Vol. 4 Special Issue. 1
International Journal of Data Science and Advanced Analytics (ISSN: 2563-4429)

Filling The Gap of Attacker Behavior Analysis in Organizing and Understanding CTF Games Among Students

Thana Jebbr[1], Intan Farahana Binti Kamsin[2], Salmiah Binti Amin[3], Nur Khairunnisha Binti Zainal[4]
Asia Pacific University of Technology & Innovation, Technology Park Malaysia, Bukit Jalil, Kuala Lumpur, Malaysia.
[1]thanamajid27@gmail.com, [2]intan.farahana@apu.edu.my, [3]salmiah@apu.edu.my,
[4]khairunnisha.zainal@apu.edu.my

Abstract— CTF games are gaining popularity in educating young cybersecurity students to train them in cyber-attack and defence. This research aims to propose an effective IoT solution to solve issues that arise from the absence of attacker behaviour analysis (ABA) in organizing and analysing CTF games as a part of students cybersecurity education. This study follows a quantitative methodology and respondents are selected by purposive judgemental sampling method to acquire knowledgeable APU students. ABA is more than relevant to cyber defence research and can positively contribute to more effective and successful CTF events.

Index Terms - CTF, effect, students, attacker behaviour analysis (ABA)

1.0 Introduction

As cyber-attacks grow more violent, cyber defence grows more important to secure systems against attacks. Cybersecurity is integrating CTF games, Capture-The-Flag events where participants must solve in order to capture a flag, by applying rigorous offensive practices of cryptography, web applications analysis among others, to train cybersecurity students and interest them in understanding cyber-attacks more. These simulations are all faked and created for the participants only. Alternately, ABA, or Attacker Behaviour Analysis is the study targeting cyber-attackers behaviour and trends to better understand how they attack system vulnerabilities. ABA is utilized in research and security units to understand their vulnerabilities and secure their systems against them more effectively.

2.0 Literature Review

2.1 Domains

2.1.1 CTF Games

Growing more common in the world of promoting cybersecurity, CTF games are Capture the Flag events where participants create teams to solve challenges of varying complexities to capture a flag and submit it for rewarding points on a scoreboard [5]. These challenges can be solved by implementing various invasive, investigative, and analytical cybersecurity and computer analysis concepts [10].

These games are conducted online and can span from over a course of hours to days. Teams gather worldwide to compete and learn more invasive techniques and apply them to hack and become a part of the cybersecurity community. CTF games come under Jeopardy styled challenges, where a team solve multiple individual challenges, and Attack and Defense style, where teams must infiltrate and attack other team's individual systems while defending their own [4]. These events are constantly being integrated into education and used to encourage students to pursue cybersecurity [5].

Summary

While researchers agree that CTF games include a vital and expressive human component, there is a lack of utilization of this resource to help organizers create better challenges to feature in these CTF events [4] [2]. Additionally, there is a bigger role that human behaviour analysis is eager to play in organizing and understanding the effects of CTF games [4]. This is known as ABA or Attacker Behaviour analysis. When performed, ABA data helps investigators better understand how cybercriminals think, behave, and strategize to better defend their systems against them. In addition, researchers have shown concern of ethical issues that arise in training students offensive hacking techniques in CTF games [6]. If ABA experts can gain control of the application of CTF events, they can better understand their effects on participants who will eventually grow up to be hacking and system vulnerability experts in the future.

Reflection

With the help of ABA, organizers can design challenges that pass ethical standards [6]. In light of this, CTF events are a perfect ground to conduct these studies and draw sound conclusions. This data can be used in constructing better targeted and calculated challenges for CTF games, and in the world of cyberwarfare. However, since there is lacking research in this domain, this paper hopes to shed light on the matter and introduce a solution that fills the need this industry thirsts for: a platform and framework that allows experts in ABA and CTF organizers to communicate and discuss key information to spread the knowledge and work together [1]. This exchange of information and

expertise allows the pooling of resources to benefit both parties.

2.1.2 Attacker Behaviour Analysis (ABA)

ABA is the analysis of attackers' behaviour to study their movements and techniques. This can be done via a platform where hackers' conducts and incentives are constantly tracked in crafted controlled scenarios, like CTF games [11]. These scenarios are systematically updated, providing a detailed understanding of the attacker's patterns and motivations, improving the effectiveness of the study and the world of cyber defence. In real-time, ABA is used to produce accurate models that realistically capture the nature of cybersecurity-related threats and attacks [7]. One of the factors that make ABA such an appropriate concept for investigating cyber attacker threats is how inclusive it represents and explains actual actions, which psychology alone falls short of [7]. As cyber threats and attacks are constantly evolving, a precise understanding of the attacker's decision-making process becomes crucial to implement successful security strategies in the near future [11].

Summary

As cyber-attacks and system vulnerabilities grow around the globe, researchers have shown keen interest in the success of ABA to investigate the behavioural indicators of hackers when attacking a system. Though this concept is new, it is evolving to service multiple researchers and security specialized companies to generate realistic and accurate cybercriminal trends and shows great promise on how to better secure system vulnerabilities. With the number of attacks increasing, more data is processed through ABA which only makes it more realistic and applicable.

Reflection

With the industry's bleeding need for realistic models that accurately depict cyber attackers behaviour and researcher's lack for such challenging studies, it is only appropriate that CTF events are linked to this phenomenon.

When implementing this analysis, cybersecurity educators can answer the questions of how cybercriminals think, and the insights gathered can be used to improve security measures around the digital world [11]. While these services are employed in real-life applications, this concept can also be deployed in the world of CTF events. These events are the perfect grounds for deploying this study and analyse how cybercriminals behave in a non-risky environment where all scenarios are faked, and fiction and analysts can alter and control [6]. ABA experts play a major role in the solution addressed, where their expertise can help organizers design CTF challenges, and in turn can monitor how participants react to them. Such pooling of resources can benefit both communities and promote successful research into ABA in controlled environments and real life.

2.2 Similar Systems

2.2.1 DISCORD

While Discord is known to be a word-wide social media platform, many CTF organizers admit to using it to discuss creating challenges and discussing the results of the CTF events. This platform has become common to easily communicate with other organizers and participants. Its chat features enable users to share images, videos, texts and even conduct online meetings with absolute ease. However, it is not specifically designed to assist organizers in designing CTF challenges or analysing the participant's attacker behaviour. Therefore, it does not fulfil the requirement that many CTF organizers admit to needing. Additionally, it does not provide any insight to its users or any information but is merely a platform like WhatsApp or Teams.

Screenshot:

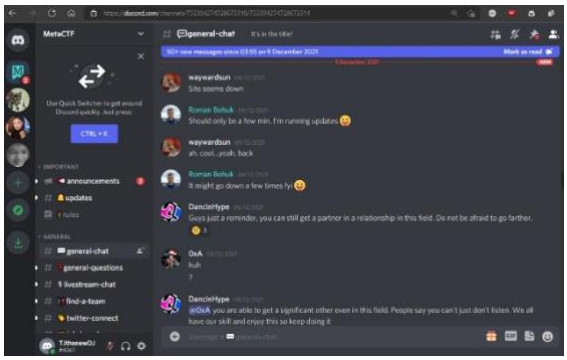

Figure 1.0 shows a screenshot of Discord's platform as used by CTF organizers.

2.2.2 Bit Trap

Bit Trap is a new-age program that can be used to conduct ABA or Attacker Behaviour Analysis. This is possible with its feature of setting bait traps for hackers like honeypot when the system is infiltrated, however, the system is rigged to record and track all the paths the hacker chose and all data they left behind. Next, Bit Trap provides in-depth analysis on the attacker's behaviour, to optimize the system's defences and neutralize the attack.

While this program provides ample analysis and useful information to help against cyberattacks, it is used only in real-time applications and real-life situations. CTF games are make-believe and controlled environments designed to teach participants and engage them, without putting any resources like financials or sensitive information up for risks. Additionally, Bit Trap does not support chat features for people to communicate, rather, it functions automatically and can be controlled by its

AI. This program is expensive, so CTF organizers may not be able to use it as often as they should.

Screenshot:

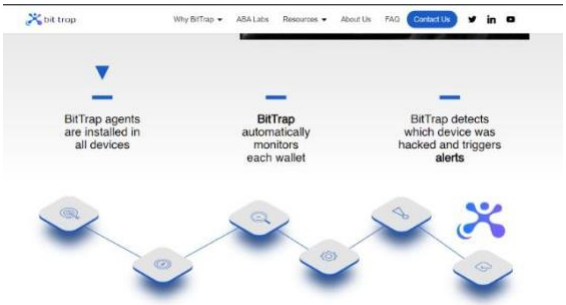

Figure 2.0 displays Bit Trap's platform as shown to users.

2.3 Table of Comparison

Table 1.0 Comparison of Important Features and Components of Similar Systems.

System/Criteria	Discord	Bit Trap
Mobile Application	YES	NO
Chat	YES	NO
AI analysis	NO	YES
IoT	YES	YES

As shown in Table 1, It can be deduced that Discord focuses on easing the process of social communication while utilizing IoT technologies to make the process as seamless as possible. Comparatively, Bit Trap is a more goal-oriented platform that provides analytical insights to ABA while utilizing IoT technologies, However, it does not provide an easy-to-use mobile application that enables its users easy and fast responses.

2.4 Conclusion

Researchers agree that chats are important to ensure communication and better pooling of information [1]. Therefore, it is key to maintain a chat feature in a proposed system. Additionally, a mobile application can prove versatile to make the system more accessible and easier to manage and access. While ABA experts play a major role in the system, an AI element is key in creating a more accurate analysis of the information provided. This can be in the form of graphs, charts and numerical analysis that can ease comparisons and make easier and more accurate deductions. Information of this magnitude must be shared using a cloud service or method of communication that can be established using Internet of Things, which proves to be a vital feature of the system. Discord seems like a similar system where it provides a platform to send messages and conduct meetings on a web and mobile application using IoT elements but lacks the aim of a system that is designed to be used to share analysis and information through comparative and deductive elements like graphs and charts. Additionally, it has no AI feature that allows the system to contribute to the analysis. On the other hand, Bit Trap is designed to monitor and track cybercriminal activity, but is not user-friendly to allow chat, meetings or features a cloud of IoT feature. With these limitation in the market, the requirement for a system that fills all these boxes is clearly shown, as this study introduces.

Reflection

The application of ABA requires vulnerabilities in a live system, only then, accurate results can be retrieved. The issue with this requirement is that no sane company would allow their systems to be hacked purposely in order to further research. To satisfy the technological world's need for ABA with the rising cybercriminal trends while fulfilling the hindering requirement for ABA studies, the solution is to conduct fake scenarios and staged system vulnerabilities. In light of this, CTF events satisfy this solution in their Attack and Defend events where participants must breach a system through exploiting its vulnerabilities in a staged scenario. Thus, ABA can be conducted in correlation with CTF events [8]. However, ethical questions have been raised about how CTF events encourage cyber attacker behaviour, and many participants complain some challenges introduced in the CTF events are not appropriate to their level (too advanced or inappropriate). As organizers find difficulties in arranging these events and generating appropriate and ethical games for the CTF events, ABA can actually assist these organizers as it can advise them on the effects of various challenges and questions and design the most appropriate and successful events.

3.0 Problem Statement

Recent years have witnessed the vast emergence of Capture the Flag (CTF) games in the educational sector of Cybersecurity. These events feature a collection of challenges participants must solve in order to capture a flag, by applying rigorous offensive practices of cryptography, web applications analysis among others. Literature [10] investigates the use of CTF games in university courses, appearing more common as it has been found to be preferred by students and faculty in the academic method. Researchers agree that integrating CTF games plays a dramatic effect in forming a learner's interest and skills in cybersecurity, as this practice involves examining system vulnerabilities, deploying strategies, and developing offensive penetration techniques [11]. CTF events share a major human component, and yet, they rarely integrate attacker behaviour and psychology studies in organizing the event or analysing its aftermath. Kimberly et al. explain how crucial cyber attacker

behaviour is in better understanding how attacks happen and are strategized to develop more resilient, reliable, and adaptive defences (Kimberly et al., 2020) and suggests this may be the key to answering some of cybersecurity's unanswered questions, like why offensive hackers become who they are and how. Additionally, in an effort to better understand the effect of CTF games on learners, research [6] in 2020 explores the ethical problems that may arise when learners are being served system vulnerabilities and taught offensive hacking methods, but overall research is lacking [6]. To better understand these risks, researchers recommend investigating psychological indicators and attacker behaviour of CTF participants to better understand the influence of CTF games on students.

Essentially, CTF games are questions participants must answer by solving riddles of cybersecurity nature, however, organisers often find difficulties creating questions that cater to the learners' demographic. Selecting and designing questions to suite the participants best has proven to be a difficult task [6] and [1]. If not constructed right, these questions may not suit the participants' skill set, or draw ethical concerns toward their offensive nature. Note that these are the very questions that play a major role in the learners' learning experience and training, forming them for who they will be.

Evidently, a platform that allows the integration of attacker behaviour analysis (ABA) with CTF game organizers can greatly assist in creating the more calculated and designed questions, and at the end of the games, analyse the participants results. With ethical issues in question regarding the effects of rigorous offensive activities, organizers can design ethically sound challenges which may not cause ethical concern for encouraging offensive behaviour and creating a dark side of a hacker.

Additionally, the results of the questions can be utilized to better comprehend how hackers respond to system vulnerabilities, and how the human brain creates offensive strategies [4]. This resource has been otherwise wasted, but now can help analysts better understand how hackers behave. Thus, these conclusions can be used to research how to better secure systems and improve defensive measures in the world against cyberhackers.

4.0 Research Aim

This research aims to propose an effective IoT solution to solve issues that arise from the absence of attacker behaviour analysis (ABA) in organizing and analysing CTF games as a part of students cybersecurity education by focusing on the following objectives:

5.0 Research Objectives

- To build the foundation for a platform that helps CTF game organizers create calculated challenges for participants by communicating with ABA experts.

- To encourage the integration of ABA in CTF game environments organizers to facilitate in creating controlled challenges and analyse participants' responses from the CTF games.

6.0 Research Significance

The findings of this research would help benefit 3 main parties in the world of cybersecurity: 1) CTF organizers who find difficulties in creating questions that benefit students and do not raise any ethical issues. By shedding light on this matter, CTF organizers can comprehend how the challenges they create in the CTF events play a key role in shaping attacker behaviour. 2) Students who participate in CTF events: when better challenges are created for these participants, they are more likely to learn more about cybersecurity and forensics defences and gain more profound skills, which is the real intention for these events to be held. As participants have a more productive and enjoyable event, they are more likely to join more. 3) Attacker Behaviour Analysts: CTF games are the perfect scene to conduct analysis on AB and how cyberattacks are strategized. This knowledge can be used to further studies and understand how system defences should be used to prevent attackers more efficiently. Therefore, an integrated platform for CTF organizers and AB analysts to share knowledge will help create successful CTF events and encourage students to learn more effectively and gain positive experiences, while helping analysts understand how hackers behave in controlled environments.

7.0 Research Methodology

7.1 Respondents Identification

The optimal respondents that this study is based on are students within Asia Pacific University, Malaysia that study IT-related programs and have a connection to the computer science world such to become aware/involved in CTF games. While deeper reasoning is elaborated in the Data Collection section, these target participants provide the most relevant input for the study. As established earlier, there is lacking research in the field of this study, and when investigating appropriate research methodologies, quantitative methodologies are

noted to be applied when there is such a case [12]. Additionally, quantitative research is regularly conducted to provide the researcher with a better controlled environment, where they can control variables and research questions and answers respondents provide [12]. In this study, this is most relevant to correlate past research and connect it with new and existing theories. While quantitative research is known to utilize higher number of respondents, it will be hard to interview respondents, thus surveys and numerical data is most appropriate for this study and can provide the insights of students' experience of CTF events for accurate measuring and comprehension of results. In light of this, higher number of respondents is needed to establish reliability and accuracy, which is key for this research. Thus, Quantitative method will be applied here.

7.2 Sampling Method

This study reflects on technical aspects of the educational sector in university, specifically among students participating in CTF events and their experiences. Due to the nature of this study, it is paramount that proper judgement is used in selecting respondents who can contribute their experiences and opinions to be studied. Thus, it would be remiss to use probability in respondent selection, and the sampling method chosen must be purposive judgemental sampling method. It is established that this method selects respondents by choosing individuals that fit a criterion and is applied to researcher bias by targeting students who actively take part in CTF events or have a required level of knowledge and experience to answer the survey accurately and provide computable and meaningful insights [9]. While there is criticism due to bias on the researcher's part, it is believed this bias is in favour of the research to integrate the most relevant subjects for the study [14].

7.3 Data Collection Method: Survey

While most CTF competitions collect rudimentary data that can provide insights to participants' experience, it has no actual contributions from the participants themselves, sourced from automatically stored meta data [4] and instead show participant success based on extracted flags. Alternately, this study aims to explore the attacker behaviour exhibited by participants when taking part in CTF events, which involves the psychological and more social aspects of an experience [5]. This data will be collected from the respondents involved in the study by completing a survey. As this study aims to be quantitative and involve a high number of 100 respondents, alternative collection methods like interviews will be difficult to acquire due to their lengthy nature. Since the study is aimed at APU students (CYB and Forensics), the collection method must be chosen most appropriate to them. these students are faced by time constraints as they are in the middle of a semester, have busy schedules and are overseas from the researcher [13]. Thus, an impersonal but more convenient and quick method is deemed necessary, like a survey, which is the adopted data collection method of this study. Additionally, since this study borderlines on behavioural data, this can become a sensitive topic for some participants to share honestly in an interview, thus, an online e-survey using Google Forms is most appropriate to encourage a safe and judgement-free zone a respondent would feel more comfortable at [13].

The survey's make-up will involve multiple closed-ended questions similar to Likert scaled queries and simple Yes or No answers, while incorporating some open-ended questions, varying on objective and subjective questions for respondents to provide simple explanations for their answers regarding the proposed system.

The survey will undergo a refining process to ensure its accuracy, validity and relevance to the respondents and study, therefore, a pre-test can be initially done with a select number of respondents. Their feedback is thoroughly evaluated to improve the survey for a pilot test, which is conducted with the same respondents. After reviewing their feedback and ensuring the survey is at its best, the survey will be final and ready for the whole sample size of the respondents.

8.0 Overview of The Proposed System

To illuminate onto the proposed system, a use-case diagram was drawn to identify the functionalities of the system and its users. A user login is implemented to identify the users and verify their identities, as sensitive information will be used. Additionally, different prefaces of the system must be utilized to account for different types of users accessing the system as they use functionalities related to their role only.

In this research, a use-case was implemented, which is a diagram used to provide a view of the whole system functionalities in relation to their users, also known as actors, who use the system for their purpose and how they interact with it. System developers create use-cases as a steppingstone to

create the system design before the later steps of system development [3].

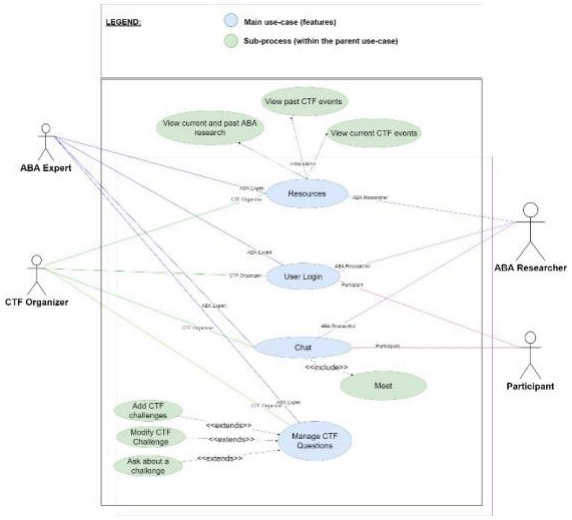

Figure 4.0: Use Case Diagram of the Proposed System

Figure 4 depicts the Use-case diagram illustrating the proposed system functionalities and users showing the functionalities of the system in blue, and the users (actors) outside the black rectangular outline that represents the full extent of the system, and anything outside is an external user.

The proposed system users are:

1- ABA Expert:

This user can login as an ABA expert and will have access to resources including CTF events and ABA research to review and refer to reliable information at hand. These resources can be stored in a virtual cloud/library for secure access. Additionally, since these experts are qualified and verified, they have access to the CTF challenges to review and modify them with the organizers. Since this platform encourages communication, ABA experts can also chat with other users in the system, discuss information and meet with other professionals. This is to encourage the honest and secure information exchange between users. This user is indicated in blue.

2- CTF Organizer

This user login as a CTF organizer. They have the same access as ABA experts and can browse through resources, chat, and meet with users and modify CTF challenges. This user is coloured green.

3- ABA Researcher

This user logs in as an ABA researcher. Since this system encourages the integration of ABA into CTF events and their application in cybersecurity education, new up-and-coming researchers are encouraged to use the system and take advantage of it. However, since they do not qualify as experts, they are limited to accessing resources and chatting with other users. This way, new researchers can collect information and conduct studies without interfering with the CTF events or challenges. This user is coloured purple.

4- Participant

This user can log in as a participant and converse with other users to share information and experiences. They have no other role to play to ensure the validity of the resources and CTF questions. This user is coloured orange.

9.0 Conclusion

As CTF events become more common in training young cybersecurity experts and cyberattacks grow more violent, cybersecurity becomes more relevant to today's world. ABA serves as a way for companies to defend against cyberattacks, therefore, contributing to this topic will strengthen and support it. While CTF organizers find difficulties creating adequate and ethical challenges for participants, this research aims to integrate them with ABA experts in a platform to assist them in creating the challenges while contributing to ABA research.

By conducting further research into ABA and connecting CTF events to cyber attacker behaviour, this realm can be amended to further cyber defence strategies. Developing a system that allows for the integration of CTF and ABA is recommended to understand how these subjects interlock into the world of cybersecurity.

Future recommendations pertaining to the system proposed imply deploying a system that provides both chat features that allow communication while enabling an AI analytical program that assists in data processing. Furthermore, supporting the platform for mobile use is a good way to make the technology more accessible to its users.

Reference

[1] Carlisle, B., Reininger, M., Fox, D., Votipka, D., & Mazurek, M. L. (2020). On the other side of the table: Hosting capture the flag (ctf) competitions. In Proceedings of the 6th Workshop on Security Information Workers, ser. WSIW (Vol. 20). https://wsiw2020.sec.uni-hannover.de/downloads/WSIW2020-On%20the%20Other%20Side%20of%20the%20Table.pdf

[2] Chung, K., & Cohen, J. (2014). Learning obstacles in the capture the flag model. In 2014 USENIX Summit on Gaming, Games, and Gamification in Security Education (3GSE 14). https://www.usenix.org/conference/3gse14/summit-

program/presentation/chung

[3] Ruparelia, N. B. (2010). Software development lifecycle models. ACM SIGSOFT Software Engineering Notes, 35(3), 8-13. https://sci-hub.st/10.1145/1764810.1764814

[4] Ferguson-Walter, K., Major, M., Van Bruggen, D., Fugate, S., & Gutzwiller, R. (2019, December). The World (of CTF) is Not Enough Data: Lessons Learned from a Cyber Deception Experiment. In 2019 IEEE 5th International Conference on Collaboration and Internet Computing (CIC) (pp. 346-353). IEEE. https://ieeexplore.ieee.org/abstract/document/8998505/

[5] Hanafi, A. H. A., Rokman, H., Ibrahim, A. D., Ibrahim, Z. A., Zawawi, M. N. A., & Rahim, F. A. (2021). A CTF-Based Approach in Cyber Security Education for Secondary School Students. electronic Journal of Computer Science and Information Technology, 7(1). http://103.227.140.18/index.php/ejcsit/article/view/107

[6] Holmi, J. (2020). Advantages and challenges of using capture-the-flag games in cyber security education. http://jultika.oulu.fi/files/nbnfioulu-202009102920.pdf

[7] Padilla, E., Acosta, J. C., & Kiekintveld, C. D. (2021). Cybersecurity Methodology for Specialized Behaviour Analysis. In International Conference on Digital Forensics and Cyber Crime (pp. 237-243). Springer, Cham. https://link-springer-com.ezproxy.apiit.edu.my/chapter/10.1007/978-3-030-68734-2_14

[8] Peng, X., & Zhao, H. (2007, June). A framework of attacker centric cyber-attack behaviour analysis. In 2007 IEEE International Conference on Communications (pp. 1449-1454). IEEE.

[9] Sharma, G. (2017). Pros and cons of different sampling techniques. International journal of applied research, 3(7), 749-752. https://d1wqtxts1xzle7.cloudfront.net/58765080/Pros_and_cons_of_sampling-with-cover-page-v2.pdf?Expires=1644313598&Signature=So5I2stwuIOg32CN57oOxpw9Rd9pDAMJ~KbdDP0zpwe4FOkZJdcSxxcytWD2SnZqOe2eq199kmqLcnvDb9nsc48RmtVuAXNSWat-iW3KLMKMU5Ls~trStJoJtAB5XKYNv1Ln8RHw8y-jmqeS0xjodsVlNjt0Cm9swQ5AwABkXsPcy0Qa~kIWfdibsj-2CbaqUafGiQcMzVSLCIgx8rHm~zqH5aZaHDkKtZgTS3MxGodMtKpopnWhXGF89w9Y3WMJVjUy39Er3RIC0jNt4MBZDVQBZDLJ16T6xB-2yXTxI5I7ahb05hUHkcdRJ7yb~4gqNeZWcOeig~a249pOIDXXtg_&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA

[10] Vykopal, J., Švábenský, V., & Chang, E. C. (2020, February). Benefits and pitfalls of using capture the flag games in university courses. In Proceedings of the 51st ACM Technical Symposium on Computer Science Education (pp. 752-758). https://dl-acm-org.ezproxy.apu.edu.my/doi/abs/10.1145/3328778.3366893

[11] Doynikova, E., Novikova, E., & Kotenko, I. (2020). Attacker behaviour forecasting using methods of intelligent data analysis: A comparative review and prospects. Information, 11(3), 168. https://www.mdpi.com/2078-2489/11/3/168

[12] Rutberg, S., & Bouikidis, C. D. (2018). Focusing on the fundamentals: A simplistic differentiation between qualitative and quantitative research. Nephrology Nursing Journal, 45(2), 209-213. https://www-proquest-com.ezproxy.apu.edu.my/intermediateredirectforezproxy

[13] Rhodes, S. D., Bowie, D. A., & Hergenrather, K. C. (2003). Collecting behavioural data using the world wide web: considerations for researchers. Journal of Epidemiology & Community Health, 57(1), 68-73. https://jech.bmj.com/content/57/1/68.short

[14] Elfil, M., & Negida, A. (2017). Sampling methods in clinical research; an educational review. Emergency, 5(1). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5325924/