Know Your Enemy: User Segmentation Based on Human Aspects of Information Security

DAMJAN FUJS1, SIMON VRHOVEC2, AND DAMJAN VAVPOTIČ1

1University of Ljubljana, Faculty of Computer and Information Science, Ljubljana, Slovenia
2University of Maribor, Faculty of Criminal Justice and Security, Ljubljana, Slovenia

Corresponding author: Damjan Fujs (e-mail: damjan.fujs@fri.uni-lj.si).

ABSTRACT Users of information systems are the weakest link in information security. Considering their current information security performance is essential for improving information security training. User segmentation can help to improve information security training by dividing users into smaller groups based on their information security performance. In this paper, we present a segmented approach for information security training of users. To test the approach, we used data collected from students at a Slovenian university (N=165) with the Human Aspects of Information Security Questionnaire (HAIS-Q). HAIS-Q data was used to divide users into groups according to their information security performance via clustering. The proposed approach inherently balances adaptation of training to the needs of users and the efforts needed to achieve it which maximizes the key benefits of existing information security training approaches. With improved personalization, it mitigates the challenges related to training boringness and lack of user motivation which are emblematic for traditional information security training approaches. The proposed approach also offers some flexibility regarding the degree of personalization and the efforts related to information security training by fine-tuning the number of user groups. Finally, the proposed approach can help to identify beneficial software security requirements during development of new information systems.

INDEX TERMS Clustering methods, computer security, data processing, data visualization, information security, information systems.

I. INTRODUCTION

PEOPLE are known to be the weakest link in information security [1], [2]. Lack of sufficient information security can lead to loss of finances and reputation [3]. If we want to know the state of information security in an organization, we need to apply appropriate measurement methods [4] such as qualitative and quantitative metrics. Given that qualitative metrics often lead to uncertain conclusions, quantitative metrics may be more appropriate [4]. Moreover, information security has conventionally been focused on technical solutions. Yet, the importance of human factors has become increasingly recognized because technical solutions alone cannot sufficiently mitigate security vulnerabilities [1]. In addition, we are now connected more than ever and systems are becoming more complex [5] (the "smart" paradigm) – and if we have smart devices, we need to have smart people as well. Therefore, in order for users to use information systems securely, they must be properly trained.

Two key approaches for provision of training can be found in the literature, i.e., traditional and personalized training [2], [6]. The traditional (i.e., one-size-fits-all) approach follows the premise that all users should get the same training [2], [6], [7] although its content can be tailored to the needs of a specific organization. The traditional approach may not be suitable for all situations and often fails to deliver the intended benefits [2], [6], [8], [9]. For example, training should meet the needs of users (e.g., users’ knowledge, attitude and behavior related to information security [10]), and given the constantly changing cyber-environment, it should be up-to-date otherwise it becomes meaningless and/or expensive, and users may see training as an obstacle rather than as an advantage [2], [6], [11]. Tailoring training to the needs of individual users is known as the personalized approach [2], [6]. Although the
personalized approach addresses some of the issues of the traditional approach, it introduces new challenges, such as increased efforts and costs for tailoring security training to each individual, and is consequentially rarely seen in practice.

The drawbacks of both approaches can be addressed by seeking the middle grounds between generic training plans advocated by the traditional approach and individual training plans advocated by the personalized approach. In this paper, we aim to build on the strengths of existing information security training approaches, and address their challenges by proposing a novel training approach based on user segmentation. As guidance in the pursuit of our goals, we pose the following research questions:

\textbf{RQ1}: How to perform user segmentation based on their information security performance?

\textbf{RQ2}: How can user segmentation help in adapting information security training?

To answer the research questions, we perform user segmentation based on the established \textit{Human Aspects of Information Security Questionnaire} (HAIS-Q). HAIS-Q is a tool for measuring information security knowledge, attitude and behavior (i.e., information security performance) on seven different areas of information security (i.e., focus areas) \cite{10, 12}. User segmentation is done via clustering according to users’ scores in all seven HAIS-Q focus areas. Finally, we investigate how can segmentation be used to identify vulnerable users and adapt information security training accordingly.

This paper makes two key contributions to the literature. First, it proposes a new approach for provisioning information security training contributing to the information security literature. On one hand, the proposed approach provides more personalized training to users than the traditional approach as it provides groups of users with training according to their common needs. On the other hand, it requires less costs and efforts to provide personalized training, and is thus more scalable, than the personalized approach at the cost of fitting less the individual needs of users. Second, it advances the information security literature by leveraging user segmentation for tailoring the content of information security training. Clustering has been applied to HAIS-Q data before to identify groups of users according to their information security performance \cite{13}. This paper builds on these ideas, and further advances existing literature by enabling adaptation of information security training based on the segmentation of information system users according to their information security performance.

The remainder of the paper is structured as follows. In the next section, we discuss related work on human aspects of information security and user segmentation based on security-related characteristics. The proposed approach and research methodology are presented in Section III. In Section IV, the results of our study are presented and discussed. The paper conclusions are provided in Section V together with prospective directions for future research endeavors.

\section{RELATED WORK}

This section discusses existing work related to human factors in ensuring information security of organizations. It focuses on two key topics. First, it reviews the literature on human aspects of information security by emphasizing the measurement of information security performance of users of information systems. Second, it reviews the literature on user segmentation based on their security-related characteristics by focusing on works related to users and/or security that employ clustering.

\section{HUMAN ASPECTS OF INFORMATION SECURITY}

Information security has traditionally been provided by technical solutions \cite{1} in the final stages of software development. This means that security is seen as an add-on feature therefore, when a vulnerability arises, it is addressed with security patches \cite{14}. Although technical solutions in the field of information security are important, they are not sufficient. Human aspects often play an even more important role in achieving comprehensive information security \cite{15}. There are differences among users of information systems in their information security knowledge, attitude and behavior \cite{16}. HAIS-Q has been developed to measure information security performance \cite{12}. It considers information security knowledge, attitude and behavior of users in seven focus areas (i.e., password management, email use, internet use, social media use, mobile devices use, incident reporting, information handling). HAIS-Q has been well-received in scientific and professional circles, and has been applied in numerous studies. For example, users who scored higher according to the HAIS-Q performed better in a phishing experiment suggesting that HAIS-Q may be a good predictor of information security behavior \cite{10}. Similarly, studies show that several HAIS-Q factors are associated with better cyber hygiene \cite{16}. HAIS-Q was also used in a study which found that there are gender differences in information security awareness \cite{1}. In addition, the study suggests that organizational and security cultures are also important besides information security awareness. Moreover, if the organizational culture of an organization improves, so will its security culture \cite{1}. HAIS-Q is therefore a state-of-the-art questionnaire for measuring information security performance of individuals. The latest applications of HAIS-Q cover various areas of information security, such as information security practices of hospital staff \cite{17} and development of methodologies for assessing information security awareness \cite{13, 18–20}.

Most studies use the entire questionnaire without any changes. Nevertheless, there are some studies that modify HAIS-Q, e.g., by shortening it (short-HAIS-Q) \cite{21} or by adding and/or removing a particular focus area \cite{22}, etc.
Some studies combine HAIS-Q with other questionnaires by adding them without changing the composition of HAIS-Q [1], [23]–[25]. In our literature review of studies using HAIS-Q, we found a single study applying clustering on HAIS-Q data [13]. Albeit this study identifies various groups of users according to their security performance and describes them, it stops short of suggesting that such results could be leveraged for tailoring information security training. This paper builds on the ideas found in this study to improve adaptation of information security training thus advancing the state-of-the-art in this research area.

### B. USER SEGMENTATION BASED ON SECURITY-RELATED CHARACTERISTICS

Being able to categorize users is one of the key goals when adapting security requirements (either software features or information security training requirements) [26]. Information security training is crucial because only appropriate training can improve the overall information security [16]. Similarly, software security features can improve security of information systems, especially when they cover a broad range of vulnerable users. The concept of segmentation is mainly useful because it can help to identify potentially interesting (e.g., vulnerable) users [27]. In computer science, the term segmentation is used mostly in the field of computer vision albeit it was first used in marketing to define different user groups [27].

Several kinds and ways of applying user segmentation can be found in the literature. For example, Xiao [28] provides security-related segmentation of smartphone users based on their gender and type of mobile operating system. In addition, user segmentation by gender can be found in several other studies, such as [1], [10], [16]. Cain et al. [29] explore the differences between age groups and gender regarding cyber hygiene knowledge and behavior. In their study, they also found that current one-size-fits-all training approaches have no impact on knowledge and behavior. This is consistent with other literature (e.g., [2], [6]) indicating the need for improving the traditional approach to information security training. Anichiti et al. [30] provide segmentation based on users’ knowledge about cyber security in hotels. They categorize users according to their generation (i.e., generation X, Y and Z). Users have been also categorized according to the role they play in an organization [31].

State-of-the-art literature on clustering in the context of security and users of information systems can be roughly divided into three areas: works that are directly related to users, works that are directly related to their security, and works that are directly related to both. First, we focused on works related to users but are not directly security-related since such segmentation approaches can be applied to the domain of information security as well. Most identified studies employ an experimental research design. For example, clustering has been applied to various research areas, such as users of smartphone applications [32], [33], behavior of retail customers [34], downloading behavior of academic search engine users [35], online shopping customer loyalty [36], and mobile network users [37]. All these studies use automatically generated data (e.g., application usage, online shopping transactions) as input into user segmentation. Some identified studies however use data collected via a cross-sectional survey research design for user segmentation in the various research fields, such as e-commerce [38] and technology threat avoidance [39]. Second, we searched for works that are related to security but not directly user-related. For example, we identified a study that employs clustering for profiling phishing activities [40]. Similar to most user-related works, this study also uses data that is automatically generated by the users for segmentation. Third, we reviewed works that are directly related to both users and security. An experimental research design has been employed in a number of research fields, such as authentication in IoT environments [41], analysis of encrypted data [42], behavior of web users [43], and privacy on social networks [44]. These studies use automatically generated user data as input into user segmentation. Additionally, several studies employ user segmentation based on data collected with various questionnaires in cross-sectional research designs. For example, studies deal with security measures of users in the European Union [45], [46], and information security expertise of users [47].

Several segmentation approaches related to users and/or security can be found in the literature. To the best of our knowledge, there is however no approach that would allow for an *a priori* prediction of optimal groups of information system users according to their information security performance. Therefore, groups need to be identified based on real data of specific organizations and their information system users. Albeit user segmentation has been used in the context of information security, it has not been applied to the context of information security training.

### III. HAIS-Q BASED USER SEGMENTATION

In this section, we present the proposed segmented training approach employing user segmentation based on HAIS-Q data for the selected population, and the research methodology used to test it.

#### A. RESEARCH DESIGN

We employed a cross-sectional survey research design to collect HAIS-Q data from the selected population. HAIS-Q data is then used for testing the proposed training approach. The entire research process starting with context selection and ending with clustering is shown in Fig. 1.

HAIS-Q was selected since it is one of the most established instruments for measuring information security performance, has been empirically validated several times, and comprehensively covers different information security focus areas. For the needs of our research, HAIS-Q was translated from English into Slovenian. A standard
questionnaire translation procedure was conducted: 1) three independent translations, 2) consolidation of independent translations, and 3) back-translation to ensure consistency between the English original and the Slovenian translation of the questionnaire. The Slovenian questionnaire was qualitatively reviewed by three academic peers who made recommendations for improving unclear questionnaire items. The final version of the Slovenian questionnaire was developed by considering the feedback obtained during the pre-test.

B. DATA COLLECTION AND SAMPLE

The study was conducted among students at one of the faculties at the second largest public university in Slovenia. Students were chosen because they are the largest group of users of university information systems. They are very familiar with technology and use it on a daily basis [48], will soon start their professional career [22], and are perceived as a homogeneous social group with related background leaving limited space for diverse beliefs about information security [49]. Furthermore, studies report that students are a comparable population to "non-students" [50] and that they can be considered as a valid population for experiments in this field [51]. It is also important that lack of information security awareness prevails among students which makes them well suited for information security research [52].

The survey questionnaire was hosted on the local hosting platform 1ka (https://1ka.arnes.si/). Respondents were informed that participation in the study was voluntary and anonymous. The survey was conducted in February 2021. An invitation to complete the survey was sent to e-mail addresses of 964 students. A total of 165 respondents completed the survey providing for a response rate of 17.1 percent. 83.0 percent of respondents were full-time students, 70.9 percent female, and 55.2 percent living in an urban area. The average age of respondents was 24 years.

C. DATA ANALYSIS AND PROPOSED APPROACH

The data collected in the survey were qualitatively checked before use. No entries were removed since all answers were consistent and no values were missing. Reverse-coded questionnaire items (about a half of HAIS-Q items is reverse-coded [10]) were recoded prior to data analysis. When using 5-point Likert scale, this means that scores 1, 2, 3, 4 and 5 are recoded to 5, 4, 3, 2 and 1, respectively. Aggregated scores for the seven HAIS-Q focus areas covered by HAIS-Q were calculated based on each respondent’s scores for knowledge, attitude and behavior (3 items for each dimension). HAIS-Q data for each respondent (63 variables) were therefore aggregated into seven new variables by averaging all 9 items for each of the seven HAIS-Q focus areas. The data was analyzed with IBM SPSS Statistics version 28.

The main data analysis procedure closely followed the proposed approach as presented in Fig. 2. Since the purpose of this study was not to find the best clustering algorithm for user segmentation but rather to propose and test a novel information security training approach, we employed standard clustering methods to perform user segmentation according to their information security performance (i.e., information security knowledge, attitude and behavior).

First, we conducted a preliminary hierarchical cluster...
to determine the most appropriate number of clusters within the collected data by following the procedure described in [53]. A visual analysis of the dendrogram indicated three clusters.

Next, the TwoStep clustering algorithm was employed for user segmentation based on HAIS-Q. TwoStep clustering is based on two steps – an algorithm similar to $k$-means is first performed, followed by the formation of homogeneous clusters [54]. Analysis with TwoStep clustering has several advantages. First, it enables automatic determination of the discretization intervals by considering the unique data distribution property of each attribute [55]. Second, it handles both categorical and numerical variables. Third, it calculates the predictor importance of variables in a cluster solution. We compared solutions with two, three, four and five clusters by conducting non-parametric tests. For each solution, a Kruskal-Wallis $H$ test was first performed to determine whether there were differences in the distributions of aggregated HAIS-Q focus area scores among all clusters in a solution. Further insights were obtained by performing post-hoc Mann-Whitney $U$ tests among pairs of clusters in a solution. Non-parametric tests indicated that the solutions with two and three clusters had clear differences between all clusters. Solutions with four and five clusters however did not have clear differences between all clusters (e.g., there were no statistically significant differences between cluster 3 and 4 for any of the HAIS-Q focus areas). This analysis further confirmed that the solution with three clusters was the most suitable.

To validate the quality of the solution with three clusters, the Silhouette measure of cohesion and separation was considered. Silhouette measure of cohesion and separation can take a value from $-1$ to $1$. In the best case, the within-cluster distances are short, and the between-cluster distances are wide, resulting in a silhouette measure close to the maximum value of $1$ [56]. Values above $0.20$ are considered as adequate as they indicate that there is a fair separation distance between clusters [57]. In our case, the Silhouette measure of cohesion and separation was $0.44$ suggesting that the solution with three clusters is adequate. Finally, we considered variable importance for each HAIS-Q focus area to determine whether it should be excluded from further analysis. Variable importance is related to some of the outcomes of the TwoStep clustering procedure in SPSS, namely within-cluster importance (i.e., features are ranked according to their importance for each individual cluster) and overall importance (i.e., the calculated overall importance of a feature for all clusters). The rule of thumb is to exclude variables with overall importance below $0.40$ from further analyses [53]. Since the lowest score for overall importance was $0.68$ which is well above the threshold, we did not exclude any variables from the analysis. It may be worth noting that even if a HAIS-Q focus area would be excluded from user segmentation, training for the excluded focus area would still need to be provided in alternative ways.

VOLUME 4, 2016
IV. RESULTS AND DISCUSSION

In this section, we present key results for the three user groups (i.e., low, moderate and high risk users) obtained with TwoStep clustering of users’ HAIS-Q data. Fig. 3 shows spider charts visualizing information security performance covering all seven HAIS-Q focus areas for all three identified user groups. Table 1 presents cluster means, within-cluster importance and overall importance for variables representing the seven HAIS-Q focus areas.

Low risk users achieve the best results in all HAIS-Q focus areas. Moderate and high risk users are the most interesting in terms of information security training since their scores are noticeably lower than the scores of low risk users. Moderate risk users have somewhat lower values than low risk users for all HAIS-Q focus areas. High risk users however have considerably lower values than low risk users and moderate users for all HAIS-Q focus areas. The aim of the first step of information security training may be to improve their information security performance to be at least comparable to the group of low risk users. After improving the security performance of moderate and high risk users, information security training may focus on gradually improving information security performance of all user groups. These results indicate that high risk users need substantial training in all seven HAIS-Q focus areas with special attention to social media use, internet use and incident reporting as their average scores for these focus areas were lower than 2 on the scale from 1 to 5. As this group consists of only 4 users, it may be reasonable to limit their access to the information system until their information security performance improves if possible. Moderate risk users have considerably better information security performance than high risk users. Nevertheless, comparing them to low risk users shows that improvements are possible, and desirable, in all seven HAIS-Q focus areas. Even though moderate risk users should also attend training in all seven HAIS-Q focus areas like high risk users, it should be adapted to their level of information security performance.

To better demonstrate the differences between existing training approaches (i.e., traditional and personalized) and the proposed segmented approach, we visualize them in Fig. 4.

According to the traditional training approach, the degree of personalization is typically very low or non-existent as training is provisioned in the same way for everyone [2], [7]. Trainees who are already familiar with some or many of the topics given at time-consuming uniformed training are bored and unmotivated [2], [7], [58]. In large companies with many employees, this leads to many wasted work hours. As typical information security training tends to be high cost [58], [59], it is essential for users to attend only training on topics they need. In the best case, traditional training approaches consider users’ information security performance (e.g., HAIS-Q data) and provide training only in the areas that users are noticeably lacking. For example, in our case, training may be provided for HAIS-Q focus areas with an overall mean below 4 (i.e., password management, email use, internet use, social media use, and incident reporting). Therefore, all users would have to attend information security training for all selected focus areas. In the worst case, the selection of training topics would be done without any insights into information security performance of users.

By following the personalized training approach, a training plan would be developed for each individual user [2], [6]. For example, 165 different information security training plans would have to be developed in our case if we consider only the sample that provided responses in our survey, and 964 information security training plans otherwise. This approach may help to deal with training boredom and to improve the motivation of trainees. Since users would be trained only in topics they need, it would also result in less wasted work hours. Nevertheless, developing an information security training plan for each individual user may consume considerable resources hindering the scalability of such endeavors. The provision of training to each individual may be similarly challenging. Although both personalization and provision of personalized information security training may be automated, its potential success is yet unclear. Since adequate motivation of trainees may be needed for the success of such training and motivation is a general issue in information security [2], it seems quite unlikely that such information security training would succeed in practice, especially for individuals with poor information security performance.

The proposed segmented training approach has several advantages in comparison to existing approaches. First, it balances training personalization and the efforts needed to achieve it thus enabling more efficient training plan development. In our study, we identified three user groups. For each user group, a single tailored information security training plan needs to be developed. This is approaching key benefits of the personalized approach albeit for a fraction of efforts needed to develop individual information security training plans. Second, it deals with boredom and improves the motivation of trainees by ensuring that all users attend only training on topics they need. Third, the number of clusters can be selected according to a number of factors besides cluster quality, such as overall information security performance of users, available resources (e.g., maximum number of user groups), desired level of personalization, etc. This is important in the context of training planning as each additional tailored training causes extra costs on one hand but enables more focused training of users on the other hand. Finally, the proposed approach can serve as a starting point for the definition of security requirements for an information system. Usually, these requirements are defined in line with requirements engineering practices that often neglect security issues which are then considered only in later stages of software development [1]. Requirements engineering generically covers activities related to software
FIGURE 3. Information security performance comparison between user groups (i.e., clusters). The smaller the surface, the better information security performance of users in the group. HAIS-Q focus areas are arranged in descending order by their overall importance clockwise, starting with SMU. Notes: SMU – social media use, IU – internet use, IR – incident reporting, IH – information handling, MDU – mobile devices use, EU – email use, PM – password management.

TABLE 1. Within-cluster and overall variable importance with mean values for each user group (i.e., cluster). HAIS-Q focus areas are arranged in descending order by their overall importance. Note: $M$ – mean, $V_{oi}$ – overall importance, $V_{wcr}$ – within-cluster importance.

| HAIS-Q focus area       | $V_{oi}$ | $V_{wcr}$ | $M$   | $V_{wcr}$ | $M$   | $V_{wcr}$ |
|-------------------------|----------|-----------|-------|-----------|-------|-----------|
| Social media use (SMU)  | 1.00     | 4.03      | 2     | 3.24      | 1.72  | 2         |
| Internet use (IU)       | 0.97     | 3.98      | 1     | 3.12      | 1.44  | 1         |
| Incident reporting (IR) | 0.90     | 4.02      | 3     | 3.31      | 1.47  | 3         |
| Information handling (IH)| 0.79  | 4.40      | 4     | 3.61      | 2.00  | 5         |
| Mobile device use (MDU) | 0.71     | 4.43      | 6     | 3.66      | 2.47  | 4         |
| Email use (EU)          | 0.71     | 4.13      | 5     | 3.37      | 2.14  | 7         |
| Password management (PM)| 0.68     | 4.05      | 7     | 3.68      | 2.31  | 6         |

FIGURE 4. Comparison between traditional, personalized and the proposed segmented training approaches.

requirements (i.e., what the system should contain or what the software should do in general) [60]. Security requirements engineering specifically addresses security requirements based on risk analysis (i.e., what the system should contain or what the software should do to make it secure) [61]. It is especially important how a system is being designed strategically (i.e., keeping users of information systems in mind during all phases). The proposed approach can be leveraged to partially mitigate this situation as it enables the identification of both, big groups of highly vulnerable users and most problematic HAIS-Q focus areas. This information can be used to implement specific software security requirements that address one or more of the identified issues. The implementations of such security requirements would complement tailoring of information security training and further enhance information security.
V. CONCLUSIONS AND FUTURE WORK

The purpose of this paper is to propose a novel training approach based on user segmentation. The proposed approach combines HAIS-Q and clustering to identify key groups of information system users according to their information security performance. We tested the proposed training approach with real-world HAIS-Q data collected from university students.

We answered to RQ1 (How to perform user segmentation based on their information security performance?) by showing how users can be grouped according to their information security performance by applying the TwoStep clustering algorithm on HAIS-Q data. We opted for HAIS-Q which is a widely used instrument for measuring information security performance. Nevertheless, HAIS-Q has some inherent weaknesses as it is a self-administered questionnaire. For example, users are free to choose whether they want to take the survey or not. This may result in some users being left out of the segmentation. Even if the survey is mandatory, some users may not engage with the survey and/or provide non-meaningful responses. This could result in incomplete data that may leave out of segmentation a significant part of information system users or significant noise that could distort the results. Another issue is that surveys cannot be done continuously. If periods between consecutive surveys are too short, users may remember their previous responses and/or learn how to respond correctly to the items included in HAIS-Q without actually improving their information security performance. These shortcomings may be addressed by incorporating other instruments for measuring information security performance, preferably more objective and automatic. Besides the TwoStep clustering algorithm, other clustering algorithms may be used to group users according to their information security performance. Further real-world studies would be needed to determine which clustering algorithms perform well in organizational settings and on which type of information security performance data, and which do not.

To answer RQ2 (How can user segmentation help in adapting information security training?), we developed and tested a novel approach for information security training based on user segmentation. User segmentation noticeably helps in adapting information security training to the needs of information system users. Several advantages of the proposed approach can be identified compared to existing approaches. For example, it inherently balances adaptation of training to the needs of users and the efforts needed to achieve it. With improved personalization, it mitigates the challenges related to training boredom and lack of user motivation which are emblematic for traditional approaches.

The proposed approach also offers some flexibility regarding the degree of personalization and the efforts related to information security training. The benefits of the proposed approach may go beyond information security training as it can help to identify beneficial software security requirements during development of new information systems. The success of adapting information security training largely depends on the quality of user segmentation. Therefore, its weaknesses are inherently also weaknesses of training adaptation based on user segmentation. Real-world studies of training provision would be highly beneficial to determine any other challenges of the proposed approach.

This work has some limitations that the readers should note. First, the proposed approach focuses on seven HAIS-Q focus areas. Additional information security aspects may exist that were not included in this study. Future studies may extend the proposed approach to include them. Second, our study investigated a single user group. Studies including different and more heterogeneous user groups would offer more insights into challenge related to segmentation and training of such groups, including users that are less skillful with information technology. The proposed approach focuses on training although information security involves both people and technology. Future works may explore how to integrate it with approaches for ensuring information security from the technical perspective. For example, future studies may investigate how development of software security features may complement information security training, and answer questions, such as, is it more efficient to implement an additional software security feature or train users to improve their information security performance related to the software security feature.

ACKNOWLEDGMENT

We would like to thank all participants in our study that took their time for answering the questionnaire. We would also like to thank the reviewers for their insightful suggestions which helped to improve the paper.

We declare that the research was conducted in the absence of any commercial or financial relationships that could be interpreted as a potential conflict of interest.

REFERENCES

[1] A. Wiley, A. McCormac, and D. Calic, “More than the individual: Examining the relationship between culture and Information Security Awareness,” Computers & Security, vol. 88, 2020.

[2] S. Furnell and I. Vasileiou, “Security Education and Awareness: Just Let Them Burn?” Network Security, vol. 2017, no. 12, pp. 5–9, 2017.

[3] K. R. Sarkar, “Assessing insider threats to information security using technical, behavioural and organisational measures,” Information Security Technical Report, vol. 15, no. 3, pp. 112–133, 2010.

[4] K. Prislan, A. Mihelčič, and I. Bernik, “A real-world information security performance assessment using a multidimensional socio-technical approach,” PLoS ONE, vol. 15, pp. 1–28, 2020.

[5] Z. Lv, W. Mazurczyk, S. Wendzel, and H. Song, “Guest Editorial: Recent Advances in Cyber-Physical Security in Industrial Environments,” IEEE Transactions on Industrial Informatics, vol. 15, no. 12, pp. 6468–6471, 2019.

[6] I. Vasileiou and S. Furnell, “Personalising Security Education: Factors Influencing Individual Awareness and Compliance,” in Information Systems Security and Privacy. Cham: Springer International Publishing, 2019, pp. 189–200.

[7] Z. J. Zhang, W. He, W. Li, and M. Abdous, “Cybersecurity awareness training programs: a cost–benefit analysis framework,” Industrial Management and Data Systems, vol. 121, no. 3, pp. 613–636, 2021.

[8] D. Fuijs, S. Vrhovec, and D. Vavpotič, “Bibliometric mapping of research on user training for secure use of information systems,” Journal of Universal Computer Science, vol. 26, no. 7, pp. 764–782, 2020.
Statistics. Springer-Verlag Berlin Heidelberg, 2011, vol. 53, no. 9, ch. Cluster An, pp. 1689–1699.

[55] S. C. Tan, “Improving association rule mining using clustering-based discretization of numerical data,” in 2018 International Conference on Intelligent and Innovative Computing Applications, ICONIC 2018. Mon Tresor, Mauritius: IEEE, 2019, pp. 18–22.

[56] A. E. M. Celestino, D. A. M. Cruz, E. M. O. Sánchez, F. G. Reyes, and D. V. Soto, “Groundwater quality assessment: An improved approach to K-means clustering, principal component analysis and spatial analysis: A case study,” Water (Switzerland), vol. 10, no. 4, pp. 1–21, 2018.

[57] A. Tkaczyński, “Segmentation Using Two-Step Cluster Analysis,” in Segmentation in Social Marketing: Process, Methods and Application, 2016, pp. 109–125.

[58] N. Chowdhury and V. Gkioulos, “Cyber security training for critical infrastructure protection: A literature review,” Computer Science Review, vol. 40, p. 100361, 2021.

[59] B. H. Kim, K. C. Kim, S. E. Hong, and S. Y. Oh, “Development of cyber information security education and training system,” Multimedia Tools and Applications, vol. 76, no. 4, pp. 6051–6064, 2017.

[60] A. M. Nhlabatsi, K. M. Khan, J. B. Hong, D. S. D. Kim, R. Fernandez, and N. Fetais, “Quantifying Satisfaction of Security Requirements of Cloud Software Systems,” IEEE Transactions on Cloud Computing, vol. XX, pp. 1–18, 2021.

[61] R. Laborde, S. T. Bulusu, A. S. Wazan, A. Oglaza, and A. Benzekri, “A Methodological Approach to Evaluate Security Requirements Engineering Methodologies: Application to the IREHDO2 Project Context,” Journal of Cybersecurity and Privacy, vol. 1, no. 3, pp. 422–452, 2021.

DAMJAN FUJS is Assistant and Ph.D. student at the Faculty of Computer and Information Science, University of Ljubljana, Slovenia. His research interests include cyber security, security requirements engineering and software development methodologies. He has published several scientific papers in the field of information security during his doctoral studies. Currently, he is also a member of the program committee of two conferences: 1) 28th Dnevi Slovenske Informatike (DSI 2021) and 2) The Sixth International Conference on Cyber-Technologies and Cyber-Systems (CYBER 2021).

SIMON VRHOVEC is Associate Professor at the University of Maribor, Slovenia. He received his PhD degree in Computer and Information Science from the University of Ljubljana (Slovenia) in 2015. He is editorial board member of the Journal of Cyber Security and Mobility, Frontiers in Computer Science, EUREKA: Social and Humanities, and International Journal of Cyber Forensics and Advanced Threat Investigations (CFATI). He serves or has served as guest editor for IEEE Security & Privacy, Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), and Journal of Universal Computer Science (J.UCS). He is in the steering committee of the European Interdisciplinary Cybersecurity Conference (EICC) since 2019, and co-chaired the Central European Cybersecurity Conference (CECC) in 2018 and 2019. His main research interests are in human factors in cybersecurity, secure software development, agile methods, and change management. Contact him at simon.vrhovec@um.si.

DAMJAN VAVPOTIČ is Associate professor and head of Information Systems Laboratory at the Faculty of Computer and Information Science, University of Ljubljana. While his main area of expertise is in the field of business information systems, in recent years his research also focused on analysis of longitudinal geospatial data in tourism and healthcare. He has published in a range of scientific journals including Business & Information Systems Engineering, International Journal of Project Management, Information and Software Technology, International Journal of Environmental Research and Public Health, and Tourism Economics where he received 2019 Tea Sinclair Award for Article Excellence.