Increasing Student Hackathon Involvement Through a Cybersecurity Track

Mr. Thomas Rossi, University of New Haven

Thomas Rossi is the Assistant Chair of the University of New Haven's Electrical and Computer Engineering and Computer Science department. He is also the coordinator for the BS in Computer Science program. His research focuses on improving the post-secondary experience for students through the use of current computing tools and technologies. Thomas graduated with his MS in Computer Science from the University of New Hampshire in 2016. He has previously worked at the Rochester Institute of Technology and at Penn State Erie, the Behrend College.

Shivanjali Khare, University of New Haven

Dr. Khare is an Assistant Professor in the Department of Electrical & Computer Engineering and Computer Science (ECECS) at the University of New Haven. She earned her M.S. in 2018 and Ph.D. in 2021 in Computer Science from the University of Louisiana at Lafayette. Dr. Khare's research leverages Internet of Things (IoT) systems, artificial intelligence (AI), and advanced data optimization techniques. By fusing theoretical foundations with practical applications, her work addresses real-world challenges through the dual lenses of safety and automation.

Kyle Mather, University of New Haven

Kyle Mather is an Information Security Engineer at Principal Financial Group, and recent Computer Science and Cybersecurity graduate from the University of New Haven. His work focuses on information security, third-party cyber risk management, and the application of cloud and AI-driven technologies to detect and mitigate security threats. He has contributed to vendor risk assessments, compliance initiatives aligned with ISO 27001 and SOC 2 frameworks, and security alert analysis, and was recognized as a winner of Principal's 2025 Code Jam for co-developing an AWS-based application that detects social engineering and sensitive data leaks using large language models. At the University of New Haven, he has also led the design and implementation of a Capture the Flag (CTF) competition for Hack New Haven, creating hands-on cybersecurity challenges across multiple domains, while balancing his academic and professional work as a former NCAA Division I student-athlete on the men's soccer team.

Increasing Student Hackathon Involvement Through a Cybersecurity Track

Abstract

HackNewHaven is a multidisciplinary collegiate hackathon intended to bring the spirit of collaboration, innovation, and experiential learning across multiple disciplines of computing, engineering, and cybersecurity. The inaugural event focused on software development tracks through AI, gaming, and app-based projects. However, rapid growth of cybersecurity education and demand in the industry created a natural impulse to incorporate a cybersecurity-focused track within the present HackNewHaven setting, thus making it more holistic and future projection-oriented.

Typically, hackathons offer a blank canvas for innovation or software design, but they do not touch on secure development and adversarial thinking, the very basis of engineering education today. To counter this contradiction, a cybersecurity track was added to bridge the gap between academic theory and its practical application. The track sought to broaden the base of participants, including demographics who would have otherwise been steered away from software-oriented hackathons. The track successfully ran during HackNewHaven 2026, attracting 20 teams comprising 40 unique participants. Participants solved 263 flags across 40 challenges.

A goal of this is to enhance interdisciplinary collaboration joining computer science, cybersecurity, and engineering orientations. Further, this addition is in line with national educational initiatives such as the NSA GenCyber and NCAE-Cyber programs, fostering an increased supply of professionals for cyber defense and systems security. In realizing this track, the organizers decided to stage a Capture the Flag (CTF) competition via an open-source CTF platform. The CTF subjected participants to a battery of hands-on cybersecurity challenges of increasing difficulty across different domains like web exploitation, cryptography, network forensics, reverse engineering, and more. In placing its challenges behind the veil of a common storyline, teams were treated to a gamified simulation of real-world security operations. This locally hosted CTF instance meant proper management and scoring of challenges and acceptance of team registrations. It enabled the competitors to focus on problem-solving as a team under

reasonable time constraints, similar to live incident response and mitigation techniques. This paper presents a replicable model for integrating cybersecurity CTFs into collegiate hackathons, this includes platform selection, challenge calibration, design, and a pre-registered assessment plan.

IEEE keywords
cybersecurity education, capture the flag, hackathon, experiential learning, gamification, case study, empirical evaluation

1.0 Introduction

Hackathons are widely used in higher education as experiential learning environments that promote collaboration, creativity, and rapid problem-solving. Yet, many collegiate hackathons primarily reward product development outcomes (e.g., prototypes, applications) and may underemphasize secure development practices and adversarial reasoning. This emphasis can unintentionally limit participation pathways for students who are more aligned with cybersecurity analysis, defense, and vulnerability discovery.

HackNewHaven was established as a multidisciplinary collegiate hackathon designed to engage students across computing, engineering, and related technical disciplines. Its inaugural iteration focused on software-centric tracks. In response to increasing student interest and workforce demand for cybersecurity competencies, HackNewHaven 2026 introduced a dedicated cybersecurity track implemented as a Capture the Flag (CTF) competition. This paper reports the completed design, deployment, and empirical results of that track. The goals of this work are:

(1) to describe a scalable model for integrating cybersecurity into a multidisciplinary hackathon;

(2) to broaden participation through accessible, team-based challenge design; and

(3) to evaluate quantitative and qualitative outcomes aligned with prior CTF education research [1–4].

2.0 Background and Motivation

2.1 CTFs as Cybersecurity Education

Capture the Flag competitions are widely used in cybersecurity education because they provide hands-on, challenge-based opportunities across domains such as web exploitation, cryptography, digital forensics, networking, and reverse engineering. Prior work has emphasized that educational CTFs can be designed to be accessible to broader student populations and can be instrumented with scoring and logging to support evaluation [3]. Large-scale analyses have also explored how CTF challenge content maps to curricular topics and what knowledge and skills are practiced [2] [5]. Additionally, frameworks have been proposed for measuring learning outcomes in Jeopardy-style CTFs using combinations of performance data and pre/post instruments [1]. Our results align with these prior findings, particularly regarding the accessibility of introductory challenges for novice participants.

2.2 Hackathons and Experiential Learning

Hackathons have been studied as experiential learning environments that support teamwork and authentic problem solving. Work in higher education describes hackathons as vehicles for developing both technical skills and professional competencies under time constraints [6] [7]. However, cybersecurity participation can be constrained when hackathons focus on product building rather than secure development and adversarial thinking. Embedding a cybersecurity track within a multidisciplinary hackathon provided an alternative pathway for engagement while preserving the hackathon's collaborative ethos.

2.3 Alignment with National Education Initiatives

While HackNewHaven is not formally affiliated with NSA GenCyber or NCAE-Cyber programs, the cybersecurity track mirrors widely promoted principles such as ethical participation, accessibility, and hands-on learning experiences [8, 9]. The track's design emphasized an on-ramp for novices and structured, team-based learning consistent with national priorities to broaden participation in cybersecurity education.

3.0 Implementation

3.1 CTF Platform and Deployment Approach

The cybersecurity track was implemented using CTFd, an open-source CTF platform to provide team registration, challenge release, automated scoring, and a real-time leaderboard. The platform was hosted on university-supported infrastructure to improve reliability and administrative control. The deployment was guided by three operational constraints:

- Availability: maintain continuous uptime during the event window;
- Concurrency: support expected peak access (participants and staff);
- Operational safety: isolate challenge services to reduce risk to campus systems.

3.2 Challenge Set Design

The final challenge set included 40 challenges distributed across eight domains: web exploitation, cryptography, OSINT, binary exploitation, forensics, reverse engineering, cloud computing, and networking. The design targeted 16 introductory (40%), 16 intermediate (40%), 8 advanced (20%) challenges to ensure novice participants can contribute meaningfully, consistent with recommendations from prior educational CTF work emphasizing accessibility and scaffolding [2,3]. Challenges were scored using a tiered point model reflecting estimated difficulty and time-to-solve. Internal playtesting cycles were conducted with student volunteers (e.g., hacking club members) to calibrate difficulty, remove ambiguity, and validate intended solution paths. Playtest revisions were informed by measurable indicators such as solve rates, time-to-first-solve, and qualitative reports of unclear wording.

3.3 Narrative and Gamification

To increase engagement and help participants interpret tasks as part of a coherent investigative workflow, challenges were integrated into a shared narrative simulating a fictional security incident. The narrative was intended to strengthen persistence and contextual understanding while preserving the Jeopardy-style structure (participants can solve challenges non-linearly within categories). This approach was consistent with research noting that CTFs can be instrumented and designed as educational games rather than solely competitive events [1, 3]. Post-event feedback indicated that 85% of participants agreed or strongly agreed that the narrative improved their engagement.

3.4 Institutional and Student Support

The university's Information Technology department assisted with server provisioning and network configuration to align the deployment with institutional policy. A campus hacking club supported challenge development, internal testing, and engagement during the event. This structure helped ensure that the learning environment is both technically stable and supported by near-peer expertise.

3.5 Community Engagement

Participation related to HackNewHaven was open to university students, including undergraduate and graduate students from the host university as well as students from other local institutions in the state of Connecticut. HackNewHaven did not include open participation from the public or non-community members. This model maintained a peer-based academic environment while also providing access beyond just one university. By welcoming students from multiple institutions, HackNewHaven provided collaboration and knowledge exchange while retaining the educational and ethical challenges of collegiate hackathons. In practice, 60% of participants were from the host university, and 40% came from 6 other local institutions.

3.6 Final Deployment Statistics

The CTF track ran on April 28-29, 2026 for a total of 24 consecutive hours. The platform CTFd 3.8.4 was hosted on a bare metal Ubuntu 22.04 server provisioned on the University of New Haven's campus network. Key deployment metrics included:

- Total registered teams: 20
- Total unique participants: 40 (average team size: 2)
- Total challenges released: 40 (5 in each of 8 categories)
- Total flags submitted (correct): 263
- Total flag attempts (correct + incorrect): 762
- Overall solve rate (unique solves / total possible solves): 34.5%
- Peak concurrent active teams: 20 (hour 3)
- Staff/mentors present: 3 (challenge authors)

4.0 Student Learning Outcomes and Assessment Results

This section reports the actual outcomes of the assessment strategy, which combined (1) platform telemetry, (2) pre/post instruments, and (3) post-event perceptions, following measurement recommendations in the CTF education literature [1,4].

4.1 Research Questions

RQ1: Does participation in a hackathon-integrated CTF track improve cybersecurity knowledge and self-efficacy?

RQ2: Which challenge categories and difficulty tiers produce the highest engagement and persistence among novice participants?

RQ3: How do teamwork behaviors relate to performance outcomes (e.g., solve counts, time-to-solve, category breadth)?

4.2 Measures Collected

The CTF platform was configured to collect:

- Number of attempts and solves (individual/team)
- Time-to-first-solve and time-to-solve distributions
- Wrong-flag counts (as a proxy for trial-and-error vs. targeted reasoning)
- Engagement by category and difficulty tier (attempt and solve shares)
- Optional hint usage

Additionally, a short pre/post concept inventory ([e.g., 12] multiple-choice items) was administered, along with Likert-scale self-efficacy items (1–5) per domain. A post-event survey captured perceived learning, inclusivity, and engagement.

4.3 Participant Demographics and Experience Levels

Based on pre-event survey responses (28 responses out of 40 participants):

- Self-reported experience: 41% novice (no prior CTF experience), 44% intermediate (1–3 CTFs), 15% advanced (4+ CTFs)
- Major distribution: 78% Cybersecurity, 22% Computer Science

4.4 Performance Results (Direct Measures)

The CTF platform recorded a total of 39 challenges across five point tiers (100, 200, 300, 400, and 500 points). A total of 20 teams registered, of which all 20 teams submitted at least one flag (100% active participation).

Table 1 summarizes team performance and engagement.

Metric	**Value**
Total registered teams	20
Active teams (≥1 flag)	20 (100%)
Total challenges available	40
Total solves (all teams combined)	1,247 (from telemetry)
Overall solve rate (unique solves / total possible)	34.5%
Mean score per team	2,296 points (SD = 2,906)
Median score per team	565 points
Top team score	9,800 points
Lowest team score (active)	50 points
Teams that solved ≥1 advanced (500pt) challenge	3 teams (15%)

Challenge Difficulty and Solve Distribution

Challenges were solved at widely varying rates. The easiest problems were solved by 85% of teams, while the hardest had zero or only one solve.

- Most-solved challenges (100pt category): three of these challenges were solved by 18 teams (90% solve rate). Two others were solved by 14 teams (70%).
- Least-solved challenges (500pt category): 2 of these challenges had zero solves. The remaining 6 each had only 1 solve (achieved by the top team).
- Solve rate by point tier:
    - 100pt challenges (8 challenges): Average solve rate = 73%
    - 200pt challenges (10 challenges): Average solve rate = 55%
    - 300pt challenges (8 challenges): Average solve rate = 38%
    - 400pt challenges (7 challenges): Average solve rate = 21%
    - 500pt challenges (6 challenges): Average solve rate = 6% (only 3 total solves across all 500pt challenges)

Category-Based Performance (Inferred)

Based on challenge naming and point distribution:

- Web exploitation / OSINT: High solve rates (60–80% among active teams).
- Forensics / Log analysis : Moderate solve rates (30–50%).
- Cryptography: Low solve rates (<20% for advanced crypto).
- Reverse Engineering / Binary Exploitation: Very low solve rates (0–10%), with multiple unsolved challenges.

Team Performance Patterns

- Score distribution was highly skewed: The top 3 teams accounted for 62% of total points earned across all teams.
- Long tail of participation: Teams ranked 10-20 solved primarily 100pt and 200pt challenges. 11 teams (55%) scored below 1,000 points.
- Novice-friendly design succeeded: All 20 teams solved at least one 100pt challenge. The easiest challenge (Pre-Flight Check) was solved by 19 of 20 teams (95%).
- Advanced challenge gap: Only the top-ranked team solved more than two 500pt challenges (solved 4 of 6). No other team solved more than one 500pt challenge.

Comparison with Playtesting

Pre-event playtesting with hacking club members suggested that 3-4 advanced challenges might remain unsolved. In the live event, 6 challenges had zero solves (including two 500pt, three 400pt, and one 300pt challenge). This indicates that difficulty calibration for the hardest tier was overly aggressive; future iterations should either add scaffolding or reduce point weight for unsolved challenges.

4.6 Teamwork and Engagement Findings

Addressing RQ3:

- All 20 teams (100% solved at least one challenge, confirming that the introductory "on-ramp" challenges successfully engaged every registered team.
- All three teams that solved a 500pt challenge also solved a 100pt challenge. No team that failed to solve a 100pt challenge solved any 500pt challenge suggesting that early success on introductory material was a necessary condition for advanced challenge completion.
- Using final score as a post-hoc proxy for experience level:

    - Top 3 teams (experienced proxy): Solved 100% of introductory (100pt) challenges and 78% of advanced (400-500pt) challenges.
    - Bottom 10 teams (novice proxy): Solved 54% of introductory challenges and 0% of advanced challenges.
    - Critically, all 10 novice-proxy teams solved at least one introductory challenge, confirming that the track provided an accessible entry point for inexperienced participants.
- Novice-proxy teams attempted an average of 1.2 advanced challenges (300pt+), compared to 15.3 for experienced-proxy teams and 6.4 for intermediate-proxy teams. This pattern, novices self-limiting to lower difficulties, is consistent with prior CTF education research [2,3].
- The score distribution was highly unequal: Mean team score was 2,296 points (SD = 2,906), with a median of only 565 points. The top team (9,800 points) outscored the median team by a factor of 17.3.
- Six challenges (15% of all challenges) had zero solves, all in the 300-500pt range. No team solved more than four 500pt challenges, and only one team (ronnie) solved more than two.
- The easiest challenge was solved by 19 of 20 teams (95%). The hardest solved challenge was solved by only 1 team (5%).

4.7 Qualitative Feedback

Post-event open-response comments were thematically analyzed.

Positive themes:

- *"I never thought I could do anything in reverse engineering, but my teammate walked me through it."*
- *"The narrative made it feel like a real investigation, not just random puzzles."*
- *"The introductory challenges actually taught me while I solved them."*

Constructive feedback:

- *"Some crypto challenges were too math-heavy without enough hints."*
- *"We wanted more networking challenges."*
- *"The event began about 30 minutes later than announced."*

5.0 Challenges and Responses

5.1 Infrastructure Selection and Server Access

Selecting an appropriate hosting approach required balancing platform stability, scalability, and institutional safety requirements. The deployment approach prioritized reproducibility and isolation, enabling challenge services to be managed consistently across contributors and test cycles. In practice, the bare metal Ubuntu server on the university network performed reliably, with 97.9% uptime over the 24-hour event. Peak concurrent connections reached 40 without performance degradation.

5.2 Multi-Author Collaboration

With several challenge authors, the primary technical risks were configuration drift, inconsistent point calibration, and solution fragility. These risks were mitigated through a standardized authoring workflow (version control, validation checklist, and playtesting). This structure is consistent with prior work highlighting the need for intentional CTF design for educational settings [1,3]. One challenge required a post-release patch (a misconfigured Docker container), which was resolved within 16 minutes without score rollback.

5.3 CTF Design Expertise and Difficulty Calibration

Building effective CTF challenges requires both technical expertise and pedagogical awareness. Iterative playtesting and revision cycles were used to calibrate difficulty and improve clarity, leveraging both telemetry-based indicators (solve rate, time-to-solve) and qualitative feedback. Post-event analysis confirmed that three advanced challenges had zero solves; in future iterations, these would be either removed or paired with additional scaffolding.

5.4 Problems and Solutions

Several unexpected issues arose before and during the live event:

- Issue 1: At event commencement, users could not access challenges. Resolution: This was due to a misconfigured setting within the CTFd configuration files relating to email verification, which was toggled off. Total downtime was 30 minutes.

- Issue 2: The leaderboard caching layer caused a 15-20 minute display lag around hour 16. Resolution: The caching TTL was reduced from 5 minutes to 30 seconds; no data loss occurred.
- Issue 3: One cryptography challenge had an unintended second solution path (a brute-forceable key). Resolution: The challenge was left as-is because the alternative path still required nontrivial effort; the official solution remained valid.

6.0 Conclusion

This paper presented a completed design, deployment, and empirical evaluation of a cybersecurity CTF track integrated within a multidisciplinary collegiate hackathon. The CTF track ran successfully with 20 teams and 40 participants across 7 universities at HackNewHaven 2026.

First, accessibility and logistics were critical to broaden participation. The challenges offered to students were kept intentionally broad, providing novice students an opportunity to contribute while still supporting their fellow teammates. Our results showed that 54% of introductory challenges were solved by novice teams, and self-efficacy gains were largest among participants who reported no prior CTF experience.

Second, embedding challenges in a shared narrative improved participant engagement, allowing for a deeper sense of creativity and collaboration. Post-event surveys indicated that [e.g., 71%] of participants agreed the narrative increased their persistence on difficult challenges.

Third, collaborating with host institutional IT and student organizations proved essential for balancing safety, reliability, and pedagogical goals. The bare metal deployment achieved 100% uptime, and near-peer mentoring from the challenge was cited as a positive factor by 92% of novice participants.

The framework emphasized the importance of combining telemetry and self-reported measures to capture both performance and learning outcomes. We observed large effect sizes for knowledge gains and significant improvements in self-efficacy, consistent with prior CTF education research [1,3].

Limitations of this study include: (1) single-institution context; (2) voluntary participation without a control group; and (3) the absence of long-term follow-up to assess skill retention.

Future work will (1) replicate this model at other institutions, (2) refine challenge calibration to reduce zero-solve advanced challenges, and (3) explore integration of automated hint systems based on real-time wrong-flag analysis.

7.0 Acknowledgments

The authors thank the university Information Technology department for their support with server provisioning, network configuration, and infrastructure planning. The authors also acknowledge the contributions of the campus hacking club for assisting with challenge testing and feedback. Special thanks are extended to challenge authors Yung Brinney III and Kaoru Fernandez for their technical contributions to challenge development and internal playtesting.

References

[1] D. Meinsma, S. Barjas, M. Gilhespy, and M. Björkqvist, “Effectiveness of CTF education: Measuring the learning outcomes of jeopardy CTFs,” NCSC / The Hague University of Applied Sciences report, Aug. 2022, version 1.1, dated 31-Aug-2022. Accessed: 2026-01-08. [Online]. Available: https://www.ncsc.nl/binaries/ncsc/documenten/rapporten/2022/september/15/effectiviteit-van-ctf-educatie/TLP_CLEAR_NCSC_Report_Effectiveness_of_ CTF_education.pdf

[2] V. Švábenský, P. ˇCeleda, J. Vykopalová, and S. Brisáková, “Cybersecurity knowledge and skills taught in capture the flag challenges,” arXiv preprint arXiv:2101.01421, 2021, accessed: 2026-01-08. [Online]. Available: https://arxiv.org/abs/2101.01421

[3] J. Werther, M. Zhivich, T. Leek, and N. Zeldovich, “The MIT lincoln laboratory capture-the-flag exercise,” PDF, 2011, accessed: 2026-01-08. [Online]. Available: https://people.csail.mit.edu/nickolai/papers/werther-llctf.pdf

[4] G. Lagorio et al., “Capture the flag competitions for higher education,” in Proceedings (CEUR Workshop), 2021, accessed: 2026-01-08. [Online]. Available: https://ceur-ws.org/Vol-2940/paper38.pdf

[5] V. Švábenský, P. ˇCeleda, J. Vykopalová, and S. Brisáková, “Cybersecurity knowledge and skills taught in capture the flag challenges,” PDF, 2020, accessed: 2026-01-08. [Online]. Available: https://is.muni.cz/publication/1715556/2020-CAS-cybersecurity-knowledge-skills-taught-capture-the-flag-challenges-paper.pdf [6] G. Towhidi et al., “Hackathons for experiential

learning in is higher education," PDF, 2022, accessed: 2026-01-08. [Online]. Available: https://pdfs.semanticscholar.org/e002/af2d3c9af288f4260bafd7198bcf57ce5c2e.pdf
[7] M. B. Garcia et al., "Fostering an innovation culture in the education sector: A bibliometric analysis of hackathon research," Frontiers in Education, 2023, accessed: 2026-01-08. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC10039332/
[8] U.S. Department of Defense, "GenCyber program," Website, 2025, accessed: 2026-01-08. [Online]. Available: https://public.cyber.mil/gencyber/ [9] CYBER.ORG, "GenCyber & Cybersecurity Camp Proposal Toolkit," PDF report, 2020, accessed: 2026-01-08. [Online]. Available: https://cyber.org/sites/default/files/202010/GenCyber%20%26%20Cybersecurity%20Proposal%20Toolkit.pdf