Identifying the Root Cause of a Cyber Attack Through Log Data Analysis: An Overview of the Challenges Faced by Novice Learners

Priya V. Nagvekar, Syaamantak Das(✉), and Sridhar Iyer

Centre for Educational Technology, Indian Institute of Technology Bombay, Mumbai, India
{priyalotlikar,syaamantak.das,sri}@iitb.ac.in

Abstract. Cyber security has become a critical concern in today's digital age, and log data analysis plays a pivotal role in investigating the root causes of cyber attacks. However, the teaching and learning of the log data analysis process is intricate, especially in root cause analysis (RCA), necessitating diverse domains of expertise and background knowledge. This study addresses a significant research gap in the existing literature by a comparative analysis of cyber attack RCA performed by novices and experts through log data analysis. The study design process initially involved novices with limited cyber security experience (senior undergraduate Computer Science students) performing a log data analysis of a cyber attack, followed by industry professionals (subject matter experts) performing the same log data analysis. In the research, the participants identified the root causes of a cyber attack through log data analysis, culminating in generating an attack tree as an outcome. The objective was to assess novices' ability to identify and deduce the root causes of cyber attacks from log data and generate a comprehensive attack tree through causal reasoning. The follow-up observation and reflections were supplemented by artefact analysis and simulated recall interviews. The study's findings shed light on the disparities in log data analysis skills between novices and subject matter experts. It reveals novices' key challenges compared to experts and the nature of novices' difficulties. The results of this research also offer valuable insights into the industry practices of log data analysis by experts and their advice contributing to developing more effective teaching-learning methodologies for cyber security education.

Keywords: cyber security · log data analysis · novice vs expert · root cause analysis

1 Introduction

Cyber security has become a critical concern in today's digital age, as cyber-attacks pose severe threats to individuals, businesses, and national security.

A. Karkare et al. (Eds.): COMPUTE 2024, CCIS 2400, pp. 147–160, 2025.
https://doi.org/10.1007/978-3-031-84391-4_12

Effective incident response and root cause analysis (RCA) of cyber-attacks are crucial for mitigating their impact and preventing future occurrences. RCA involves investigating the underlying causes of a security breach and identifying vulnerabilities that enabled the attack. However, performing RCA is a complex task that requires diverse domains of expertise and background knowledge, making it challenging for novice learners in cyber security. This study addresses a significant research gap by analyzing the differences between novice learners and industry professionals (subject matter experts) in performing RCA of cyber attacks using log data analysis. The primary objective is to assess novices' ability to identify and deduce the root causes of cyber attacks from log data and generate comprehensive attack trees through causal reasoning.

While there is a substantial body of literature on RCA methods and techniques, there is a lack of comparative evidence on the processes followed by experts and the specific challenges faced by novices in this domain. Understanding these differences is crucial for developing effective teaching-learning methodologies and bridging the gap between novice and expert performance in cybersecurity education. The study focuses on the following research question:

What are the various challenges and difficulties encountered by novice learners when undertaking the task of performing RCA?

To address these research questions, the study employs a mixed-methods approach. Novice participants (senior undergraduate Computer Science students) and industry professionals analyze simulated cyber attack log data using RCA methods such as the 5 Whys, fault tree analysis, fishbone diagrams, and attack tree creation. Observations, reflections, artefact analysis, and simulated recall interviews are used to collect both quantitative and qualitative data, providing insights into the decision-making processes and challenges faced by novices and experts. By identifying the root causes of the difficulties faced by novice learners and contrasting their approaches with those of experts, this study contributes to the ongoing efforts to enhance cybersecurity education and practice.

This paper is organized as follows: Sect. 2 briefly describes the existing literature and how the work connects to the skills of RCA. Section 3 details the design of the study, participants, resources used and experimental phases. Our findings are presented and discussed in Sects. 4 and 5 while highlighting various challenges faced by novices and industry experts' opinions, including the limitations involved with the study in subsections. The paper is concluded in Sect. 6.

1.1 Problem Background: Log Data Analysis Techniques for RCA

Performing vulnerability assessment through RCA is a crucial step in incident response after a cyber attack. RCA involves systematically investigating the underlying causes of a security breach to identify vulnerabilities and develop mitigation strategies [18]. In a rapidly evolving cybersecurity environment, a multifaceted approach to analyzing cyberattacks is critical to understanding their nature, origin, and potential impact. This study uses a comprehensive set of analytical tools to analyze and interpret cyber attack data, including reading

log files, the 5 Whys technique, fault tree analysis, fishbone diagrams, and attack trees. The following methods show how log data can be used for RCA.

- **Log file reading (Identifying tags):** It involves examining detailed records of system events focusing on key elements such as UserID, URLs, requests, responses, timestamps and duration. This method helps in identifying patterns and anomalies associated with cyber attacks [24].
- **5 Whys Technique:** It is a root cause analysis technique that involves repeatedly asking "why" questions to dig deeper into the root cause of a problem. In cyber security, this helps uncover the underlying reasons for an attack and go beyond the surface-level symptoms to identify key vulnerabilities [25].
- **Fault Tree:** Constructing a fault tree for a scenario is a deductive failure analysis method. It helps to visualize the logical combinations of different factors that could result in a successful attack [26].
- **Fishbone Diagram:** Also known as the Ishikawa diagram, is a cause-effect diagram that helps in identifying potential causes for a problem. In cyber security, it can be used to categorize and analyze various factors that contribute to security breaches, such as environmental factors, people, technology and processes [27].
- **Attack Tree:** It is a conceptual diagram that shows how an asset can be attacked through various routes. Starting with the target as the root node, the tree branches out to show different ways as attacker can reach the target, which helps in threat modeling and risk assessment [28].

Various frameworks, such as NIST SP 800-61, the SANS Incident Handling Process, and ISO/IEC 27035 [35], provide guidelines for conducting RCA and analyzing log data to determine the appropriate response.

2 Literature Review

While the importance of RCA in cyber security is well-established, the existing literature primarily focuses on the technical aspects and methodologies involved. However, there is a lack of research exploring the challenges faced by novice learners in acquiring and applying these skills effectively. Most studies in this domain concentrate on expert practitioners' perspectives and industry best practices [14,23,36], overlooking the unique difficulties encountered by novices during their learning journey.

Several studies have investigated the differences between novice and expert performance in cybersecurity-related tasks. For instance, Silva et al. [34] examined expert and novice performance within computer security incident response teams, while Koutcheme et al. [21] explored how students solve open-ended assignments in SQL injection. However, these studies do not specifically address the challenges faced by novices in performing RCA using log data analysis, which is a critical component of incident response and vulnerability assessment.

While the literature acknowledges that log analysis is an essential part of RCA [35], there is a lack of research investigating the specific difficulties novice

learners encounter when analyzing log data to identify root causes. A study was done by Scheponik et al. [36] where they interviewed 26 learners on cyber security concepts. Related similar cyber security studies were the work of Silva et.al [34] which focused on Expert-Novice performance in cyber security incidence response and the study by Koutcheme et al. [21] which explored how students solve open-ended assignments in SQL injection. Most existing studies focus on technical aspects, such as log analysis tools [1,23] or machine learning techniques [4], rather than the cognitive and pedagogical challenges faced by novices in this domain. Additionally, the literature lacks a comparative analysis of the approaches and strategies employed by experts and novices when performing RCA using non-technical methods, such as the 5 Whys, fishbone diagrams, fault tree analysis, and attack tree creation. While these methods are often recommended for novice learners [12,14,22], there is limited empirical evidence on the specific challenges encountered by novices in applying these techniques and the differences in approaches compared to experts.

Addressing this research gap is crucial for developing effective teaching-learning methodologies and interventions to bridge the gap between novice and expert performance in cybersecurity education. By identifying the root causes of the difficulties faced by novice learners and contrasting their approaches with those of experts, this study aims to provide valuable insights for enhancing introductory-level cyber security education, particularly in the context of log data analysis and RCA.

3 Study Design

In order to address the research questions, it was suggested to utilize a mixed-methods approach [6]. The study process involves several stages: Stage 1 includes providing log data of cyber attacks to both novice students and industry experts. Stage 2 consists of an orientation session, detailing the specific instructions for student participants (novices) and industry experts, respectively. Following the orientation, Stage 3 involves analyzing log data using the 5 why method, identifying and creating the fault tree, also focusing on constructing a fishbone model by identifying all potential causes, and developing an attack tree to identify the root cause of the attack. The process was followed by stage 4 - observations and reflections by the participants which were supplemented by artefact analysis [34] and a simulated recall survey [7]. Figure 1 shows the entire study design process.

For this study, the quantitative data is collected through the log data analysis and the fishbone and attack tree models, while the qualitative data is collected through observations, reflections, artefact analysis, and simulated recall interviews. The 5 Why method and fault tree analysis are quantitative techniques that help to identify the root cause of the attack. The fishbone model and attack tree are visual tools that help to organize and analyze the data. The observations and reflections of the participants provide qualitative insights into the decision-making process and the effectiveness of the techniques used. The artefact analysis and simulated recall interviews validate the findings and provide additional insights into the participants' experiences and perspectives.

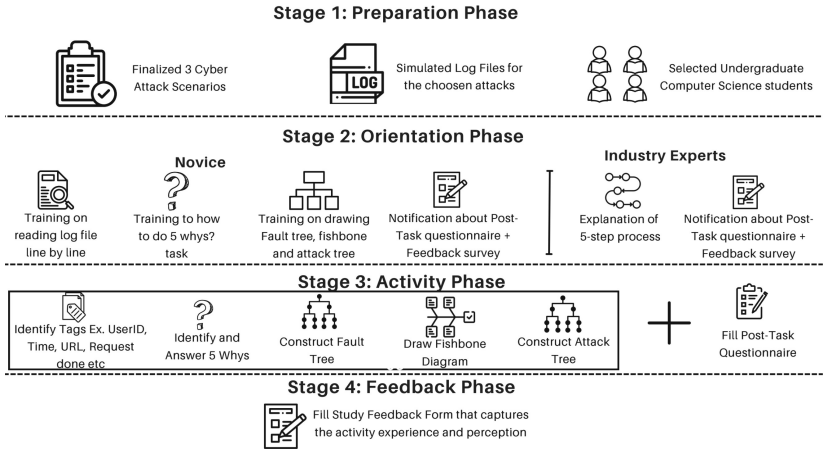

Fig. 1. Study design process

3.1 Datasets and Resource Material

To ensure a fair assessment of expertise, participants were required to have completed standard undergraduate-level Computer Science Network courses (semester subjects) as a prerequisite. This baseline knowledge was deemed essential for all participants. The study included a diverse group of Indian students, categorized as follows in Table 1. This breakdown of participants allows for a comparison between novice and expert approaches to cyber attack log analysis.

Table 1. Participant Profile.

Participant Type	Gender	Age Range	No.
Novice	Male	21–24	17
Novice	Female	21–24	7
Expert	Male	35–40	3

3.2 Sample Log Data

Below is a sample log file of a simulated DoS attack:

```
Timestamp: 2023-11-05T13:03:12-08:00
IP Address: 192.168.1.100
User Agent: "Mozilla/5.0(X11; Linux x86\_64)AppleWebKit/537.36
(KHTML, like Gecko) Chrome/107.0.0.0 Safari/537.36"
Request: GET /index.php HTTP/1.1
Response: 200 OK
```

This log file records a single HTTP request for the page index.php from the IP address 192.168.1.100. The response code is 200 OK, indicating a successful request.

Root Cause: The root cause of the DoS attack is the attacker flooding the server with numerous requests for the same page. This overloads the server, rendering it inaccessible to other users.

The following logfile parameters help determine the root cause of the DoS attack:

- **Timestamp:** Shows when the request was received.
- **IP Address:** Identifies the attacker's location.
- **User Agent:** Reveals the browser and operating system the attacker is using i.e. tools being used.
- **Request:** Indicates the specific page the attacker is targeting, aiding in pinpointing the vulnerability being exploited.
- **Response:** The HTTP response code 200 OK suggests that the attacker successfully exploited the vulnerability.

Analysis of web server log files can reveal patterns indicative of Denial of Service (DoS) attacks, distinguishing them from regular web traffic. In this study, we examined log data from an experimental setup designed to demonstrate DoS attack patterns to students. The logs exhibited several key characteristics consistent with a DoS attack: high-frequency requests from a single IP address, with multiple identical GET requests targeting the same URL and using the same user agent, all occurring within seconds. Notably, these requests consistently resulted in 499 status codes, suggesting the server's inability to process the high volume or potentially invalid nature of the requests. The uniformity of these requests—identical IP, URL, and user agent—maintained over an extended period, as evidenced by the broader dataset, serves as a significant indicator of malicious activity. While log files do not inherently differentiate between regular access and DoS attacks, the presence of these patterns provides strong evidence for distinguishing between benign traffic and potential security threats.

3.3 Experiment Phase

Phase 1: Pilot Study Phase. This study was conducted as a proof of concept, to get better clarity on various aspects of complexity that a student might face while doing the activity. It was an open resource study, where participants were allowed to access the internet to seek help on certain concept explanations. However, they had to mention in their answer sheets wherever they used the internet for help mentioning the details. Approximately 120 min were given to a student to perform the entire activity.

- **Participants: -** 7 novices and 1 expert
- **Purpose: -** Face validity testing

– **Task: -** Participants were given log files of three distinct cyber attacks(SQL Injection, XSS, DoS) to identify the tentative root cause of an attack using the 5 Why method. They were also tasked with generating a Fault tree, preparing a fishbone model, and creating an Attack tree. This was followed by a post-task questionnaire to gather their observations and reflections. The questionnaire consisted of 9 open-ended questions based on the five methods performed and the challenges the participant faced while working with the 5 Why methods. For instance, *"What are the five whys you identified, and why did you choose those?", "What Challenges did you encounter when using the 5 whys method and why?"*. Participants were needed to answer the questionnaire after every attack analysis, thus having 3 responses for all three attacks. Once the analysis task was completed, all the participants had to answer the feedback survey, which was designed to gather the various challenges faced while doing the activity using all 5 methods and their perspective on the entire RCA activity. The feedback form comprised of 16 choice-based and 1 open-ended questions.

Phase 2: Extensive Study Phase

– **Participants: -** 24 novices and 3 experts
– **Task: -** Similar to Phase 1, participants were asked to identify the root cause of an attack using the 5 Why method, generate a Fault tree, prepare a fishbone model, and create an Attack tree. However, this phase included structured questions for each method. There were two sets of questions: Question Set 1 focused on the procedural aspects of the task, for e.g. *"What were the 5 Whys you identified, and why did you choose those?"*; while Question Set 2 addressed the challenges and obstacles faced by the participants during task execution, regardless of the outcome. For e.g. *"What challenges did you encounter when using the 5 Whys method, and why?"*.

4 Data Analysis and Result

4.1 Methodology

The research methodology employed a mixed-methods approach, combining quantitative and qualitative analysis to investigate challenges faced by novice students in the cybersecurity domain. The data sources used for analysis included student and industry expert answer sheets, post-task questionnaire responses and post-study feedback responses from both student and industry experts. Initially, 4 teacher interviews were conducted to identify student challenges in cyber security problem analysis, which followed the approach suggested by Seidman [29]. These insights, along with a comprehensive literature review [30–32], informed the development of a set of categories and themes as shown in Table 2 capturing various challenges faced by students and teachers. This framework was then used as the basis for a deductive coding process on student and Industry experts' post-task questionnaires and feedback responses as outlined by Saldana

[33]. The deductive coding process aimed at identifying common categories and themes of challenges that students could encounter during the RCA process and also identifying the gap between novice and industry expert approaches for RCA. To identify challenges faced by novices while doing RCA, the following meta-analysis approaches were employed on the student answer sheets: 1) Structural Analysis: i) Examining the completeness of each diagram. ii) Evaluating the coherence and logical flow of the 5 whys responses. 2) Content analysis: i) Analyze the accuracy and relevance of identified causes and effects. ii) Evaluate the use of cyber security terminology and concepts. 3) Pattern identification: i) identify common mistakes or misunderstandings across artefacts. ii) Identifying gaps in knowledge. iii) Identify patterns of misinterpreted log data. 4) Time-based analysis: analyzing time spent on each method. 5) Thematic analysis: Identify recurring predefined themes and categories.

Table 2. An Overview of Challenges faced by Novices - Themes and Categories

Themes	Categories
Teaching Methods	Traditional Teaching methods, Practical Teaching
Learning Challenges	Knowledge Gap, Curriculum constraints, Teaching challenges
Pre-requisite Knowledge	Programming skills, OS understanding, CS foundations
Motivation	Interest and importance, Emerging career options, Passion for cyber security

4.2 Result

Time Duration Analysis for Each Attack. The distribution of time for experts was more consistent, as evidenced by the lower standard deviation (17.85) compared to novices (32.60). Both groups had a similar minimum time (15 min), indicating that the fastest individuals in each group solved problems at a comparable speed. Table 3 shows a comparison of the time taken for problem analysis between Novice and Experts.

Table 3. Comparison of Time taken for problem analysis between Novice and Experts

Metrics	Group	
	Novice	Experts
Mean(mins)	57.05	41.67
Standard Deviation(mins)	32.60	17.85
Min(mins)	15	15
Max(mins)	141	60
25th Percentile(mins)	35	25
50th Percentile (Median - mins)	45	45
75th Percentile(mins)	85	60

However, the maximum time for novices (141 min) was much higher than for experts (60 min), suggesting that some novices took considerably longer to solve problems. The median time for both groups was 45 min, indicating that half of all participants, regardless of their expertise, solved the problems within this time frame.

The first problem (SQL injection) proved to be the most time-consuming, with each student spending over an hour on it. However, the time required for subsequent attacks (XSS and DoS) showed a significant reduction, indicating a learning curve as participants became more familiar with the analytical methods.

5 Discussion

5.1 Challenges Faced by Novices

To seek answers to the given Research Question and get deeper insights into the challenges faced by novices while doing RCA using log files, the study analyzed the post-task questionnaire. The key findings of the meta-analysis on challenges faced by novice learners are as follows:

1. Difficulty in identifying the attack activities i.e. difficulty in understanding the meaning, which required additional effort to comprehend each line of the log file and the mode of attack.
2. Superficial Analysis: Stopping the analysis too early and not addressing deeper root causes. Not digging deep enough to uncover the root cause.
3. Assumption and Bias: Influence of preconceived notions or biases on the analysis.
4. Single-Cause Bias: Oversimplification of complex problems by attributing them to a single cause.
5. Lack of Cross-Functional Input: Narrow analysis due to not involving various other computer science concepts.
6. Not familiar with XSS-based malware attack. Almost 71% of students found this as a major challenge.
7. Complexity of the System: Difficult to capture all potential failure modes accurately in complex systems with numerous dependencies and interactions.
8. Assumption of Linear Causality: Oversimplification by assuming linear cause-and-effect relationships, potentially overlooking interconnected factors. This also shows a lack of critical thinking ability.
9. Inadequate Expertise: Lack of necessary expertise to conduct a thorough analysis.
10. Insufficient Information: Incomplete or ambiguous data due to limited access to information, complex systems, or lack of comprehensive documentation.

At the end of the study, the student also had to answer a feedback survey which was designed to gather their experience and perception about the RCA activity using log files. Based on the analysis of student responses, four main themes were identified: teaching challenges, knowledge gaps, curriculum constraints, and resource limitations. The findings reveal that teaching challenges

were the most prominent issue, accounting for 43.9% of the responses. This was closely followed by knowledge gaps, which represented 39.0% of the feedback. Curriculum constraints were cited in 14.6% of the responses, indicating a less significant but still notable concern. Resource limitations appeared to be the least prevalent issue, mentioned in only 2.4% of the responses. Figure 2 demonstrates the graphical representation of the analysis. These results suggest that improvements in teaching methods and addressing students' knowledge deficits should be prioritized to enhance the effectiveness of log analysis activities also emphasizing the need for more orientation sessions on how to read and interpret log files, while curriculum adjustments and resource allocation may require attention but to a lesser extent.

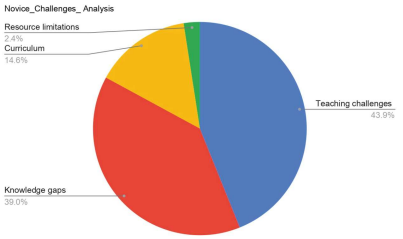

Fig. 2. Challenges faced by Novice learners

5.2 Analysis of Industry Experts Study

Based on the expert analysis, the methods for RCA of cyber attack i.e. 5 whys, fault tree, fishbone and attack tree, using log files each have their strengths and limitations. Common challenges across these 5 methods as mentioned by the experts in their responses include incomplete data resources, time constraints, and the need for deep technical knowledge. The analysis suggests that these methods can be subjective, potentially leading to biased interpretations. They also struggle to capture the full complexity of modern cyber attacks, which often involve multiple layers of technology and diverse techniques. The experts noted that creating comprehensive analyses is challenging without complete information, and that staying updated with emerging attack techniques is crucial. Based on the analysis of industry experts' responses to the post-activity feedback survey, several key themes emerged. The most significant issue identified was the lack of expertise, accounting for 15.6% of the responses. This was followed by three equally prominent themes: lack of foundational knowledge, log data analysis tools, and simulation demos, each representing 9.4% of the feedback. Several themes tied at 6.3% of responses, including cyber security upskilling, continuous learning, emerging challenges, introductory SIEM education, and real-world examples. Less frequently mentioned but still notable were themes such as infrastructure upgradation, curriculum enhancement, student growth and success, pas-

sion for cyber security, interest and importance, operating systems understanding, networking expertise, and cyber security foundations, each accounting for 3.1% of the responses. Figure 3 summarizes the analysis graphically. These findings suggest that while addressing the lack of expertise is crucial, a multi-faceted approach encompassing foundational knowledge, practical tools, and hands-on experience is necessary to improve log analysis education and training in the cyber security field.

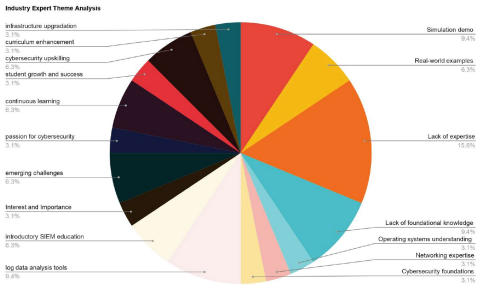

Fig. 3. Themes from industry expert analysis

Suggestions by Industry Experts. In response to the open-ended question in the feedback form *"What additional resources or support would help a novice better understand and apply these analysis methods to cyber attacks?"*, the following themes emerged:

1. Use of tools in classroom - experts recommend teaching tools like machine learning techniques [37], and SIEM tools [1,23] to analyze log data.
2. Inclusion of log analysis in the curriculum - experts proposed revising the digital forensics curriculum to include data analysis.
3. Experts also suggest More practices and case studies with log data.

5.3 Takeaway for the Computer Science Education Community: Pedagogy Perspective

Novice learners face distinct challenges in cyber security problem solving compared to experts. The analysis reveals the following pedagogical factors:

1. Lack of Foundational Knowledge: The pieces of evidence indicate a significant gap in basic cyber security knowledge among novice learners, which hinders their ability to tackle complex problems effectively. The frequent mention of "*lack of theoretical knowledge*" in student responses, coupled with the industry perspective emphasizing the need for concept understanding in Networks and Data Communication, underscores the critical need for a stronger theoretical base in cyber security education.

2. Technical Complexity and Comprehension: Novice participants struggle with interpreting technical data, in particular the log files. This difficulty is evident from both students ("*Wasn't able to understand the URL so I used Chatgpt*")and industry ("*Difficulty in interpreting data*") responses. The fact that some students adopted tools like ChatGPT for understanding URLs suggests a need for more training sessions on how to read and interpret log files, as comprehending and analysing log data is fundamental in cyber security.
3. Methodology Implementation: The evidence shows that novices struggle with formulating appropriate "Why" questions and constructing logical flow charts ("*finding the proper questions starting from why was challenging.*", "*To link the questions that sequence wise explains the attack.*"). This indicates that while the students are taught about these methodologies, they need more practice and guidance in applying these methods to real-world scenarios.
4. Analytical Thoroughness: The analysis reveals that novices face difficulties in connecting events and synthesizing information into a coherent analysis. Given that cyber security issues/attacks are often interconnected, the challenge of synthesizing information is notable. As reported by the student, "*Cannot synthesize the attack information*", mandates the need for developing a more holistic view of cyber security scenarios.
5. Information Synthesis and Interpretation: Developing a comprehensive analytical model from sparse information is a skill that novices find challenging. The evidence of students being "*unable to create final output*" points out to a gap in their ability to make informed inferences and construct meaningful analyses from limited data resources. In real-world cyber security attack scenarios, this skill is crucial where complete information is rarely available.

5.4 Limitations of the Study

The study has limitations due to the small size of the expert group, the use of simulated data, and insufficient training sessions for students. However, it provides valuable insights for cyber security education and suggests the need for further research to address challenges faced by novice learners. Future research could involve expanding the scope of participants from Indian undergraduate students to computer science students worldwide.

6 Conclusion

This research contributes in the quest to enhance cyber security education and practice through an in-depth analysis of novice challenges and industry expert insights. The study identifies the areas where novices face the most difficulties and highlights essential elements for effective log analysis.

The findings emphasize on the identification of key parameters for analyzing cyber attack log data. By addressing the cognitive biases and technical obstacles that novices encounter, this research also offers valuable guidance for educators and curriculum developers to enhance cybersecurity training programs.

References

1. Abbott, R.G., McClain, J., Anderson, B., Nauer, K., Silva, A., Forsythe, C.: Log analysis of cyber security training exercises. Procedia Manufacturing **3**(2015), 5088–5094 (2015)
2. Ahmad, A., Desouza, K.C., Maynard, S.B., Naseer, H., Baskerville, R.L.: How integration of cyber security management and incident response enables organizational learning. J. Assoc. Inf. Sci. Technol. **71**(8), 939–953 (2020)
3. Aldawood, H., Skinner, G.: Educating and raising awareness on cyber security social engineering: a literature review. In: 2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE), pp. 62–68. IEEE (2018)
4. Švábensky, V., Vykopal, J., Čeleda, P., Tkáčik, K., Popovič, D.: Student assessment in cybersecurity training automated by pattern mining and clustering. Educ. Inf. Technol. **27**(7), 9231–9262 (2022)
5. Angelini, M., Blasilli, G., Catarci, T., Lenti, S., Santucci, G.: Vulnus: visual vulnerability analysis for network security. IEEE Trans. Vis. Comput. Graph. **25**(1) , 183–192 (2018)
6. Buchanan, L., D'Amico, A., Kirkpatrick, D.: Mixed method approach to identify analytic questions to be visualized for military cyber incident handlers. In: 2016 IEEE Symposium on Visualization for Cyber Security (VizSec), pp. 1–8. IEEE (2016)
7. Burden, S., Topping, A., O'Halloran, C.: The value of artefacts in stimulated-recall interviews. Nurse Researcher **23**, 1 (2015)
8. Chockalingam, S., Pieters, W., Teixeira, A., Khakzad, N., van Gelder, P.: Combining Bayesian networks and fishbone diagrams to distinguish between intentional attacks and accidental technical failures. In: Cybenko, G., Pym, D., Fila, B. (eds.) GraMSec 2018. LNCS, vol. 11086, pp. 31–50. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-15465-3_3
9. Conti, M., Dragoni, N., Lesyk, V.: A survey of man in the middle attacks. IEEE Commun. Surv. Tutor. **18**(3), 2027–2051 (2016)
10. Anderson Bergamini de Neira, Burak Kantarci, and Michele Nogueira. 2023. Distributed denial of service attack prediction: Challenges, open issues and opportunities. Computer Networks 222 (2023), 109553
11. Dong, S., Abbas, K., Jain, R.: A survey on distributed denial of service (DDoS) attacks in SDN and cloud computing environments. IEEE Access **7**(2019), 80813–80828 (2019)
12. Fantin, I.: Applied Problem Solving: Method, Applications, Root Causes, Countermeasures, Poka-Yoke and A3 (2014). Ivan Fantin
13. Flores, R., Namin, A.S., Tavakoli, N., Siami-Namini, S., Jones, K.S.: Using experiential learning to teach and learn digital forensics: educator and student perspectives. Comput. Educ. Open **2**(2021), 100045 (2021)
14. Furnell, S.: The cybersecurity workforce and skills. Comput. Secur. **100**(2021), 102080 (2021)
15. Goel, S.: National cyber security strategy and the emergence of strong digital borders. Connections **19**(1), 73–86 (2020)
16. Grammatikakis, K.P., Koufos, I., Kolokotronis, N., Vassilakis, C., Shiaeles, S.: Understanding and mitigating banking trojans: from zeus to emotet. In: 2021 IEEE International Conference on Cyber Security and Resilience (CSR), pp. 121–128. IEEE (2021)

17. Hatfield, J.M.: Social engineering in cybersecurity: the evolution of a concept. Comput. Secur. **73**(2018), 102–113 (2018)
18. Hellesen, N., Torres, H., Wangen, G.: Empirical case studies of the root-cause analysis method in information security. Int. J. Adv. Secur. **11** (2018)
19. Kanta, A., Coisel, I., Scanlon, M.: A survey exploring open source Intelligence for smarter password cracking. Forensic Sci. Int. Digit. Investig. **35**(2020), 301075 (2020)
20. Khan, F., Kim, J.H., Mathiassen, L., Moore, R.: Data breach management: an integrated risk model. Inf. Manag. **58**(1), 103392 (2021)
21. Koutcheme, C., Tilantera, A., Peltonen, A., Hellas, A., Haaranen, L.: Exploring how students solve open-ended assignments: a study of SQL injection attempts in a cybersecurity course. In: Proceedings of the 27th ACM Conference on on Innovation and Technology in Computer Science Education, vol. 1, pp. 75–81 (2022)
22. Lallie, H.S., Debattista, K., Bal, J.: A review of attack graph and attack tree visual syntax in cyber security. Comput. Sci. Rev. **35**(2020), 100219 (2020)
23. Landauer, M., Skopik, F., Wurzenberger, M., Rauber, A.: System log clustering approaches for cyber security applications: a survey. Comput. Secur. **92**(2020), 101739 (2020)
24. Kent, K., Souppaya, M.: Guide to Computer Security Log Management. NIST Special Publication 800-92 (2006)
25. Serrat, O.: The five whys technique. In: Knowledge Solutions, pp. 307–310. Springer, Singapore (2017)
26. Ericson, C.A.: Fault tree analysis – a history. In: Proceedings of the 17th International System Safety Conference (1999)
27. Ishikawa, K., Loftus, J.H.: Introduction to quality control, vol. 98, p. 31. Tokyo: 3A Corporation (1990). https://www.scirp.org/reference/referencespapers?referenceid=1272669, June 2024
28. Schneier, B.: Attack trees. Dr. Dobb's J. **24**(12), 21–29 (1999)
29. Seidman, I.: Interviewing as qualitative research: a guide for researchers in education and the social sciences. Teachers college press, 2006
30. Triplett, W.J.: Addressing cybersecurity challenges in education. Int. J. STEM Educ. Sustain. **3**(1), 47–67 (2023)
31. Cybersecurity in Higher Education: Problem and Solutions | Toptal®, May 2024. https://www.toptal.com/insights/innovation/cybersecurity-in-higher-education
32. Top Challenges in Faced by Students in Cybersecurity Education. https://3university.io/challenges-faced-by-students-in-cybersecurity/. Accessed May 2024
33. Saldaña, J.: The coding manual for qualitative researchers, pp. 1–440 (2021)
34. Silva, A., Emmanuel, G., McClain, J.T., Matzen, L., Forsythe, C.: Measuring expert and novice performance within computer security incident response teams. In: Foundations of Augmented Cognition: 9th International Conference, AC 2015, Held as Part of HCI International 2015, Los Angeles, CA, USA, 2–7 August 2015, Proceedings 9, pp. 144–152. Springer (2015)
35. Svacina, J., et al.: On vulnerability and security log analysis: a systematic literature review on recent trends. In: Proceedings of the International Conference on Research in Adaptive and Convergent Systems, pp. 175–180 (2020)
36. Scheponik, T., et al.: How students reason about cybersecurity concepts. In: 2016 IEEE Frontiers in Education Conference (FIE), pp. 1–5. IEEE (2016)
37. Yin, C., Minatoya, D., Zhao, F.: Factor analysis for performance prediction using e-book learning logs. In: 2023 International Conference on Artificial Intelligence and Education (ICAIE), pp. 42–43 (2023). https://doi.org/10.1109/ICAIE56796.2023.00021