a perceived risk facets perspective. *International journal of human-computer studies*, 59(4):451–474.

Forsythe, S., Liu, C., Shannon, D., and Gardner, L. C. (2006). Development of a scale to measure the perceived benefits and risks of online shopping. *Journal of interactive marketing*, 20(2):55–75.

Forsythe, S. M. and Shi, B. (2003). Consumer patronage and risk perceptions in internet shopping. *journal of Business research*, 56(11):867–875.

Furnell, S. and Clarke, N. (2012). Power to the people? the evolving recognition of human aspects of security. *Computers & Security*, 31(8):983–988.

Garson, D. (2012). Partial least squares: Regression and path modeling. *Asheboro, NC: Statistical Publishing Associates*.

Gefen, D., Karahanna, E., and Straub, D. W. (2003). Trust and tam in online shopping: An integrated model. *MIS quarterly*, 27(1):51–90.

Gefen, D. and Straub, D. W. (2000). The relative importance of perceived ease of use in is adoption: A study of e-commerce adoption. *Journal of the association for Information Systems*, 1(1):8.

Gelman, A., Carlin, J. B., Stern, H. S., and Dunson, D. B. (2014). *Bayesian data analysis*, volume 2.

Gratian, M., Bandi, S., Cukier, M., Dykstra, J., and Ginther, A. (2018). Correlating human traits and cyber security behavior intentions. *computers & security*, 73:345–358.

Haddadi, H., Howard, H., Chaudhry, A., Crowcroft, J., Madhavapeddy, A., and Mortier, R. (2015). Personal data: Thinking inside the box. *arXiv preprint arXiv:1501.04737*.

Hair, J. F., Black, W. C., Babin, B. J., and Anderson, R. E. (2010). *Multivariate Data Analysis*. Prentice-Hall, Inc, Upper Saddle River, NJ, USA, 7 edition.

Hair, J. F., Ringle, C. M., and Sarstedt, M. (2011). Pls-sem: Indeed a silver bullet. *Journal of Marketing theory and Practice*, 19(2):139–152.

Hair Jr, J. F., Hult, G. T. M., Ringle, C., and Sarstedt, M. (2016). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage Publications.

Hasan, B. (2006). Delineating the effects of general and system-specific computer self-efficacy beliefs on is acceptance. *Information & Management*, 43(5):565–571.

Heckerman, D., Geiger, D., and Chickering, D. M. (1995). Learning bayesian networks: The combination of knowledge and statistical data. *Machine learning*, 20(3):197–243.

Henseler, J., Hubona, G., and Ray, P. A. (2016). Using pls path modeling in new technology research: updated guidelines. *Industrial management & data systems*, 116(1):2–20.

Herath, T. and Rao, H. R. (2009). Protection motivation and deterrence: a framework for security policy compliance in organisations. *European Journal of Information Systems*, 18(2):106–125.

Hof, H.-J. (2015). User-centric IT security-how to design usable security mechanisms. *arXiv preprint arXiv:1506.07167*.

Holden, H. and Rada, R. (2011). Understanding the influence of perceived usability and technology self-efficacy on teachers' technology acceptance. *Journal of Research on Technology in Education*, 43(4):343–367.

Hong, W., Thong, J. Y., and Wai-Man Wong, K.-Y. T. (2002). Determinants of user acceptance of digital libraries: an empirical examination of individual differences and system characteristics. *Journal of Management Information Systems*, 18(3):97–124.

Howe, A. E., Ray, I., Roberts, M., Urbanska, M., and Byrne, Z. (2012). The psychology of security for the home computer user. In *Security and Privacy (SP), 2012 IEEE Symposium on*, pages 209–223. IEEE.

Huth, D. (2017). A pattern catalog for gdpr compliant data protection.

Igbaria, M., Zinatelli, N., Cragg, P., and Cavaye, A. L. (1997). Personal computing acceptance factors in small firms: a structural equation model. *MIS quarterly*, pages 279–305.

Izquierdo-Yusta, A., Olarte-Pascual, C., and Reinares-Lara, E. (2015). Attitudes toward mobile advertising among users versus non-users of the mobile internet. *Telematics and Informatics*, 32(2):355–366.

Jacoby, J. and Kaplan, L. B. (1972). The components of perceived risk. *ACR Special Volumes*.

Jason, B., Calitz, A., and Greyling, J. (2010). The evaluation of an adaptive user interface model. In *Proceedings of the 2010 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists*, pages 132–143. ACM.

Jeyaraj, A., Rottman, J. W., and Lacity, M. C. (2006). A review of the predictors, linkages, and biases in it innovation adoption research. *Journal of Information Technology*, 21(1):1–23.

Juárez-Ramírez, R., Navarro-Almanza, R., Gomez-Tagle, Y., Licea, G., Huertas, C., and Quinto, G. (2013). Orchestrating an adaptive intelligent tutoring system: towards integrating the user profile for learning improvement.

Procedia-Social and Behavioral Sciences, 106:1986–1999.

Judson, R., Elloumi, F., Setzer, R. W., Li, Z., and Shah, I. (2008). A comparison of machine learning algorithms for chemical toxicity classification using a simulated multi-scale data model. *BMC bioinformatics*, 9(1):241.

Kainda, R., Flechais, I., and Roscoe, A. (2010). Security and usability: Analysis and evaluation. In *Availability, Reliability, and Security, 2010. ARES'10 International Conference on*, pages 275–282. IEEE.

Kaplan, L. B., Szybillo, G. J., and Jacoby, J. (1974). Components of perceived risk in product purchase: A cross-validation. *Journal of applied Psychology*, 59(3):287.

Kim, J. W., Lee, B. H., Shaw, M. J., Chang, H.-L., and Nelson, M. (2001). Application of decision-tree induction techniques to personalized advertisements on internet storefronts. *International Journal of Electronic Commerce*, 5(3):45–62.

Koller, D., Friedman, N., Getoor, L., and Taskar, B. (2007). Graphical models in a nutshell. *URL http://www. seas. upenn. edu/taskar/pubs/gms-srl07. pdf*.

Kumaraguru, P., Cranshaw, J., Acquisti, R., Cranor, L., Hong, J., Blair, M. A., and Pham, T. (2009). A real-word evaluation of anti-phishing training.

LaRose, R., Rifon, N., Liu, S., and Lee, D. (2005). Understanding online safety behavior: A multivariate model. In *The 55th annual conference of the international communication association, New York city*.

LaRose, R., Rifon, N. J., and Enbody, R. (2008). Promoting personal responsibility for internet safety. *Communications of the ACM*, 51(3):71–76.

Lee, M.-C. (2009). Factors influencing the adoption of internet banking: An integration of tam and tpb with perceived risk and perceived benefit. *Electronic commerce research and applications*, 8(3):130–141.

Lee, Y. and Kozar, K. A. (2008). An empirical investigation of anti-spyware software adoption: A multitheoretical perspective. *Information & Management*, 45(2):109–119.

Lin, J. C.-C. and Lu, H. (2000). Towards an understanding of the behavioural intention to use a web site. *International journal of information management*, 20(3):197–208.

Lin, W.-S. (2012). Perceived fit and satisfaction on web learning performance: Is continuance intention and task-technology fit perspectives. *International Journal of Human-Computer Studies*, 70(7):498–507.

Liu, B., Andersen, M. S., Schaub, F., Almuhimedi, H., Zhang, S. A., Sadeh, N., Agarwal, Y., and Acquisti, A. (2016). Follow my recommendations: A personalized privacy assistant for mobile app permissions. In *Twelfth Symposium on Usable Privacy and Security ({SOUPS} 2016)*, pages 27–41.

Lu, H.-P., Hsu, C.-L., and Hsu, H.-Y. (2005). An empirical study of the effect of perceived risk upon intention to use online applications. *Information Management & Computer Security*, 13(2):106–120.

Lu, J., Lu, C., Yu, C.-S., and Yao, J. E. (2014). Exploring factors associated with wireless internet via mobile technology acceptance in mainland china. *Communications of the IIMA*, 3(1):9.

Madsen, A. L., Jensen, F., Kjaerulff, U. B., and Lang, M. (2005). The hugin tool for probabilistic graphical models. *International Journal on Artificial Intelligence Tools*, 14(03):507–543.

Maguire, M. (2001). Context of use within usability activities.

Mezhoudi, N., Perez Medina, J. L., Khaddam, I., et al. (2015). Context-awareness meta-model for user interface runtime adaptation. *International Journal of Software Engineering*, 2.

Milne, G. R., Labrecque, L. I., and Cromer, C. (2009). Toward an understanding of the online consumer's risky behavior and protection practices. *Journal of Consumer Affairs*, 43(3):449–473.

Mitnick, K. D. and Simon, W. L. (2011). *The art of deception: Controlling the human element of security*. John Wiley & Sons.

Morris, M. G. and Venkatesh, V. (2000). Age differences in technology adoption decisions: Implications for a changing work force. *Personnel psychology*, 53(2):375–403.

Mun, Y. Y. and Hwang, Y. (2003). Predicting the use of web-based information systems: self-efficacy, enjoyment, learning goal orientation, and the technology acceptance model. *International journal of human-computer studies*, 59(4):431–449.

Nadkarni, S. and Shenoy, P. P. (2004). A causal mapping approach to constructing bayesian networks. *Decision support systems*, 38(2):259–281.

Ng, B.-Y. and Rahim, M. (2005). A socio-behavioral study of home computer users' intention to practice security. *PACIS 2005 Proceedings*, page 20.

Nielsen, J. (1994). *Usability engineering*. Elsevier.

Notario, N., Crespo, A., Martín, Y.-S., Del Alamo, J. M., Le Métayer, D., Antignac, T., Kung, A., Kroener, I., and Wright, D. (2015). Pripare: integrating privacy best practices into a privacy engineering methodology. In *Security and Privacy Workshops (SPW), 2015 IEEE*, pages 151–158. IEEE.

Omidosu, J. and Ophoff, J. (2016). A theory-based review of information security behavior in the organization and home context. In *Advances in Computing and Communication Engineering (ICACCE), 2016 International Conference on*, pages 225–231. IEEE.

Özkan, S., Bindusara, G., and Hackney, R. (2010). Facilitating the adoption of e-payment systems: theoretical constructs and empirical analysis. *Journal of enterprise information management*, 23(3):305–325.

Parsons, K., McCormac, A., Butavicius, M., and Ferguson, L. (2010). Human factors and information security: Individual, culture and security environment. Report, DTIC Document.

Pearson, S. (2013). *Privacy, security and trust in cloud computing*, book section 1, pages 9–13. Springer.

Pituch, K. A. and Lee, Y.-k. (2006). The influence of system characteristics on e-learning use. *Computers & Education*, 47(2):222–244.

Raghu, T., Kannan, P., Rao, H. R., and Whinston, A. B. (2001). Dynamic profiling of consumers for customized offerings over the internet: A model and analysis. *Decision Support Systems*, 32(2):117–134.

Rainie, L., Kiesler, S., Kang, R., Madden, M., Duggan, M., Brown, S., and Dabbish, L. (2013). Anonymity, privacy, and security online. *Pew Research Center*.

Ramayah, T. (2006). Doing e-research with e-library: Determinants of perceived ease of use of e-library. *International Journal of Technology, Knowledge and Society*, 1(4):71–82.

Ringle, C. M., Wende, S., and Becker, J.-M. (2015). Smartpls 3. *Boenningstedt: SmartPLS GmbH, http://www. smartpls. com*.

Ross, R. S. and Johnson, L. A. (2010). *Guide for Assessing the Security Controls in Federal Information Systems and Organizations: Building Effective Security Assessment Plans*. National Institute of Standards and Technology.

Sakellaropoulos, G. and Nikiforidis, G. (2000). Prognostic performance of two expert systems based on bayesian belief networks. *Decision Support Systems*, 27(4):431–442.

Sarstedt, M., Henseler, J., and Ringle, C. M. (2011). *Multigroup analysis in partial least squares (PLS) path modeling: Alternative methods and empirical results*, pages 195–218. Emerald Group Publishing Limited.

Schneier, B. (2011). *Secrets and lies: digital security in a networked world*. John Wiley & Sons.

Schwartz, A. M. (2011). Cybersecurity, innovation, and the internet economy. Report.

Shaughnessy, P. and Livingston, G. (2005). Evaluating the causal explanatory value of bayesian network structure learning algorithms. *Research paper*, 13.

Sheng, S., Holbrook, M., Kumaraguru, P., Cranor, L. F., and 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*, pages 373–382. ACM.

Shin, D.-H. (2009). Towards an understanding of the consumer acceptance of mobile wallet. *Computers in Human Behavior*, 25(6):1343–1354.

Suh, B. and Han, I. (2003). The impact of customer trust and perception of security control on the acceptance of electronic commerce. *International Journal of Electronic Commerce*, 7(3):135–161.

Sun, H. and Zhang, P. (2006). The role of moderating factors in user technology acceptance. *International journal of human-computer studies*, 64(2):53–78.

Tenenhaus, M., Vinzi, V. E., Chatelin, Y.-M., and Lauro, C. (2005). Pls path modeling. *Computational statistics & data analysis*, 48(1):159–205.

Thong, J. Y., Hong, W., and Tam, K.-Y. (2002). Understanding user acceptance of digital libraries: what are the roles of interface characteristics, organizational context, and individual differences? *International journal of human-computer studies*, 57(3):215–242.

Thong, J. Y., Hong, W., and Tam, K. Y. (2004). What leads to user acceptance of digital libraries? *Communications of the ACM*, 47(11):78–83.

Topa, I. and Karyda, M. (2015). Identifying factors that influence employees' security behavior for enhancing isp compliance. In *International Conference on Trust and Privacy in Digital Business*, pages 169–179. Springer.

Tsai, H.-y. S., Jiang, M., Alhabash, S., LaRose, R., Rifon, N. J., and Cotten, S. R. (2016). Understanding online safety behaviors: A protection motivation theory perspective. *Computers & Security*, 59:138–150.

Tsanas, A. and Xifara, A. (2012). Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. *Energy and Buildings*, 49:560–567.

Urbach, N. and Ahlemann, F. (2010). Structural equation modeling in information systems research using partial least squares. *JITTA: Journal of Information Technology Theory and Application*, 11(2):5.

Venkatesh, V. and Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management science*, 46(2):186–204.

Venkatesh, V. and 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*, pages 115–139.

Venkatesh, V., Morris, M. G., Davis, G. B., and Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, pages 425–478.

Whitten, A. and Tygar, J. D. (1999). Why Johnny can't encrypt: A usability evaluation of PGP 5.0. In *Proceedings of the 8th Conference on USENIX Security Symposium - Volume 8*, SSYM'99, pages 14–14, Berkeley, CA, USA. USENIX Association.

Wong, T. (2008). On the usability of firewall configuration. In *Symposium on usable privacy and security*.

Woon, I., Tan, G.-W., and Low, R. (2005). A protection motivation theory approach to home wireless security. *ICIS 2005 proceedings*, page 31.

Xu, D. J. (2006). The influence of personalization in affecting consumer attitudes toward mobile advertising in china. *Journal of Computer Information Systems*, 47(2):9–19.

Xu, D. J., Liao, S. S., and Li, Q. (2008). Combining empirical experimentation and modeling techniques: A design research approach for personalized mobile advertising applications. *Decision support systems*, 44(3):710–724.

Yiu, C. S., Grant, K., and Edgar, D. (2007). Factors affecting the adoption of internet banking in hong kong—implications for the banking sector. *International journal of information management*, 27(5):336–351.

Zurko, M. E. and Simon, R. T. (1996). User-centered security. In *Proceedings of the 1996 Workshop on New Security Paradigms*, NSPW '96, pages 27–33, New York, NY, USA. ACM.

A Survey Instrument, Descriptions and References for Measured Items

Part 1 – Demographic Profile/ External Variables
Essential for defining personal aspects of users in specific contexts (Lu et al., 2005; Juárez-Ramírez et al., 2013).

Individual Differences – Demographics	**Options**
Gender What is your gender?	A. Male B. Female C. Prefer not to say
Age In which category is your age?	A. 18-24 years B. 25-34 years C. 35-44 years D. 45-64 years E. 65-74 years F. 75 years or older
Education What is the highest degree or level of education you have completed*?* If currently enrolled, mark the previous grade or highest degree received.	A. 12th grade or less (no diploma) B. High school diploma C. Some college, no degree D. Associate or technical degree E. Bachelor's degree F. Graduate degree/professional
Employment Status	A. Employed for wages B. Self-employed C. Out of work and looking for work D. Out of work but not currently looking for work E. A homemaker F. A student G. Retired H. Unable to work
Income What category best describes your annual household income?	A. Less than $10,999 B. $11,000 to $49,999 C. $50,000 to 99,999 D. $100,000 or more
Ethnicity How would you classify yourself?	A. Arab B. Asian/Pacific Islander C. African/Black D. Caucasian/White E. Hispanic F. Latino G. Multiracial H. Other:.....
Physical Environment/Location Please indicate how often you use a notebook computer in the following locations.	A. Home: B. Apartment Lounge: C. Friend's house: D. Coffee Shop: E. Students Residence Halls: F. Classrooms/ Lecture Halls G. Other:.....
Experience and/or Frequency of use The set of questions here will be used to determineusers level of experience with web browser security settings as well as actualusage (Chang, 2004; Ng and Rahim, 2005). How many times do you use web browsers during a week?	A. not at all B. once/week C. several times/week D. less than once/day E. once/day F. 2-3/day G. bseveral times/day
Which of the following web browsers are you most familiar with?	A. Internet Explorer B. Google Chrome C. Firefox D. Other:.....

Which of the following web browser design do you prefer and/or find enjoyable to use?	A. Internet Explorer B. Google Chrome C. Firefox D. Other:.....
How often do you change security settings on your web browser?	A. not at all B. once/week C. several times/week D. less than once/day E. once/day F. 2-3/day G. several times/day
Domain Knowledge(DK) Adapted from Milne et al. (2009). DK_1: I have hadsignificant experience with configuring my browser security settings in thepast. DK_2: I am knowledgeable about cybersecurity and privacy related technologies. DK_3: I am skilled at avoiding dangers while browsing the internet	5-point Likert scale type strongly agree — strongly disagree

Individual Differences – Descriptive Charateristics
SE and SBCL are PMT constructs used to examine the mediating effects of participant's protection motivation on cybersecurity behaviours. The set of questions here are used to examine users level of experience with their preferred web browser as well as exposure to web browser security issues and protection motivation levels (Chang, 2004; Ng and Rahim, 2005). SE items are adapted from the instrument developed and empirically validated by (Compeau et al., 1999) while SBCL items are adapted from (Herath and Rao, 2009).

Self-Efficacy (SE)
I could optimise my web browser security settings ...
SE_1: ... if I had only the web browser manuals for reference.
SE_2: ... if I had seen someone else doing it before trying it myself (Reverse Coded)
SE_3: ... if there was no one around to tell me what to do as I go

Security Breach Concern Level (SBCL)
SBCL_1: Cybersecurity issues affects me directly
SBCL_2: Cybersecurity threats are exaggerated (Reverse Coded)
SBCL_3: I think cybersecurity issues should be taken seriously
SBCL_4: Security breaches are only targeted at organizations (Reverse Coded)

System Characteristics (SC) — SC assesses participants view on the user friendliness of their preferred web browser and are measured using items from (Thong et al., 2002, 2004). The construct is used to elicit individual preferences in terms of the Design, Terminology/ Language and Navigation of the browser security interface/ user interactions with the following items:
IC_1: I understand the terms used on my preferred browser security interfaces
IC_2: Layout of the browser security interface is clear and consistent
IC_3: The sequence of screens for security settings are difficult to navigate (Reverse Coded)
IC_4: Security functions are well depicted by buttons and symbols

Part 2 (A) – User Perceptions (TAM & PMT)
Perceived Ease of Use (PEOU) – is "the degree to which an individual believes that using a particular system would be free of physical and mental effort (Davis, 1989)." Likert type statements were adapted from previously validated measurement inventory of TAM variables and rephrased for web browser security settings (Davis et al., 1989; Lu et al., 2014;

Thong et al., 2002; Venkatesh and Davis, 2000).
PEOU_1: Learning to configure a browser security settings is easy for me
PEOU_2: Interacting with the interface for web browser security settings does not require a lot of my mental effort
PEOU_3: My interaction with web browser security settings is clear and understandable
PEOU_4: I find it easy to optimise my web browser security to the level of protection I want for my computer and privacy

Perceived Usefulness (PU) – which is also adapted from TAM's scale items is the degree to which a person believes web browser security settings would improve their protection against cyber-attacks (Davis, 1989).
PU_1: Web browser security functionalities gives me greater control over my safety and privacy online
PU_2: Overall, I find browser security settings useful in protecting my computer from cyber attacks
PU_3: Optimising my browser security settings gives me peace of mind when I am working with the internet
PU_4: The sensitive nature of information I search for and/or store on my personal computer requires me to optimise my web browser security settings

Perceived Risk (PR) – Questionnaire items for perceived risk was adapted from (Lu et al., 2005). Their research findings indicate that perceived risk indirectly impacts intentions to use an online application under security threats.
PR_1: Security functionalities embedded in web browsers are not adequate for preventing cyber attacks
PR_2: It is important to optimise browser security when visiting sites that requires data input
PR_3: I can make mistake whiles configuring my browser settings which can cause damage to my computer

Value for Personalization (VFP) – in this study VFP refers to the level of appreciation that a user has for all types of personalization possibilities within cyberspace. Items were adapted from the value of online personalisation scale developed and validated by Chellappa and Sin (2005).
VFP_1: I value online applications that are personalized based on information that is collected automatically (such as IP address, pages viewed, access time) but cannot identify me as an individual.
VFP_2: I value products and services that are personalized on information that I have voluntarily given out (such as age range, salary range, Zip Code) but cannot identify me as an individual.
VFP_3: I value application interfaces that are personalized for the device (e.g. desktop, mobile phone, tablet, etc.), browser (e.g. Internet explorer, Chrome, Firefox, etc.) and operating system (e.g. Windows, Unix) that I use.

Part 2 (B) — Attitude to Personal Data (APD)
To minimize survey fatigue, the APD scale adopted from (Addae et al., 2017) is simplified based overall cluster membership predictor importance of the APD factors as well as reliability score of the measured items.
Protection
PDP_1: I regularly look out for new policies on personal data protection
PDP_2: I consider the privacy policy of institutions where I give out such personal details
PDP_3: I don't always optimize my privacy settings when I create an online profile (Reverse Coded)
Awareness
PDA_1: Such details about me are of value to external organizations
PDA_2: Researchers don't need my consent to access my personal details (Reverse Coded)
PDA_3: Data collection organizations need to disclose the way the data are collected pro-

cessed and used.
Privacy Concern
PRI_1: I am sensitive about giving out information regarding my preferences
PRI_2: I am concerned about anonymous information (information collected automatically but cannot be used to identify me, such as my computer, network information, operating system, etc.) that is collected about me.

Part 3 — Cybersecurity Behavioural Intentions
Personalized Cybersecurity Adoption Intention (BI) — Items used to examine participants' general attitude to personalized adaptive web browser security are adapted from (Lu et al., 2014; Ng and Rahim, 2005).
BI_1: I am likely to accept personalized browser security update notification
BI_2: It is possible that I will allow adjustments to my web browser security settings to improve my safety online
BI_3: I am certain that I will pay attention to cybersecurity alerts tailored to my personal preference
Actual Cybersecurity Behaviour (ACB) – Items determining user interaction with web browser security settings were selected and adapted from the list of strategies people adopt to protect themselves online identified by (Rainie et al., 2013).
ACB_1: I have used service that allows me to browse the web anonymously
ACB_2: I don't set my browser to disable or turn off cookies (Reverse Coded)
ACB_3: I regularly clear cookies and browser history while I use the internet
ACB_4: I sometimes encrypt my communications while using the internet

Part 4 - Components of personalization
Items were adapted from (Xu et al., 2008) to acquire participants' ratings of the personalization dimensions identified for the purposes of building a BN-based model for adaptive cybersecurity.
User preference
1. Please indicate the importance of the following user interface characteristics to be considered in personalizing your web browser security and privacy settings:

a Language
b Presentation style (popup, icon change etc.)
c Navigation style (buttons, drop down etc.)
d Level of Information (Detailed vs. simplified)
e Others (please specify)

Adaptive Cybersecurity
2. Please indicate the importance of the following characteristics of an adaptive cybersecurity to be considered in personalizing your web browser security and privacy settings.

a User Effort Required
b Benefit of the security configuration
c Cost of the automated configuration
d Others (please specify)

Context
3. Please indicate the importance of the following contextual factors , which should be taken into consideration in personalizing your web browser security and privacy settings.

a Browser Type
b Enabled Browser Extensions
c Location
d Time
e Others (please specify)

User Goals/Needs
3. Please indicate the importance of the following user actions, which should be taken into consideration in personalizing your web browser security and privacy settings.

a Active Browsing session

b Browser History
c Explicit security/privacy queries
d Previous acceptance of personalized cybersecurity
e Others (please specify)