Design and Use of Privacy Capture-the-Flag Challenges in an Introductory Class on Information Privacy and Security

Wolfgang Vigl
wolfgang.vigl@uibk.ac.at
Department of Computer Science
University of Innsbruck
Innsbruck, Austria

Svetlana Abramova
svetlana.abramova@uibk.ac.at
Department of Computer Science
University of Innsbruck
Innsbruck, Austria

ABSTRACT

With the advancement of data-intensive technologies and online tracking opportunities, education on privacy is becoming a bigger priority in cybersecurity curricula. Owing to its multidimensional and context-dependent nature, privacy and its protection goals can be introduced to computer science students as collective efforts of system designers to enable data protection, user's informational control, and privacy-preserving data sharing with third parties. Since engaging course participants into hands-on and gamified exercises is generally known to enhance learning experience and effect, we adopted the teaching practice of using Capture-the-Flag (CTF) security challenges and validated its applicability in the privacy education domain. This work presents a pioneering set of 8 CTF-style tasks intentionally designed for the study and demonstration of handpicked privacy concepts, techniques, and attacks. Drawing upon both quantitative and qualitative feedback collected from 27 course participants, this format of homework exercises is found to increase students' confidence in this knowledge domain and perceived as an enjoyable, motivating, and engaging way of learning about information privacy.

CCS CONCEPTS

• **Applied computing** → **Interactive learning environments**; • **Security and privacy**; • **Social and professional topics** → **Student assessment**;

KEYWORDS

Privacy, Capture-The-Flag, CTF, Education, Cybersecurity, Student engagement

ACM Reference Format:
Wolfgang Vigl and Svetlana Abramova. 2024. Design and Use of Privacy Capture-the-Flag Challenges in an Introductory Class on Information Privacy and Security. In *Proceedings of the 2024 Innovation and Technology in Computer Science Education V. 1 (ITiCSE 2024), July 8–10, 2024, Milan, Italy.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3649217.3653572

1 INTRODUCTION

Historically, *information privacy* has received little attention in computer security curricula, mostly oriented towards teaching the fundamental protection goals of system security, common vulnerabilities, and technical countermeasures [3, 32]. However, with the ever-increasing volume of online sensitive data, the availability of user tracking capabilities, and the emergence of data protection laws (e. g., EU's General Data Protection Regulation, California Consumer Privacy Act, Brazil's General Data Protection Law), integrating privacy and its engineering paradigms within a broader scope of security courses has become imperative in higher education [7, 28]. In particular, trending discussions of the utility of data analytics need to be juxtaposed with considerations of privacy risks inherent to data collection, processing, and transfer [24]. Offering a broader and unbiased perspective on this fundamental dilemma will allow universities to prepare computer science graduates who can blend technical skills with critical reasoning about ethical, legal, and social implications of emerging and future technologies.

The conceptual interpretation of privacy has evolved over time, initially articulated by legal scholars as 'the right to be let alone' [35] and later re-framed as 'the right to informational self-determination', i. e., *"an entity's ability to control how, when, and to what extent personal information about it is communicated to others"* [36]. Computer scientists and security researchers think of privacy through the lens of adversarial models and protection means, which enable data confidentiality and anonymity, user's informational control and decision making, and transparency and accountability associated with data access and processing [31]. This reflects a so-called engineering perspective on privacy, advocating the idea of shifting both autonomy and responsibility for protection of personal data from users to system controllers and designers [29]. In light of this background, introducing privacy to computer science students as a knowledge area which drives the need for data protection can lead to new ways of designing, building, and operating services.

Teaching privacy in combination with security as part of one introductory course is, however, not a straightforward task, as there exist many and varied classes of privacy issues and technologies [31]. Typically, standalone courses on privacy cover such fundamental topics as regulatory frameworks, user preferences, data protection and anonymization, digital traces and online tracking, and anonymous communication systems. Despite this variety of topics and constraints of an introductory course, instructors should educate students to recognise privacy issues, to describe them using proficient terms, and to suggest appropriate prevention or mitigation strategies. These learning goals can be reached by complementing theoretical contents with engaging exercises and scenarios showcasing the real-world application of the taught concepts, potential privacy infringements, and countermeasures.

ITiCSE 2024, July 8–10, 2024, Milan, Italy

ACM ISBN 979-8-4007-0600-4/24/07.
https://doi.org/10.1145/3649217.3653572

One of the popular hands-on and engaging methods to teach cybersecurity skills is by using 'Capture-the-Flag' (CTF) challenges. In its common form, a CTF challenge simulates a real-world cyber attack against a purposefully vulnerable system, the successful exploitation of which reveals a secretly hidden text string known as a flag. Security CTF challenges often have an unspecified solution path, stimulating participants to experiment, conduct external research, and exhibit self-learning and ingenuity in a repeated, trial-and-error manner [21]. Prior works studying the effect of CTF challenges in security education report that this format of homework exercises is generally perceived by students as an engaging and enjoyable learning activity [6, 21, 23]. We were inspired by our own multi-year experience of using security CTF challenges as a supporting educational tool and implemented a series of privacy CTF-style exercises, which students needed to solve individually within a two-week time frame. This experiment aimed to showcase the practical relevance of privacy engineering as well as to stimulate students' interest, creativity, and enjoyment of learning about it. Furthermore, our goal was to transition from pen-and-paper homework assignments and manual grading to an automated learning system by making use of the existing CTF platform.

This paper presents a pioneering set of the privacy CTF challenges designed for first-year computer science and information systems Master students enrolled in an introductory course on information privacy and security. To the best of our knowledge, our initiative is the first attempt to ideate, design and validate simple, yet instructive CTF-based challenges which illustrate selected privacy concepts and re-identification attacks. In this paper, we share our proof-of-concept experience and lessons learned and contribute to computing education research by describing the designed 8 privacy CTF challenges in detail. We also hope that our work will encourage course instructors and CTF creators to elaborate and develop CTF-style assignments beyond typical security scenarios.

The remaining paper is organized as follows. In Section 2, we briefly review related work. Section 3 is devoted to the presentation of the developed privacy CTF challenges. We report the evaluation results provided by the students who participated in our educational experiment in Section 4. Finally, we discuss the key lessons learned and limitations of our approach in Section 5, and conclude.

2 RELATED WORK

In recent decades, there has been a shift towards the inclusion of privacy education into high-school and university curricula to address the growing risks from IT systems [11]. Egelman et al. [10] present the *Teaching Privacy Project* to educate high school and college students about online privacy, and support educators with extensive teaching materials and resources. Other publications (e. g., [29, 31, 32]) emphasize the responsibility of software designers to develop privacy-friendly systems, thereby protecting the interests of online users. Yet, there is still a lack of consensus in higher education regarding privacy curriculum frameworks for technical degrees.

Most of prior works on the use of CTFs in cybersecurity education are in the form of experience reports describing positive or neutral learning effects, potential pitfalls, lessons learned, and recommended practices [2, 4, 27]. More recent studies adopt scientific experimental designs and attempt to measure the impact of CTF hands-on exercises on student performance, engagement, and satisfaction through common statistical tests and using both subjective (e. g., survey responses) and objective (e. g., assessment scores) measures [6, 23, 33].

As supported by both quantitative and qualitative evidence, students often find the CTF-based form of learning as interactive, collaborative, motivating, and helpful in developing practical security skills and self-confidence [6, 21, 23]. Concerning the effectiveness of CTF-style assignments compared to traditional in-class or home exercises, little empirical data exists to draw conclusive statements. Cole's work [6] shows that solving CTF challenges increases student motivation without negative effects on learning outcomes. However, further empirical studies based on rigorous research methods are needed before these results can be generalized.

Since CTF systems have been in use in cybersecurity education for some time, Švábenský et al. [38] have analyzed the distribution and evolution of topics represented in the existing CTF challenges. The authors have found that technical knowledge areas (e. g., data, connection, or system security) prevail, whereas non-technical aspects, such as usable or human-centered security, privacy, ethics, and law, are usually neglected by CTF creators. This state of the affairs is probably due to the non-trivial complexity of designing jeopardy-style and fun-to-solve assignments for originally non-technical concepts, as this task by default requires a high level of creative thinking. Our work seeks to address this apparent gap by presenting a series of the validated privacy CTF challenges.

3 PRIVACY CTF CHALLENGES

This section outlines a total of 8 CTF challenges created for hands-on learning of privacy concepts, practices, and attacks. The challenges can be categorized into three thematic categories: *Informational Control*, *Online Tracking*, and *Privacy-Preserving Data Sharing*, and are described by the following characteristics:

- *Concept(s)*: privacy concepts or attacks being addressed.
- *Scenario*: the story, context and approach to get the flag.
- *Implementation*: relevant technical details including cheating prevention measures.

As illustrated in Figure 1, the CTF platform's dashboard displays a title, a teaser image, a short description (and solution hints, if necessary), a difficulty tag (*easy*, *medium*, or *hard*) and an addressed concept for each challenge. All challenges are run on the host platform as standalone dockerized services. Python is used for server-side logic and responses to HTTP requests, while frontends and client-side functionality are implemented in Angular[1].

3.1 Informational Control

The two challenges presented below are designed under the privacy as informational control paradigm and introduce a few techniques allowing users to express their privacy preferences on the web.

Alternative Signals for Privacy Preferences

[1]Source code files and instructions for setting up the CTF platform and challenges are stored in a GitHub repository accessible to course instructors and available to other instructors upon a written request.

Figure 1: Example of a CTF challenge

Concept(s): Privacy preference signals, i. e., digital representations of how users want their personal data to be processed [19].

Scenario: A custom website presented to students uses "*Do Not Track*" and "*Global Privacy Control*" signals as alternatives to GDPR consent dialogues shown to EU visitors. The student's task is to activate both signals via the respective HTTP headers and to opt out of tracking and sharing of personal data.

Cookie Consent Bypassing

Concept(s): Dark design patterns coercing end-users into desired behavior and actions that are often not in their best interest [34].

Scenario: Students engage in a cookie consent dialogue game featuring several yes/no choices and related dark patterns (e. g., obfuscating text descriptions, pre-ticked check boxes, highlighted options etc). Rejecting consent to all cookie options within 90 seconds reveals the flag.

3.2 Online Tracking

The following two challenges introduce students to browser fingerprinting, a common technique used by websites for online user tracking and targeted advertising.

Two-Factor-Authentication (2FA) Login

Concept(s): Browser fingerprinting, i. e., a collection of a user's device and browser attributes, which together generate a fingerprint and may uniquely identify a user's browser [22].

Scenario: A flag hidden in the admin account is protected by a login system that enforces 2FA, if the user's browser fingerprint differs from the expected one upon a successful password-based login. The admin password is easily guessable, and students are expected to bypass 2FA by debugging the expected fingerprint in a browser console and manually setting the required system and browser settings (e. g., a timezone, window size, and language).

Implementation: A manual spoofing of an HTTP request is intentionally prevented by sending a calculated hash value of the user's fingerprint components for comparison purposes to the server.

TruthTunnel

Concept(s): Fingerprint entropy, which is used to quantify the level of a fingerprint's identifying information [15].

Scenario: Students are asked to access a whistleblower portal in their browser. This portal is accessible only if the entropy of specific fingerprint components is below an individually defined threshold. Students are expected to change their browser fingerprints by adjusting the values of the indicated browser components with a high entropy to the most common values.

Implementation: The values of a user's fingerprint components are fetched using *fingerprintjs* library [2] (v2.1.4). Once sent to the server, component entropies are calculated using the *Cover your Tracks* [3] API offered by the *Electronic Frontier Foundation*. The API operational status is verified on a system's launch. API congestion is mitigated by caching entropy values in the server's database and the browser's local storage.

3.3 Privacy-Preserving Data Sharing

This thematic category of the CTF challenges is devoted to basic practices for privacy-preserving data disclosure and the demonstration of interactive and non-interactive data inference attacks.

Anonymity Amongst Offenders

Concept(s): k-anonymity, i. e., each record in a dataset cannot be differentiated from at least $(k-1)$ other records with respect to a set of certain quasi-identifiers [30]. l-diversity, i. e., each equivalence class of quasi-identifiers has l distinct values for a sensitive attribute [25].

Scenario: Five k-anonymous, synthetic datasets containing criminal records of thousands of individuals are published. Students are expected to calculate and submit k-values for each dataset in order to reveal the flag. The next challenge is unlocked on success, with the similar task applied to l-diversity.

Implementation: Unique data sets are generated for each user upon clicking on a download button. The risk of brute-force solving is reduced by using multiple datasets and asking students to submit a combination of k-values. With $k, l \in [2, 16]$ and a total of five datasets, there are 537,824 possible combination solutions.

Census and Statistics

Concept(s): Database reconstruction attack, i. e., reconstructing the underlying microdata that is consistent with published statistical tables (e. g., by creating mathematical constraints and solving them via SAT solvers) [13].

Scenario: Students are presented with aggregated data (counts, average and median ages) of different population subgroups of individuals surveyed in a census study. By defining mathematical constraints using Satisfiability Modulo Theory solvers, individual data records can be recovered, whose upload to the CTF platform in a predefined format reveals the flag.

Implementation: For each student, unique census data is generated and aggregated for the display. File uploads undergo a series of format, size, and header checks.

Public Health Dataset

[2] https://github.com/fingerprintjs/fingerprintjs
[3] https://coveryourtracks.eff.org

Concept(s): Record linkage attack, i. e., an attacker manages to de-anonymize a certain record (or an individual) in a supposedly anonymized dataset by leveraging auxiliary information from another database [26].

Scenario: A student downloads two synthetic datasets: the first one contains the ICD disease codes of thousands of individuals, and the second provides voter information including identifiers such as names. A record linkage attack using the quasi-identifiers shared by both datasets de-identifies a single person whose ICD code serves as a flag (see Figure 1).

Implementation: Two unique datasets are generated for each user.

StatDB

Concept(s): Statistical inference attack, i. e., leveraging aggregate data and statistics to derive sensitive data of individuals [20] through interactive queries to a database, and a countermeasure "Query-Set-Size Control" mechanism that restricts a query's result depending on the number of retrieved records [8].

Scenario: The student's task is to statistically derive the salary of an employee with a predefined ID using an SQL-based API that provides statistical insights into the staff salaries. Students need to navigate through three increasing levels of the query restrictions. Initially, the concept is familiarised by only blocking direct access to elements, necessitating groupings and aggregations. The next levels implement the concept of query set size restrictions: first, by only allowing queries with the result set having two or more records, and then by additionally requiring query set sizes less than $(N - 2)$ records, where N equals the total number of records.

Implementation: Restrictions are enforced using *sqlglot* library to parse and verify a user's database query. Pre-query-execution checks verify the proper usage and prevent direct access. Query-set size restrictions are enforced after the query has been executed.

4 EVALUATION RESULTS

Learning experience and effectiveness of using privacy CTF challenges for educational purposes were evaluated by 27 participants who have attended the university's course *"Introduction to Information Privacy and Security"* in the winter term 2023/2024. This course is recommended for first-semester students enrolled in the Master's programs in *Software Engineering, Computer Science* or *Information Systems*. In terms of prerequisites, course participants are expected to have prior coding experience and basic knowledge of computer architectures, database systems, networking, and high-school maths. Students were presented with the CTF platform in an information session at the start of the competition. After two weeks, all the challenges and solutions were discussed in an interactive class session moderated by the course instructors.

4.1 Survey Methodology

Two in-class surveys were administered to students in order to assess the overall learning experience and effectiveness of this educational experiment. The pre-challenge questionnaire elicited students' privacy concerns and self-evaluated confidence in their abilities to explain privacy concepts before accessing the CTF system. The post-challenge survey gathered students' feedback on the challenges themselves and their perceived difficulty, and measured subjective satisfaction, motivation, and learning experience.

Table 1 presents the respective items used in the post-challenge questionnaire. Following the work of Zhang et al. [37], we borrowed five attitudinal variables that can influence technology-supported learning: learning flexibility (*LF*; the extent of freedom in terms of time, pace, and location students have in completing assignments outside a classroom), learning climate (*LC*; the learning atmosphere created by an educational platform), motivation (*MOT*), satisfaction (*SAT*), and task-technology fit (*TTF*; the extent to which the platform's functionality matches user's task requirements and expectations). The original scales were adapted to our privacy-centered context and the use of a CTF platform. In addition, we developed two specific factors to evaluate learning effect and student's understanding of privacy in a personal (*LEP*) or corporate (*LEC*) context. The respective items are shown in Table 1. Participants' privacy concerns were measured by including the common Internet User's Information Privacy Concerns (IUIPC-8) scale [17] in both questionnaires. Self-reported confidence (*CON*) of students in relation to privacy concepts, attacks, and countermeasures was measured in both questionnaires by a three-item scale adapted from Leune et al. [23]. All items were answered on a 5-point Likert scale (1 = *strongly disagree*, 5 = *strongly agree*).

Both questionnaires were administered using LimeSurvey platform licensed by the participating institution. Out of 33 registered participants, 28 and 27 completed the pre- and post-challenge questionnaire, respectively. We were able to cross-link 18 responses from both questionnaires using an anonymous ID code generated individually by each student according to the given instructions. For the linked responses, a directed paired Wilcoxon signed-rank test, chosen for non-normal data distributions, assessed whether *CON* and *IUIPC* improved after participating in the experiment. A descriptive analysis of the responses was done for the constructs surveyed only in the post-challenge questionnaire.

Our study was approved by the institution's research ethics committee. At the beginning of both surveys, each participant was asked to provide explicit consent to participation as well as informed of their right to withdraw from the study at any time without consequences and with their data being deleted. In addition, survey responses were stored on a national server adhering to the local data protection regulations.

4.2 Survey Results

4.2.1 Privacy Concerns. Although the mean scores across all IUIPC items increased after participation with a medium effect size [5], the results are not statistically significant, as showed by the Wilcoxon signed-rank test. On average, students reported increased concerns about control of personal information (pre: $M = 3.78$, $SD = 0.77$; post: $M = 3.94$, $SD = 1.11$; $r = .29$, $p = 0.219$) and awareness of conditions and practices (pre: $M = 4.14$, $SD = 0.87$; post: $M = 4.33$, $SD = 0.857$; $r = .35$, $p = 0.138$). Furthermore, concerns about data collection relative to the benefits received rose on average (pre: $M = 3.56$, $SD = 0.85$; post: $M = 3.67$, $SD = 1.03$; $r = .31$, $p = 0.194$).

4.2.2 Confidence and Learning Success. In contrast to privacy concerns, participation in the CTF competition significantly increased students' reported self-confidence, as evidenced by the following measures: $M = 3.72$, $SD = 0.46$ in the post-challenge responses as compared to the levels reported before the experiment $M = 3.19$,

Table 1: Constructs and items used in the post-challenge questionnaire.

Scale Item	1[a]	2[a]	3[a]	4[a]	5[a]	Mean	SD
Confidence (CON) [b]							
I am confident in my abilities to explain online privacy concepts.	0	0	7	20	0	3.74	0.45
I am confident in my abilities to recognize online privacy risks.	0	1	9	16	1	3.63	0.63
I am confident in my abilities to protect my own online privacy.	0	3	13	6	5	3.48	0.94
Learning Climate (LC)							
I had fun solving the privacy CTF challenges.	0	1	5	9	12	4.19	0.88
I found the privacy CTF challenges enjoyable.	0	3	2	11	11	4.11	0.97
I found the practice of solving privacy CTF challenges as an engaging way of learning.	0	2	4	10	11	4.11	0.93
Learning Effectiveness in terms of Corporate Privacy (LEC)							
Solving privacy CTF challenges helped me to ...							
... reason about privacy risks that businesses face.	0	6	6	11	4	3.48	1.01
... reason about implications of privacy risks that businesses face.	0	4	11	10	2	3.37	0.84
... identify measures that businesses should take to mitigate privacy risks.	1	4	9	11	2	3.33	0.96
Learning Effectiveness in terms of Personal Privacy (LEP)							
Solving privacy CTF challenges helped me to							
... learn about online privacy risks.	0	2	2	15	8	4.07	0.83
... reason about personal privacy risks.	0	3	5	13	6	3.82	0.92
... reason about implications of personal privacy risks.	0	2	4	18	3	3.82	0.74
... defend against personal privacy risks.	1	3	6	16	1	3.48	0.89
Learning Flexibility (LF)							
The CTF system gave me more flexibility in learning the course content.	0	3	6	13	5	3.74	0.90
The CTF system allowed me to learn and complete the assignments at my own pace.	0	2	3	6	16	4.33	0.96
The CTF system allowed me to arrange my study for the class more effectively.	2	3	7	6	9	3.63	1.28
Motivation (MOT)							
The privacy CTF challenges motivated me to learn the course content.	0	1	7	8	11	4.07	0.92
I feel my motivation to learn the course content increased by solving the privacy challenges.	0	3	2	10	12	4.15	0.99
The privacy CTF challenges helped enhance my motivation in learning the course content.	1	2	3	12	9	3.96	1.06
Satisfaction (SAT)							
Overall, I am pleased with the privacy CTF challenges.	0	2	1	14	10	4.19	0.83
Overall, the privacy CTF challenges are pleasant to me.	0	1	5	13	8	4.00	0.92
Overall, I am satisfied with the privacy CTF challenges.	0	1	5	14	7	4.00	0.78
Overall, the privacy CTF challenges satisfy my learning needs.	1	3	6	12	5	3.63	1.04
Task-Technology-Fit (TTF)							
The concepts addressed by the privacy CTF challenges							
... match my learning needs.	0	2	4	15	6	3.93	0.83
... are compatible with my learning needs.	0	0	5	17	5	4.00	0.62
... comprehensively address my learning needs.	0	2	11	10	4	3.59	0.84
My learning goals and needs are met by solving the privacy CTF challenges.	1	1	12	8	5	3.56	0.97

[a] Columns "1" to "5" show participant's response counts for the answer options "Strongly disagree" (1), "Disagree" (2), "Neutral" (3), "Agree" (4) and "Strongly agree" (5).
[b] The construct *Confidence (CON)* was measured both in the pre- and post-challenge questionnaires.

$SD = 0.59$ ($r = 0.68$, $p < 0.01$). The effect size of this change is considered to be high according to Cohen [5].

Figure 2 displays the score distributions of each attitudinal construct measured in the post-challenge questionnaire as a violin plot. Additionally, the mean scores are represented as red diamonds. Participants' construct scores were calculated by averaging all the respective items. All variables have an average score above 3, indicating a moderately high level.

4.2.3 Difficulty and Qualitative Feedback. The difficulty of the challenges was rated as neither too easy nor too difficult. This is evident in the responses to the statement *"In overall, I found the difficulty of the privacy challenges contained in the CTF system as ... "*, reported on a 3-point scale (1 = *too easy*, 2 = *just right*, 3 = *too hard*), with the mean score of 2.07 ($SD = 0.38$). At the same time, participants' responses exhibited a neutral stance regarding the general balance of the challenges in terms of their difficulty, as indicated by an average score of 3.37 ($SD = 0.97$) to the statement *"The privacy CTF challenges were balanced in terms of their difficulty"*. The existence of sub-challenges, where solving easier ones unlocks more

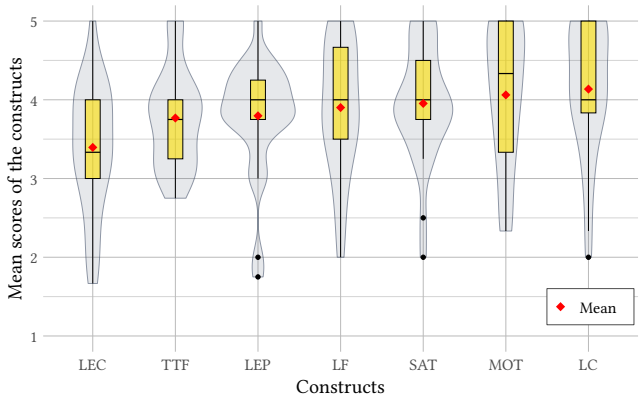

Figure 2: Mean scores of the attitudinal constructs.

difficult ones, had a neutral impact on students' learning process, as reflected in their responses to the statement *"The provision of easy to hard privacy sub-challenges made my learning process easier"* ($M = 3.33$, $SD = 1.04$).

Open-ended responses to a question about students' overall learning experience revealed a number of positive and negative aspects. The most frequently appreciated aspects were self-paced learning (12), enjoyment (8), task variety (6), motivation and effectiveness (4), hands-on approach (3), and choice of topics (2). Additionally, two course participants endorsed the competitive nature of the challenges, as reflected in the following comment: "*The competitiveness and curiosity the CTF system created made it kind of addictive.*" Conversely, students expressed dissatisfaction due to insufficient challenge descriptions (4), a lack of hints (3), too difficult tasks (3), or missing information about what tools to use (2).

5 DISCUSSION

In this paper, we present the implementation and validation of a set of privacy CTF challenges, that were evaluated by Master students with a technical background in an educational experiment. In the following, we share our proof-of-concept experience by summarizing our findings and sharing lessons learned as a roadmap for course instructors and developers of CTF platforms.

Overall, students' feedback on our challenges was consistently positive, as evidenced by moderately high scores across the most attitudinal constructs elicited in the post-challenge questionnaire, particularly in terms of learning climate, learning flexibility, motivation, and satisfaction. Students' higher ratings after the use of the CTF system and understanding of privacy and associated risks in a personal context are also reflected in the open-ended question, in which one participant explicitly expressed a desire to learn more about privacy-preserving techniques employed in organizations. Additionally, our experiment had a significant positive impact on students' self-reported confidence in explaining online privacy concepts, recognizing attacks, and protecting their own online privacy. Similar effects of the CTF participation on reported self-confidence were obtained in [23]. There was also a tendency, although not statistically significant, for increased students' online privacy concerns.

5.1 Lessons Learned

Besides a few recommendations on how to improve specific challenges, the students' feedback and our own reflections provided us with several suggestions for the design and use of CTF challenges.

Eliminate misunderstandings about the CTF teaching format in advance. As mentioned before, some students criticized that some task descriptions were vague and lacked precise information on what tools or techniques to use. It is worth noting here that only 21.4% of our participants were familiar with CTF competitions, as stated in the pre-challenge questionnaire. Although we had a brief introductory session and explained to students the general format and goals of a typical CTF competition, we encourage instructors to explicitly inform students of the intentionally exploratory, creative, and independent nature of such competitions before their participation in order to avoid potential misconceptions.

Ensure the correct functioning of challenges in all environments. The solution path for one of the challenges on browser fingerprinting turned out to be browser-specific, as some browsers did not offer all the necessary settings to successfully solve the task. We updated this challenge retrospectively, and recommend a comprehensive functionality check across all the potential environments.

Offer hints to students through the CTF platform. Although students were encouraged to contact course instructors by e-mail in case of any obstacles, we have not received any requests. Following students' feedback on missing hints, we suggest integrating helpful cues for the most difficult challenges within the CTF system itself.

5.2 Limitations

Our work has limitations inherent to survey based-studies. The small and homogeneous (young age, students) sample size resulting from the need for participants to be registered for the seminar limits the validity and generalisability of our results. Some of the questionnaires could not be linked due to the absence of students, which further reduced the sample size for inferential statistics. In general, reliance on self-reported data also introduces challenges, such as the potential for inaccurate responses, or strategic responses [18]. Participants' responses may also be influenced by social desirability bias, where students answer in ways they think are socially acceptable rather than reflecting their true thoughts, to please others or avoid disapproval [16].

Another limitation stems from the absence of a control group exposed to traditional pen-and-paper methods and exercise sheets. While a parallel seminar group could address this, its realisation would introduce additional educational complexity. Furthermore, we note that the effect of CTF participation on reported self-confidence may have been confounded by external factors, such as gaining further knowledge about privacy in related courses.

6 CONCLUSION

This study is intended to serve as a blueprint for future university courses with the goal of engaging and educating the next generation of computer science students on the fundamentals of privacy through the use of capture-the-flag (CTF) challenges.

As the presented tasks cover only selected aspects of privacy, future work may build up on our work and expand the selection of the taught concepts and techniques. Specifically, privacy attacks on traffic metadata (e. g., onion routing, mix networks), or on location metadata used as quasi-identifiers [9, 12] may be introduced. Additionally, privacy concerns related to neural networks and the burgeoning field of large language models [1] provide further potential fields of interest to be addressed. Finally, the desire of students to learn more about privacy in a business context may be approached by discussing secure multi-party computation, where several parties jointly compute a function over confidential inputs [14].

ACKNOWLEDGMENTS

The authors thank the staff members of the University's of Innsbruck Security & Privacy Lab and Dr. Maximilian Hils for their helpful comments and efforts in pretesting the challenges.

REFERENCES

[1] Carlini, N., Liu, C., Erlingsson, Ú., Kos, J., and Song, D. The Secret Sharer: Evaluating and Testing Unintended Memorization in Neural Networks. In *Proceedings of the 28th USENIX Security Symposium (USENIX Security 19)* (2019), pp. 267–284.

[2] Carlisle, M., Chiaramonte, M., and Caswell, D. Using CTFs for an Undergraduate Cyber Education. In *2015 USENIX Summit on Gaming, Games, and Gamification in Security Education (3GSE 15)* (Washington, D.C., 2015), USENIX Association.

[3] CC2020 Task Force. *Computing Curricula 2020: Paradigms for Global Computing Education.* Association for Computing Machinery, New York, NY, USA, 2020.

[4] Chung, K., and Cohen, J. Learning Obstacles in the Capture The Flag Model . In *Proceedings of 2014 USENIX Summit on Gaming, Games, and Gamification in Security Education (3GSE 14)* (San Diego, CA, 2014), USENIX Association.

[5] Cohen, J. *Statistical power analysis for the behavioral sciences.* Lawrence Erlbaum Associates, 2013.

[6] Cole, S. V. Impact of Capture-The-Flag (CTF)-style vs. Traditional Exercises in an Introductory Computer Security Class. In *Proceedings of the 27th ACM Conference on on Innovation and Technology in Computer Science Education Vol. 1* (2022), pp. 470–476.

[7] CS2023 Task Force. *Computer Science Curricula 2023: Version Beta.* Association for Computing Machinery, IEEE-Computer Society, and Association for Advancement of Artificial Intelligence, 2023.

[8] Denning, D. E. R. *Cryptography and Data Security*, vol. 112. Addison-Wesley Reading, 1982.

[9] Douriez, M., Doraiswamy, H., Freire, J., and Silva, C. T. Anonymizing NYC Taxi Data: Does It Matter? In *2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA)* (2016), pp. 140–148.

[10] Egelman, S., Bernd, J., Friedland, G., and Garcia, D. The Teaching Privacy Curriculum. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education* (2016), pp. 591–596.

[11] Fischer-Hübner, S. Teaching Privacy as a Part of the Computer Science Curriculum. *The Impact of Information Technology: From practice to curriculum* (1996), 47–53.

[12] Gambs, S., Killijian, M.-O., and del Prado Cortez, M. N. De-Anonymization Attack on Geolocated Data. *Journal of Computer and System Sciences 80*, 8 (2014), 1597–1614.

[13] Garfinkel, S., Abowd, J. M., and Martindale, C. Understanding Database Reconstruction Attacks on Public Data. *Communications of the ACM 62*, 3 (2019), 46–53.

[14] Goldreich, O. Secure Multi-Party Computation. *Manuscript. Preliminary version 78*, 110 (1998), 1–108.

[15] Gómez-Boix, A., Laperdrix, P., and Baudry, B. Hiding in the Crowd: An Analysis of the Effectiveness of Browser Fingerprinting at Large Scale. In *Proceedings of the 2018 World Wide Web Conference* (2018), pp. 309–318.

[16] Grimm, P. Social Desirability Bias. *Wiley International Encyclopedia of Marketing* (2010).

[17] Gross, T. Validity and Reliability of the Scale Internet Users' Information Privacy Concerns (IUIPC). *Proceedings of the Symposium on Privacy Enhancing Technologies* (2021).

[18] Hansen, M. H., Hurwitz, W. N., Marks, E. S., and Mauldin, W. P. Response Errors in Surveys. *Journal of the American Statistical Association 46*, 254 (1951), 147–190.

[19] Hils, M., Woods, D. W., and Böhme, R. Privacy Preference Signals: Past, Present and Future. *Proceedings on Privacy Enhancing Technologies*, 4 (2021), 249–269.

[20] Hylkema, M. A Survey of Database Inference Attack Prevention Methods. *Educational Technology Research* (2009).

[21] Kaplan, Z., Zhang, N., and Cole, S. V. A Capture The Flag (CTF) Platform and Exercises for an Intro to Computer Security Class. In *Proceedings of the 27th ACM Conference on on Innovation and Technology in Computer Science Education Vol. 2 (ITiCSE 2022)* (2022), ACM, pp. 597–598.

[22] Laperdrix, P., Bielova, N., Baudry, B., and Avoine, G. Browser Fingerprinting: A Survey. *ACM Transactions on the Web (TWEB) 14*, 2 (2020), 1–33.

[23] Leune, K., and Petrilli Jr, S. J. Using Capture-the-Flag to Enhance the Effectiveness of Cybersecurity Education. In *Proceedings of the 18th Annual Conference on Information Technology Education* (2017), ACM, pp. 47–52.

[24] Li, T., and Li, N. On the Tradeoff between Privacy and Utility in Data Publishing. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (New York, NY, USA, 2009), KDD '09, Association for Computing Machinery, p. 517–526.

[25] Machanavajjhala, A., Kifer, D., Gehrke, J., and Venkitasubramaniam, M. l-diversity: Privacy beyond k-anonymity. *ACM Transactions on Knowledge Discovery from Data (TKDD) 1*, 1 (2007), 3–es.

[26] Merener, M. M. Theoretical Results on De-anonymization via Linkage Attacks. *Transactions on Data Privacy 5*, 2 (2012), 377–402.

[27] Mirkovic, J., and Peterson, P. A. H. Class Capture-the-Flag Exercises. In *2014 USENIX Summit on Gaming, Games, and Gamification in Security Education (3GSE 14)* (San Diego, CA, 2014), USENIX Association.

[28] Nurse, J. R., Adamos, K., Grammatopoulos, A., and Di Franco, F. Addressing the EU Cybersecurity Skills Shortage and Gap Through Higher Education. *European Union Agency for Cybersecurity (ENISA) Report* (2021).

[29] Spiekermann, S., and Cranor, L. F. Engineering Privacy. *IEEE Transactions on Software Engineering 35*, 1 (2008), 67–82.

[30] Sweeney, L. k-anonymity: A Model for Protecting Privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-based Systems 10*, 05 (2002), 557–570.

[31] Troncoso, C. *The Cyber Security Body of Knowledge v1.1.0, 2021.* University of Bristol, 2021, ch. Privacy & Online Rights. KA Version 1.0.2.

[32] Vaidya, J., Shafiq, B., Lorenzi, D., and Badar, N. Incorporating Privacy into the Undergraduate Curriculum. In *Proceedings of the 2013 on InfoSecCD'13: Information Security Curriculum Development Conference* (2013), pp. 1–7.

[33] Vykopal, J., Švábenskỳ, V., and Chang, E.-C. Benefits and Pitfalls of Using Capture the Flag Games in University Courses. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education* (New York, NY, USA, 2020), SIGCSE '20, Association for Computing Machinery, p. 752–758.

[34] Waldman, A. E. Cognitive Biases, Dark Patterns, and the 'Privacy Paradox'. *Current Opinion in Psychology 31* (2020), 105–109.

[35] Warren, S. D., and Brandeis, L. D. The Right to Privacy. *Harvard Law Review 4*, 5 (1890), 193–220.

[36] Westin, A. F. Privacy and Freedom. *Washington and Lee Law Review 25*, 1 (1968), 166.

[37] Zhang, Y. G., and Dang, M. Y. Understanding Essential Factors in Influencing Technology-Supported Learning: A Model toward Blended Learning Success. *Journal of Information Technology Education. Research 19* (2020), 489.

[38] Švábenský, V., Čeleda, P., Vykopal, J., and Brišáková, S. Cybersecurity Knowledge and Skills Taught in Capture the Flag Challenges. *Computers & Security 102* (2021), 1–14.