

Problem-Based Cybersecurity Lab with Knowledge Graph as Guidance

Yuli Deng, Zhen Zeng, Kritshekhar Jha, and Dijiang Huang
Arizona State University, USA

(Received 05 December 2021; Revised 14 December 2021; Accepted 16 December 2021; Published online 13 February 2022)

Abstract: Lecture-based teaching paired with laboratory-based exercises is mostly used in cybersecurity instruction. However, it focuses more on theories and models but fails to provide learners with practical problem-solving skills and opportunities to explore real-world cybersecurity challenges. Problem-based learning (PBL) has been identified as an efficient pedagogy for many disciplines, especially engineering education. It provides learners with real-world complex problem scenarios, which encourages learners to collaborate with classmates, ask questions, and develop a deeper understanding of the concepts while solving real-world cybersecurity problems. This article describes the application of the PBL methodology to enhance professional training-based cybersecurity education. The authors developed an online laboratory environment to apply PBL with Knowledge Graph (KG)-based guidance for hands-on labs in cybersecurity training. Learners are provided access to a virtual lab environment with KG guidance to simulated real-life cybersecurity scenarios. Thus, they are forced to think independently and apply their knowledge to create cyber attacks and defend approaches to solve problems provided to them in each lab. Our experimental study shows that learners tend to gain more enhanced learning outcomes by leveraging PBL with KG guidance, become more aware of cybersecurity and relevant concepts, and express interest in keep learning of cybersecurity using our system.

Key words: laboratory; knowledge graph; cybersecurity

I. INTRODUCTION

Existing cybersecurity instruction materials are mainly managed in a traditional lecture-centric fashion [16], in which instructors arrange learning materials and corresponding lab content based on the list of selected topics. However, the interlab dependencies are usually complicated and unclear, hindering both learners and instructors from managing learning and teaching materials coherently. As a result, most of the available cybersecurity hands-on labs fall short of learner needs. These labs typically follow a step-by-step approach and consist of series of tasks presented to learners to work on. Still, they are not designed to build problem identification and solving skills and fall short of helping learners to gain a deep understanding of cybersecurity concepts. A few critical issues that need to be addressed are as follows [6,15]:

- Current hands-on lab projects do not provide sufficient PBL design and problem-solving experiences to students.
- Students lack the skills and experience due to lack of practice, and hands-on lab projects need to incorporate more opportunities to develop these skills.
- Hands-on lab projects need to introduce more awareness amongst students of cybersecurity practice and its impact.
- The existing hands-on lab projects strategies in cybersecurity need to become more student-centered.

To meet these challenges, this study aims to apply PBL [23] on cyber security hands-on lab projects. In PBL, instructor acts as a facilitator and a mentor rather than the source of the solution and presents the students with a problem instead of lectures and

assignments. As the students are not handed any content, the learning becomes more active. It encourages students to explore and work with the specific contents identified as important by the instructor to find a solution to the problem.

As the hands-on lab is the most critical learning approach for cyber security education [26], we focus on laboratory experience. To apply PBL to cybersecurity labs, we developed a cloud-based hands-on lab environment by rearranging lab materials in a problem-centric fashion. All lab materials in the system, including hands-on lab instructions and related documents, are processed, and a list of problems is identified as goals for each lab.

We then construct Knowledge Graph (KG) [18] for each lab based on these problems identified. Nodes in these KGs are cybersecurity-related concepts and domain-specific knowledge mentioned or required in each lab. These concepts and knowledge are obtained by applying a data mining algorithm when processing the lab materials. The KG also provides rich learning materials for each node, including definition, tutorial, example applications, and other external resources (Wikipedia Link, YouTube videos, etc.). Accessible to all learners in a Web UI, the constructed KG is utilized as guidance during the PBL process during each lab, which will significantly enhance learners' problem-solving experience. To better understand how our PBL lab with KG guidance approach affects the learning process of learners, we conducted a qualitative analysis of our lab system by interview users of the system. We looked at multiple factors during the interviews, including users' problem-solving behaviors, their motivations for doing the labs, how they use the KG, and their willingness to continue their learning using the system. By interviewing learners who used our PBL lab system with KG guidance during a cybersecurity professional developing event and analyzing their training performance, we found that learners tend to gain better learning

Corresponding author: Yuli Deng (email: yuli.deng@asu.edu).

outcomes, become more aware of cybersecurity, and express more interest in the cybersecurity area.

The remainder of this article is organized as follows. Section II briefly describes the background of Computer Science Education, Problem-based learning (PBL), the lab system, and KG. Section III explains the system architecture and the approaches used to construct the problem-based learning virtual lab environment with knowledge graph as guidance and how we emphasize it for cybersecurity education. Section IV reports the detail of our interview and the qualitative analysis design. Section V covers the qualitative analysis process and the analysis result. Finally, we provide a discussion and conclusion of the article in Sections VI.

II. RELATED WORK

A. COMPUTER SCIENCE EDUCATION RESEARCH

Computer science is a developing and diverse academic discipline following the rapidly changing of the technology itself. A large amount of effort and researching of issues concerning the teaching and learning of computer science are even more diverse. A lot of research is published by instructors in computer science education who shared their own experiences of teaching a certain course or using a certain tool. Many of them are facing challenges when developing a new course or teaching a new technology, they then implemented some innovation in the instruction process of the course. The effect of the innovation is then evaluated in a publication based on students' learning outcomes and feedback. For example, instructors created VIPLE [27], an opensource visual programming language for robotics courses, in an effort to integrate design process, workflow, programming concepts, control flow, parallel computing, event-driven programming into curriculum. New forms of instruction are also being introduced to enhancing student learning experience. Online resources like Wiki Web sites are being used during classes for collaboration and for publication of course assignments to introduces new ways of learning, improves collaboration, improves learning results, and increases self-directed learning skills. Advance technologies are also being brought into the classroom to improve learning experience [28]. Mobile phones with combinations of face detection and recognition algorithms have been used to record student activities and improve learning performance [29].

B. PROBLEM-BASED LEARNING

PBL was pioneered by Barrows and originally used for medical education [3]. Over the years, the model has been adopted to teach concepts in other disciplines like architecture, law, and business management [8,19,21]. PBL has also been identified as an efficient pedagogy for engineering education, where engineering professions constantly deal with uncertainty, and with incomplete data and competing demands from clients, governments, environmental groups, and the general public [15]. Studies show that PBL compensates learners with the engineering knowledge and skills they obtained despite the greater amount of work. Despite the greater amount of work, it was revealed that students learning with the PBL approach not only were benefited in the content area but also in generic skills such as leadership, analytical thinking, conflict management, and decision-making [17]. Another study also shows that PBL is a promising approach in limited timed training, where a short time PBL training can be very effective when associating with technology projects [24].

C. HANDS-ON LAB ENVIRONMENT

Our PBL lab system with KG guidance is built on a cloud-based cybersecurity hands-on lab environment [Reference is suppressed for review]. The goal of the lab environment was to allow students to access a remote and geographically distributed didactic lab environment with maximal flexibility. It provides a contained experimental environment for hands-on experiments using cloud-based virtualization technologies. The system can be securely accessed remotely through an interactive Web GUI by both instructors and students. This system enables learners to have full control of a lab environment. It enables them to conduct any lab tasks within the environment by accessing it through a pure Web GUI without installing any client or plugin.

D. KNOWLEDGE GRAPH

KGs play an important role in many applications such as question answering and recommendation because of their rich structural information. In the education domain, KGs are often used for subject teaching and learning and are integrated into intelligent tutoring systems [18]. To construct knowledge graphs, text mining and relation extraction are used as common approaches [13,22], but these approaches all have certain limitations [4].

III. DESIGN AND DEVELOPMENT OF THE PBL LAB SYSTEM WITH KG GUIDANCE

To enable PBL in hands-on lab projects, we transform hands-on labs into a set of problems that need to be solved in each laboratory session. These problems are gathered and managed in the KG that we use the cutting-edge natural language processing approaches [2,25] to construct for each lab. This section described how a KG is constructed for each lab and how it is applied to guide learners to enable PBL in hands-on lab projects.

The process includes 1) constructing a knowledge base by gathering knowledge in public domain and analyse the relation between different knowledge, 2) problem identification and problem statement generation from lab material, and 3) KG construction.

A. KNOWLEDGE BASE CONSTRUCTION AND RELATIONSHIP ANALYSIS

To construct the KG, a knowledge base is created by gathering and processing publicly available knowledge in the cybersecurity domain. We have designed a multiagent-based crawler toolkit developed in Python, as shown in Figure 1. The system uses Wikipedia articles and YouTube videos in the cybersecurity domain as knowledge sources in this study. Multiple agents are used to crawl the keyword glossary. All crawled data are eventually sent to a centralized server node which merge the entire data set from multiple worker nodes. Each worker instance dumps the data in a semistructured schema which is later used to create the KG.

In order to analyse the relationship between knowledge in the knowledge base, word embedding is applied to represent the concepts as vectors in low-dimensional space which is further used to detect similarities across multiple concepts. Studies on embedding with transformers have been popular in recent times. These studies leverage pretrained contextual language models, i.e., BERT [7], XLNET [25], etc., which are grounded on the transformer architecture. These solutions achieve compelling

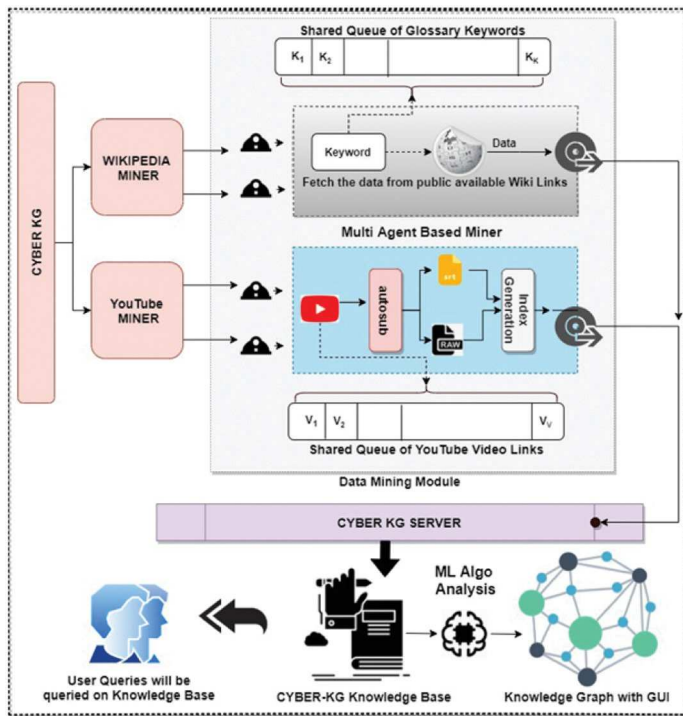


Fig. 1. System architecture.

performance on a wide variety of word embedding tasks [2]. Our solution modified the Xlnet word embedding model to derive semantically meaningful embedding from short phrases, instead of single words, as contextual phrases are a must to understand most cybersecurity concepts. It uses a cross-encoder procedure: two concepts are passed to the transformer network, and the target similarity value is predicted which is used to construct the Knowledge graph.

B. PROBLEMS IDENTIFICATION IN LABS

Within our system, we create a cybersecurity lab repository available to instructors and students in our university. We create

more than 50 cyber- security lab content based on computer network security courses in our school and used publicly available lab repositories, such as SEED lab [9]. Instructors are also able to upload their own new lab material into the lab repository at any time. To adopt these lab materials to PBL, we need to identify the cybersecurity problems implied in each lab. Thus, all lab materials are processed by instructors and broken down into a list of lab tasks. For example, "Packet Filter Firewall Lab" contains these tasks: 1. Prepare the lab environment. 2. Set up a network connection in the lab environment. 3. Install network software and service. and 4. Setup a stateless packet filter firewall.

We then transform each of these tasks into a problem statement, as shown in the second level of the knowledge in Figure 2. As these problem statements refer to real-world cybersecurity problems, our labs provide learners the experience of solving and analyzing real-world cybersecurity problems as PBL required.

C. CONSTRUCTION OF KNOWLEDGE GRAPH

To construct a knowledge graph for labs in our lab repository, each lab is tagged with key concepts by matching the lab material with concepts we identified in Section III.A. After gathering these concepts, we can calculate the similarities between concepts and lab problems we identified in section III.B using the word embedding similarity calculation model from Section III.A. After successfully mining YouTube video and cybersecurity glossary keywords from Wikipedia, we are then able to generate a knowledge graph for each lab in our system. One example Knowledge Graph is shown in Figure 2. The root node of the KG is the lab's name, second-level nodes are the problem statements identified in the lab based on lab tasks and requirements, and leaf nodes are cybersecurity concepts related to each problem statement node in the second level. This knowledge graph is then provided to learners during training in web GUI. Learners can utilize the KG as guidance to solve the problems by studying and researching the cybersecurity concepts in the graph.

D. KNOWLEDGE GRAPH AS PBL GUIDANCE

With the Knowledge Graphs stored in a graph data structure, we visualize the graphs in an interactive GUI to guide learners. In the interactive KG UI, learners are first presented with the problems

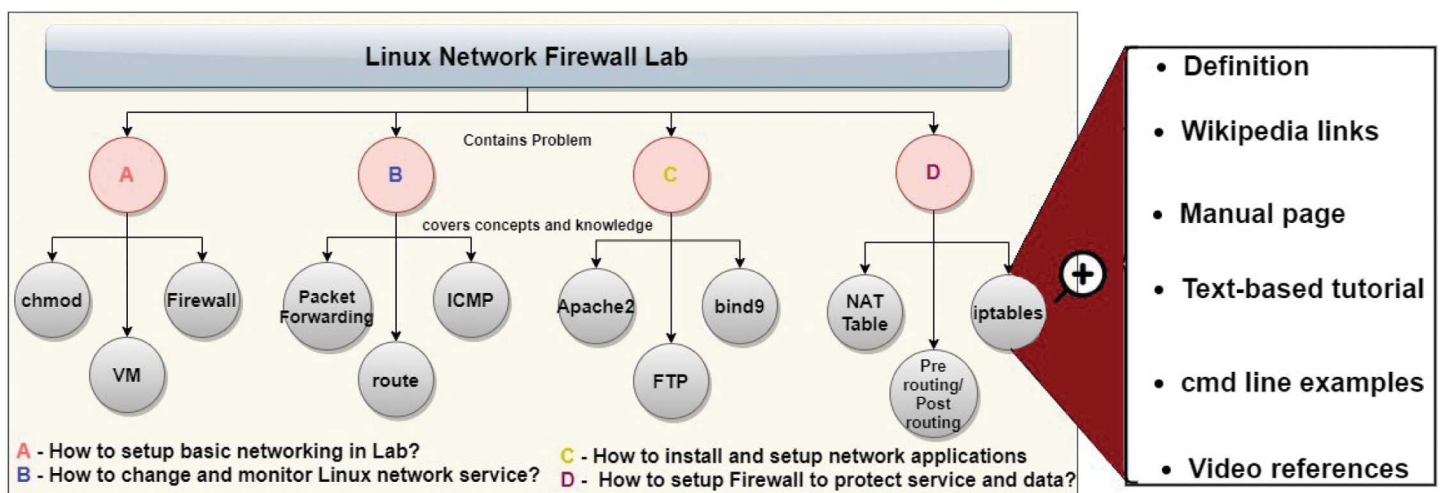


Fig. 2. Knowledge Graph of Linux Network Firewall Lab.

identified in Section III.B. These problems are the core of the PBL process, and they transform the lab system into the facilitator role in the PBL model. Since all the problem statements in each KG are identified and generated following the inner logic of the corresponding lab, KG is able to preserve the correlation and dependency between these problems. When learners follow the KG guidance to solve the problems, they are able to trace the relationship between the problems and gain cybersecurity knowledge and skills in an organized manner and better understand the correlation between different skills. As the facilitator of PBL, KG is responsible for supporting, guiding, and monitoring each learner's learning process. The KG can achieve these tasks by providing cybersecurity concepts directly related to each problem based on Section III.C result. In a KG, all related concepts are present to learners in one single place in a consolidated fashion. From the most basic concepts like IP address and user privilege to advanced knowledge like network virtualization and advanced persistent threat, Learners can quickly explore all concepts required to solve a particular problem conveniently. Learners are able to have a clear overall view of the knowledge required during the problem-solving process. To further reduce the time learners will otherwise spend investigating what they need to know and how and where to access required information, our KG UI also provides them with rich learning materials for each concept. These materials including the definition of the specific concept, tutorial/manual, example/application of the concept in the cybersecurity domain, and other resources available, like related Labs in our system, Wikipedia link, and even YouTube videos that explain the concept, as shown in the right part of Figure 2. YouTube video is to be highlighted here, as appropriate, vetted, and factually right video content can facilitate deep thinking in students to a great extent. These rich materials help learners to be more confident and more focused on the problem-solving process, key concepts, and their reflection of the problem, all three are critical elements in the PBL model. As a result, the guidance provided by KG can enhance learners' content knowledge while simultaneously allowing the development of their problem-solving, critical thinking, better understanding, and retention of knowledge and self-directed learning skills.

IV. EXPERIMENT DESIGN AND ANALYSIS

A. EXPERIMENT DESIGN

The goal of PBL lab system with KG guidance is to create a meaningful learning guide to help students address practical security issues and accomplish lab tasks. This system provides an effective platform to educate students on a comprehensive set of concepts and criteria that provide a tangible problem-based learning experience. To this end, we investigate the following research questions (RQs) for the KG-based hands-on lab system:

- RQ1: Why do professional trainees use the PBL lab system with KG guidance? (Motivation)
- RQ2: How do professional trainees solve problems in the PBL lab system with KG guidance? (Problem based learning experience)
- RQ3: Do professional trainees keep using this PBL lab system with KG guidance for professional development in the future? (User satisfaction)

We design a qualitative study that focused on investigating the re- search questions in this study. The design of semi-structured interview questionnaires in Table 1 is used to investigate

Table 1. The pre-selected interview questions in this study

Interview Questions

1. What is a specific purpose that you find yourself for using this lab system?
2. How seriously did you take the lab projects?
3. Did you learn something new from these hands-on lab projects?
4. Could you please briefly describe what you did by using this system?
5. Could you please give an example that you think the function of this system supports your learning during the lab project?
6. Do you think this knowledge graph covers the needed cybersecurity knowledge and concepts for the lab project?
7. Do you think the practice on hands-on labs helps and supports you to solve practical problems in real life?
8. Do you think the practice on hands-on lab projects helps/supports you in learning about cybersecurity? (Give detailed example)
9. Do you think the practice of hands-on lab projects improves your awareness of cybersecurity?
10. How difficult were these lab projects compared to the projects you did for the same topic?
11. What's the confident level of learning that you target?
12. What factors affected your choice of confidence levels?

professional trainees' behavior and perception on using the PBL lab system with KG guidance, which is based on the existing qualitative study on investigating students' behavior and perception on using a designed education system for computer science education [20]. We highlight participants' feedback on using the KG-based lab system for solving a specific problem. Additionally, we collect participants' background information at the beginning of the interview. Such background information contains the participants' education/professional experience in cybersecurity.

B. DATA COLLECTION

The experiment is situated in a week-long professional development event aiming to introduce researchers, educators, and other working professionals with content and domain knowledge of cybersecurity and help them develop problem-solving skills for cybersecurity. During the event, participants were asked to finish 2 to 4 hands-on labs recommended by the system. We recruited 9 of the participants from those who had completed the professional development event as interviewees as shown in Table 2, where three interviewers (around 33%) are female professionals, three interviewers work in an industry with a graduate degree in Cybersecurity. Our interviewees were a random sample, and we interviewed every trainee who expressed their interest in this research and included all interviews in the analysis.

Table 2. Participants Background

Participants Background		Number	Male/ Female
Academia	Master Students (A1, A2)	2	2/0
	PhD Students (A3-A6, A-I-1)	5	2/3
Industry	Junior Pro (I1)	1	1/0
	Senior Pro (I2, A-I-1)	2	1/1

* One female (A-I-1) Ph.D. student works also as a full-time senior professional in industry; thus, the total participants are nine.

The interview was semi-structured with some pre-selected questions listed in Table 1. During the interview, all interviewers are flexible to dive deeper into any questions. They are encouraged to provide further context or relevant information when answering the listed questions. The interview started with a demo. Each trainee shared their computer screen and showed a list of lab output/results for a specific lab they finished during the professional development event.

During the demo, the trainees were also asked to explain the results in their own words so that interviewers could better understand their thinking process. The demo was included to observe student behaviors when using our PBL lab system live, which would offer another layer of analysis and rely on their answers to our follow-up questions, especially to analyze problem-solving behaviors during the lab process. After the demo, the interviewers first asked all the trainees if they have any questions or recommendations about the PBL lab system with KG guidance during their training. After that, the interviewer went over a list of pre-selected questions about the student's experience with the PBL Lab system with KG guidance. These questions were related to the research question we aimed to investigate, as detailed in Section IV.A. The pre-selected questions used in the interview are listed in Table 1. Each interview takes approximately 15 mins, and all interview sessions are recorded for analysis later. Among the trainees, there are 3 female participants and 6 male participants, where 2 of them is a master student pursuing a master's degree in computer science while the rest 7 of them are with a graduate degree in CS. There are 3 of the trainees working as professionals in the IT industry, while the rest 6 trainees are still studying and doing research at their university.

C. ANALYSIS METHODOLOGIES

To analyze the collected data, we coded the transcripts of recorded interview audio files. We follow the open coding approach that used in the previous study [12,20] to analyze the collected data. The open coding method is the analytic process that attaching the concepts (codes) to the observed data and phenomenon in qualitative data analysis [12]. In the study, two authors coded an interview transcript independently and then discussed them together under several rounds to ensure the standardized coding framework. Then, the coders coded the collected data independently and interpreted the participant's responses by considering the semantic information of the entire interview. After coding all transcripts, the participants' comments are extracted to address the research questions. Our analysis results are based on 10 codes representing participants' feedback on using the PBL lab system with KG guidance. Table 3 shows the list of codes for each research question category. We explore the participants' learning behavior using the PBL lab system with KG guidance by analyzing the relevant codes from the participants' interview transcripts.

V. EXPERIMENT RESULTS AND ANALYSIS

This section presents the findings for the research questions by analyzing the collected data. The findings of how professional trainees use the PBL lab system with KG guidance are summarized from 3 perspectives (motivation, problem-based learning experience, and user satisfaction), which address the proposed research questions RQ1–RQ3 in Section IV.A.

A. MOTIVATION OF USING PBL LAB SYSTEM WITH KG GUIDANCE

For the motivation part, participants provide the purpose of using this PBL lab system with KG guidance that mainly focuses on creating a cybersecurity simulation environment, completing the group project practice, and supporting the learning experience. For instance, I1 said that “**it's convenient** for me to deploy any **virtual infrastructure** remotely,” and the senior professional in industry I2 stated that “(It is) a platform or tool which can quickly deploy, simulate, and verify my network and security architecture design easily.” A6 also highlights the purpose of using this system “**to practice real-world problems that are related to cybersecurity and network security.**” Differently, the A2 and A3 are motivated by the functions that support group projects. A2 said that “a passion that I also use those machines and in some of my research for work, where we need to work together as a group.” Additionally, A5 and A1 emphasized the usefulness of the KG function as “the Knowledge graph is a neat add-on to the learning experience. The mapping on related concepts and useful information at each node really helps to frame an overall concept map **better**”, and “the knowledge graph is also helpful... and provides a lot of resources for me to master this knowledge.”

Take-away for RQ1: The interview results suggest the similarity and difference between the professional participants from industry and academia on the motivation of using the PBL lab system with KG guidance. All participants expressed their perception of this system that is easy to use and useful. According to the theory of IT acceptance [1], the perceived usefulness and ease of use are the key factors that are positively associated with the usage of a new IT system. Thus, this PBL lab system with KG guidance might be accepted well by a broader range of users based on the findings in this qualitative study. Additionally, we noticed the different needs of using this system based on the role of participants, where the participants from academia emphasized the need of supporting collaborated work that is usually for a course study purpose, and the participants from industry focused on the functions that can support deploy a project. There is not a significant difference among gender in the motivation of using this system. These findings will guide our future development on this PBL lab system with KG guidance to better address the needs of professional trainees.

B. PROBLEM-BASED LEARNING EXPERIENCE

The participants in academia and industry all showed their behavior on exploring the new knowledge and skillset in the system, which are guided by the KG function. For example, I1 said that “knowledge graph provides all this information in a consolidated fashion in one single place.” I1 also explained that “the knowledge graph basically helps students understand that correlation between different skills ... I can basically trace those dependencies and learn these in an organized manner.” Regarding how to solve a problem under the KG guidance, we found that the KG function gradually guides the trainee exploring from the basic level to the advanced level of knowledge to solve a problem. The representative quotes from A3 are: “I hardly understood the Linux environment and how to do the setup ... it (the KG) shows that what commands you should use for setting up a firewall, what does a firewall means, and it also gives a description. And then, it gives a link where you can go and watch, ... and then all the related ... when the tasks were given ... it was super easy because you know that knowledge

graph actually have been traveled from the basic level to the further advanced level ... and gradually improve my knowledge in that area.”

Regarding their perceptions on the confidence level of solving a problem, most participants felt the lab practice in this study was at a moderate level and felt confident on solving a problem in this case study, where the KG guidance efficiently supports participants in learning cybersecurity knowledge and skill. The key factors that affected the choice of confidence levels were, as stated by A5, “the visualization of the related terms across concepts helps the user in identifying new concepts to learn ... KG brings all the relevant information in one place for students to enable effective learning.” Only one participant (A2) reported that he might not be so confident at the beginning, and A2 said that “I ended up retaining and learning more than I thought I was, so it was very beneficial.” A3 explained how to build up the confidence gradually under the KG guidance as: “before I started it, I wasn’t that much sure, like say 20% or so, but after reading the material once ... it was boosted my confidence to around 50–60%, but when I actually did it ... go back to the material again ... followed each step carefully, I think I could do most of the part.”

In addition, all participants agreed that their awareness of cybersecurity was increased. Specifically, the industry participants suggested that an advanced level of KG function that including attack scenarios could better support the needs of industrial practice.

Take-away for RQ2: The interview results reveal that the KG function facilitates exploring and obtaining knowledge, organizes the correlated skills for tasks, and effectively supports learning and problem-solving processes. Based on the results, the design of KG function well organized the knowledge/skills covered in a task into a big picture, for example, as shown in Figure 2, which prevented the trainee from being trapped in microscopic perspectives with an isolated problem/knowledge units and guided the trainee to deliver a comprehensive big picture for the target task. The existing study on cybersecurity education [5] also proves that the KG’s multi-layer, multi-dependencies design can help to build a knowledge network instead of isolated knowledge units. The participants suggest an advanced knowledge graph, which covers more cybersecurity concepts, dependencies, and cybersecurity attack scenarios.

In addition, we noticed that the behavior (as stated by A3) that mapping the cybersecurity knowledge learning with the hands-on problem learning contributes to increasing the trainee’s confidence level. According to the existing study, the decrease rates of newly acquired knowledge are lowered down by consolidating hands-on learning with cognitive learning [10]. Thus, this PBL lab system with KG guidance might have an advantage in supporting trainees’ learning success in the long term, which can be possible future work. All the findings from the interview bring up our future work of a more accurate way of gauging the trainee’s PBL in this system.

C. USER SATISFACTION

All participants preferred using the PBL lab system with KG guidance for other projects in the future. The participants from the industry emphasized the supports for solving real-life problems. For instance, I2 said that “the labs allow me to easily design and deploy a small to mid-scale network security architecture. I can use the deployment to verify my design for feasibility and give it a small-scale performance test.” Additionally, I1 said that this PBL lab system with KG guidance “definitely helps understand the

different ways an attacker can enter your network, and what are the actions you can take to defend against those kinds of attacks, ... so it definitely helps in improving the security of the data as well as infrastructure in my day-to-day job setting.” The participants from academia underlined the need to learn cybersecurity knowledge and resource since such supports could help them better prepare cybersecurity skills for the job market as A4 said that “as the Cybersecurity concepts are understood and visualized better with solving practical problems around the concepts. This also prepares students for better practical engineering jobs associated with graduation.

Take-away for RQ3: This part of the interview focuses on investing in the desire for system usage. The participants in this study showed differences in the interested cybersecurity tasks. The findings reveal that the trainees perceived a good performance of tasks by using this PBL lab system with KG guidance. According to the existing study, a fit between technologies and users’ tasks can enhance the task performance [11], and then motivates to use the system for learning [14]. These findings well guide our future development of this system on generating a personalized knowledge graph that better supports the individual trainee’s task.

VI. CONCLUSION AND FUTURE WORK

This study applied Problem-based learning in a cybersecurity lab environment and created a knowledge graph as PBL guidance for learners. We observed each trainee’s problem-solving process in the PBL lab system with KG guidance and studied the similarity and difference of motivations between participants from industry and academia background. We also explored how the functional design of KG facilitates the knowledge acquisition process and enhances the trainee’s confidence level by consolidating hands-on lab-based learning with cognitive learning (concepts/knowledge in KG). All participants shared the eagerness to continue training in cybersecurity and were all interested in using our PBL lab system with KG guidance for future training. The findings in this study also identified the urgency of developing a more advanced and complete cybersecurity knowledge base that covers most cybersecurity concepts and training scenarios. A personalized knowledge graph for an individual trainee is also required to gauge the problem-based learning experience in this system more accurately. Lastly, since this study only samples on a small group of participants in a professional training event, the study base is very limited. In the future, to cover more participants, especially university students, we want to carry out more studies in a university classroom environment to further investigate the effect of PBL in the cybersecurity education and keep improving our PBL lab system with KG guidance.

REFERENCES

- [1] D. A. Adams, R. R. Nelson, and P. A. Todd, 1992. “Perceived usefulness, ease of use, and usage of information technology: a replication,” *MIS Q.*, vol. 16, pp. 227–247, 1992.
- [2] E. Alsentzer, J. R. Murphy, W. Boag, W.-H. Weng, D. Jin, T. Naumann, and M. McDermott, 2019. Publicly available clinical BERT embeddings. arXiv preprint arXiv:1904.03323 (2019).
- [3] H. S Barrows, R. M Tamblyn, et al., *Problem-Based Learning: An Approach to Medical Education*, Vol. 1. New York, NY, USA: Springer Publishing Company, 1980.

- [4] P. Chen, Y. Lu, V.W. Zheng, X. Chen, and B. Yang, "Knowedu: A system to construct knowledge graph for education," *IEEE Access*, vol. 6, no. 2018, pp. 31553–31563, 2018.
- [5] J. Dai, 2018. "Situation awareness-oriented cybersecurity education," in 2018 IEEE Frontiers in Education Conference (FIE), San Jose, CA, USA: IEEE, pp. 1–8.
- [6] P. Deshpande, C.B. Lee, and I. Ahmed, 2019. "Evaluation of peer instruction for cybersecurity education," in *Proc. 50th ACM Techn. Symp. Comput. Sci. Educ.*, pp. 720–725.
- [7] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805* (2018).
- [8] E. Driessen and C. Van Der Vleuten, "Matching student assessment to problem-based learning: lessons from experience in a law faculty," *Stud. Contin. Educ.*, vol. 22, no. 2, pp. 235–248, 2000.
- [9] W. Du, "SEED: hands-on lab exercises for computer security education," *IEEE Secur. Priv.*, vol. 9, no. 5, pp. 70–73, 2011.
- [10] S. Gerstner and F. X. Bogner, "Cognitive achievement and motivation in hands-on and teacher-centred science classes: Does an additional hands-on consolidation phase (concept mapping) optimise cognitive learning at workstations?," *Int. J. Sci. Educ.*, vol. 32, no. 7, pp. 849–870, 2010.
- [11] D. L. Goodhue and R. L. Thompson, "Task-technology fit and individual performance," *MIS Q.*, vol. 19, pp. 213–236, 1995.
- [12] P. Kinnunen and B. Simon, 2010. "Building theory about computing education phenomena: A discussion of grounded theory," in 10th Koli Calling International Conference on Computing Education Research., Held in Koli, Finland, between October 28–31, 2010. pp. 37–42.
- [13] H. Liu, W. Ma, Y. Yang, and J. Carbonell, "Learning concept graphs from online educational data," *J. Artif. Intell. Res.*, vol. 55, pp. 1059–1090, 2016.
- [14] T. J. McGill and J. E. Klobas, "A task–technology fit view of learning management system impact," *Comput. Educ.*, vol. 52, no. 2, pp. 496–508, 2009.
- [15] J. E. Mills, D. F. Treagust, et al., "Engineering education—Is problem-based or project-based learning the answer," *Australas. J. Eng. Educ.*, vol. 3, no. 2, pp. 2–16, 2003.
- [16] D. Mouheb, S. Abbas, and M. Merabti, "Cybersecurity curriculum design: A survey," in *Transactions on Edutainment XV*. Z. Pan, A.D. Cheok, W. Müller, M. Zhang, A.E. Rhalibi, K. Kifayat. Springer, 2019, pp. 93–107. <https://link.springer.com/book/10.1007/978-3-662-59351-6#editorsandaffiliations>.
- [17] J. C. Perrenet, P. A. J. Bouhuijs, and J. G. M. M. Smits, "The suitability of problem-based learning for engineering education: Theory and practice," *Teach. High. Educ.*, vol. 5, no. 3, pp. 345–358, 2000.
- [18] M. Rizun, et al., "Knowledge graph application in education: A literature review," *Acta Univ. Lodziensis. Folia Oecon.*, vol. 3, no. 342, pp. 7–19, 2019.
- [19] H. G. Schmidt, S. M. M. Loyens, T. Van Gog, and F. Paas, "Problem-based learning is compatible with human cognitive architecture: Commentary on Kirschner, Sweller, and Clark," *Educ. Psychol.*, vol. 42, no. 2, pp. 91–97, 2007.
- [20] K. Shah, P. Lee, D. Barretto, and S. N. Liao, "A qualitative study on how students interact with quizzes and estimate confidence on their answers," *Proc. 26th ACM Conf. Innov. Technol. Comput. Sci. Educ.*, vol. 1, pp. 32–38, 2021.
- [21] J. E. Stinson and R. G. Milter, "Problem-based learning in business education: Curriculum design and implementation issues," *New Direct. Teach. Learn.*, vol. 68, pp. 33–42, 1996.
- [22] S. Wang, C. Liang, Z. Wu, K. Williams, B. Pursel, B. Brautigam, S. Saul, H. Williams, K. Bowen, and C. Lee Giles, 2015. "Concept hierarchy extraction from textbooks," *ACM Symposium on Document Engineering 2015*, held in Lausanne, Switzerland, between September 8–11, 2015, pp. 147–156. doi:10.1145/2682571.
- [23] D. F. Wood, "Problem based learning," *BMJ*, vol. 326, no. 7384, pp. 328–330, 2003.
- [24] A. Yadav, D. Subedi, M. A. Lundeberg, and C. F. Bunting, "Problem-based learning: Influence on students' learning in an electrical engineering course," *J. Eng. Educ.*, vol. 100, no. 2, pp. 253–280, 2011.
- [25] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. R. Salakhutdinov, and Q. V. Le, "Xlnet: Generalized autoregressive pretraining for language understanding," *Adv. Neural Inf. Process. Syst.*, vol. 32, pp. 5754–5764, 2019.
- [26] Z. Zeng, Y. Deng, I. Hsiao, D. Huang, and C.-J. Chung, 2018. "Improving student learning performance in a virtual hands-on lab system in cybersecurity education," in 2018 IEEE Front. Educ. Conf. (FIE), IEEE, San Jose, CA, USA, pp. 1–5.
- [27] Y. Chen and G. De Luca, 2016. "VIPLE: visual IoT/robotics programming language environment for computer science education," in 2016 IEEE Int. Parallel Distrib. Process. Symp. Workshops (IPDPSW), IEEE, Chicago, IL, USA, pp. 963–971.
- [28] W. T. Tsai, W. Li, J. Elston, and Y. Chen, "Collaborative learning using wiki web sites for computer science undergraduate education: A case study," *IEEE Trans. Educ.*, vol. 54, no. 1, pp. 114–124, 2010.
- [29] L. Cronjé and I. Sanders, "Semiautomated class attendance monitoring using smartphone technology," *J. Artif. Intell. Technol.*, vol. 1, no. 1, pp. 9–20, 2021.