EAGER: SaTC-EDU: Training Mid-Career Security Professionals in AI & Data-Driven Cybersecurity

Yuxin Chen, Nick Feamster, Blase Ur
University of Chicago

Project Summary

This project will perform exploratory, early-stage work to develop a curriculum for large-scale, online training of mid-career security professionals who aim to develop the skills to apply both conventional and cutting-edge machine learning tools to cybersecurity. The outcomes of the project will include curriculum development at the intersection of machine learning and security—with a focus on applications of machine learning to practical, real-world security use cases—establishing a pedagogical foundation for security researchers to evaluate and apply various potential ML-based approaches to cybersecurity. We focus in particular on training mid-career professionals who have a classical training in security (and thus understanding of practical concepts), but need to gain a stronger foundation in data-driven methods, which have become the basis for most applied cybersecurity in the past decade.

Cybersecurity and machine learning have progressed as relatively disjoint fields. Occasionally there is overlap between the two fields, generally taking the form of either (1) applications of machine learning to statistical anomaly detection (e.g., malware detection); or, on the other hand (2) adversarial attacks on machine learning detection algorithms (e.g., adversarial ML). By and large, however, the concepts and pedagogy of these two fields have evolved independently, which comes at the expense of both research and education. Both fields are also rapidly advancing, which makes education both in these respective fields and at their intersection critical.

Intellectual Merit. This effort will lead to advances in research areas of both machine learning and cybersecurity. Machine learning research is continually in search of application domains where new models and algorithms can be applied. This nexus will offer opportunities for students to learn about state-of-the-art machine learning techniques in the context of security problems. Security research and practice, on the other hand, can benefit from applications of cutting-edge machine learning models to an array of cybersecurity use cases. There is a significant opportunity for new advances that require both domain knowledge in security and intimate familiarity with machine learning fundamentals. We will also advance research on machine teaching, focusing on the optimization and customization of the teaching materials we develop.

Broader Impacts. Our advancement and re-skilling the U.S. cybersecurity workforce's through large-scale, online training in data-driven and ML methods is critical for keeping the country secure. Through formative human-subjects protocols with U.S. industry leaders, we will align the curriculum we develop with the needs and use cases of cybersecurity in practice. Furthermore, the development of a curriculum at the intersection of machine learning and cybersecurity will have broader impacts for the research community in both of these fields, opening up new research areas that require domain expertise in both areas. For example, one particular area where a significant gap exists across both areas involves representation of timeseries data for outlier detection. In the case of network traffic traces, for example, questions arise about sampling, transformations, and aggregation, and how those operations affect deployability and model accuracy. Solving these problems requires a deep understanding of *both* machine learning fundamentals *and* cybersecurity domain knowledge, which our planned curriculum aims to develop.

EAGER: SaTC-EDU: Training Mid-Career Security Professionals in AI & Data-Driven Cybersecurity

Yuxin Chen, Nick Feamster, Blase Ur
University of Chicago

1 Motivation

This project will perform exploratory, early stage work to develop a curriculum for large-scale, online training of mid-career security professionals who aim to develop the skills to apply both conventional and cutting edge machine learning tools to cybersecurity. The outcomes of the project will include curriculum development at the intersection of machine learning and security—with a focus on applications of machine learning to practical, real-world security use cases—establishing a pedagogical foundation for security researchers to evaluate and apply various potential ML-based approaches to cybersecurity. We focus in particular on training mid-career professionals, who have a classical training in security (and thus understanding of practical concepts), but need to gain a stronger foundation in data-driven methods, which have become the basis for most applied cybersecurity in the past decade.

Cybersecurity and machine learning have progressed as relatively disjoint fields. Occasionally there is overlap between the two fields, generally taking the form of either (1) applications of machine learning to statistical anomaly detection (e.g., malware detection); or, on the other hand (2) adversarial attacks on machine learning detection algorithms (e.g., adversarial ML). By and large, however, the concepts and pedagogy of these two fields have evolved independently, which comes at the expense of both research and education. Both fields are also rapidly advancing, which makes education both in these respective fields and at their intersection critical.

Security emphasizes abstraction and APIs, applies methods as black box. As a field within computer science, security has disparate roots in both math (e.g., cryptography) and systems. In terms of machine learning, however, security has generally applied "off the shelf, black box" machine learning algorithms, often without careful attention to the importance of considerations like feature engineering, representation learning, model selection, and model evaluation. Systems security focuses extensively on threat modeling and software vulnerabilities, giving rise to training in systems-oriented considerations. More generally, software development (and computer science) views abstraction as a critical building block to scalable, deployable systems. Yet, when applying machine learning algorithms, such abstraction can be harmful, if it prevents a practitioner from fully understanding the underlying characteristics of the data or model that gave rise to the predictions or understanding a limitation of the models that are being applied to the data. Specifically, a simple change to a function call can have significant implications to the resulting performance and predictions for an intrusion detection system. Yet, those functions abstract important details about the model, from how error is minimized to assumptions about the statistical distributions on data. Without an understanding of the underlying mechanics, security practitioners may apply existing algorithms and models incorrectly, or fail to recognize their limitations.

Machine learning emphasizes model optimization, often without regard to practical considerations other than accuracy. Machine learning involves developing functions that learn on data; it has roots in statistics and applied mathematics, and as such the practice and pedagogy are typically tuned on optimization and accuracy, often without careful consideration of practical, systems-oriented constraints, ranging from real-time performance constraints to ethics and privacy concerns. Yet, whether a machine learning algorithm is deployable in practice depends on how well it performs on available data—which, in a security context, may be highly constrained, due to practical systems, privacy, or other constraints.

The security and machine learning pedagogy itself is currently even more divergent. Existing security courses typically focus on conventional threat modeling, cryptography, and systems-level security issues (e.g., memory protection), with little attention paid to statistical learning and AI as the basis for securing computer systems. On the other hand, the machine learning pedagogy is by and large focused on the theory and practice of the machine learning models—but while practitioners indeed need to understand how those models work, they also need to learn how to *apply those models to data*.

2 Project Concept

The planned project involves the development of an online curriculum in Machine Learning and Cybersecurity that bridges the divide between these two disciplines. Our target audience are members of the workforce who have technical training in security operations—as system administrators, DevOps engineers, or software development—and seek to gain data-oriented skills for data-driven cybersecurity.

2.1 Bridging the Pedagogical Divide

When considering applications of machine learning to security problems, design decisions must consider what is happening "under the hood". Similarly, training in machine learning should incorporate not only the theoretical mechanics of the models themselves, but implications of the design of these models for applications to real-world datasets. This EAGER project aims to bridge this pedagogical gap, through the development of new material in both cyber security and machine learning:

- We aim to imbue statistical and machine learning methods in cybersecurity course material, so that security practitioners develop better fundamental knowledge and intuition when applying machine learning models in security contexts.
- We aim to orient machine learning concepts towards practical applications in cybersecurity, developing a new curriculum that focuses not only on the theoretical underpinnings and mechanics of the machine learning models, but also on how these models can (and cannot) be applied to real data.

2.2 Intellectual Merit

This effort will lead to advances in research areas of both machine learning and cybersecurity. Machine learning research is continually in search of application domains where new models and algorithms can be applied. This nexus will offer opportunities for students to learn about state of the art machine learning techniques, in the context of security problems. Security research, on the other hand, can benefit from applications of cutting-edge machine learning models to an array of problems, such as in anomaly detection. There is significant opportunity for new advances that require both domain knowledge in security and intimate familiarity with machine learning fundamentals. The intellectual merit in this proposal concerns three themes: (1) privacy-preserving dataset acquisition and generation; (2) automated course content generation for personalized teaching; (3) evaluation of pedagogy and educational outcomes. Each of these research themes has corresponding research questions, as described in detail in Section 4.

2.3 Broader Impacts

Curriculum development. The development of a curriculum at the intersection of machine learning and cybersecurity will open up new research areas that require domain expertise in both areas. For example, one area where a significant gap exists across both domains involves learning representations of timeseries data for outlier detection. In the case of network traffic traces, questions arise about sampling, transformations, aggregation, and how those operations affect deployability and model accuracy. Solving these problems requires a deep understanding of *both* machine learning fundamentals *and* cybersecurity domain knowledge, which our planned curriculum aims to develop.

Outreach through online courses. The project achieves a broader impact through the development of a course sequence that can be offered to a broad population in a variety of formats, and also uses machine learning to *tailor* that online content to individual students. The project will advance machine learning pedagogy through the development of new course material along four axes: (1) machine learning fundamentals; (2) fundamentals of networking and cybersecurity; (3) conceptual applications of machine learning techniques to practical problems in cybersecurity; (4) the instantiation of these concepts through real-world programs and tools, applied to real-world (and realistic) datasets. We aim to modularize the course material to reach a broad audience. Course modules and resources will be small and interchangeable to allow material to be targeted to different audiences as appropriate.

Integration with broadening participation efforts. The PIs have extensive engagement with outreach and broadening participation efforts at the University of Chicago; the PIs will also integrate their efforts and activities in this project with these outreach efforts. This summer, PI Feamster is teaching a course to 20 Chicago Public School (CPS) students on "Applications of Machine Learning to Network Security", through the University of Chicago's Upward Bound College Readiness Program. These students attend more than

ten different CPS public schools; about 80% of the students are from underrepresented minorities.

3 Project Outcomes

The project's outcomes will include: (1) curriculum development, including a four-course sequence whose curriculum and materials we will open source; (2) formative research with companies applying machine learning to cybersecurity to identify critical skills currently lacking in ML-cybersecurity professionals and develop case studies that allow students to develop skills and knowledge through hands-on assignments that provide exposure to real-world considerations; (3) toolkits for applying machine learning to real-world datasets—bridging the divide between theory, concepts, and use cases; (4) augmented teaching materials, whereby teaching materials are customized through human perception. We expand on each item below.

We will be releasing as open-source all curricula and course materials described in the subsections that follow. This will include slides, assignments, lab exercises, and (when applicable) lecture videos.

3.1 Curriculum Development

Curriculum: Re-skilling the security workforce, online. Towards this goal, we involve the following curricular elements, delivered through a combination of online videos, coursework, programming assignments, and capstone projects. All courses will integrate hands-on programming exercises from use cases inspired by the real world, where students are placed in a simulated scenario having to analyze data and make AI-assisted decisions related to cybersecurity. These use cases will encompass both those already encountered by the PIs in their work in the area, as well as identified in the formative studies we will conduct (Section 3.2). Courses will include the following:

- **Foundations of Machine Learning and Data Science**. A course in machine learning fundamentals, with applications to security. This course will offer an introduction to the data science pipeline, and teach fundamental building blocks, from data ingestion and feature engineering to machine learning model selection.
- **Data-Driven Network and Computer Security**. A system-oriented security course, with grounding in fundamentals that we expect our audience will be familiar with, pivoting towards more data-driven and oriented concepts. For example, we aim to bridge the divide between familiar concepts such as signature-oriented anomaly detection to statistical anomaly detection. The course will also provide training in the underlying mechanics of machine learning as applied to practical problems in cybersecurity, akin to an online course that Prof. Feamster has already developed for undergraduates.
- **Adversarial Machine Learning**. A course involving evasion and adversarial behavior in machine learning for security, including considerations of adversarial attempts to evade detection, pollution attacks on classifiers, backdoors in neural networks, and related topics in adversarial ML.
- **Ethics, Fairness, Responsibility, and Transparency in Data-Driven Cybersecurity**. The course will engage students in a series of programming assignments and case studies to expose them to ethical considerations associated with automated decision-making in the context of security.

The PIs have extensive experience teaching these proposed topics in other formats, and for other audiences, including in an online setting. Prof. Yuxin Chen teaches Statistical Foundations of Machine Learning; Prof. Nick Feamster teaches both Computer Security and Computer Networking (including network security); Prof. Blase Ur teaches Computer Security and Ethics, Fairness, Responsibility and Privacy in Data Science. Having taught multiple popular massive open online courses (MOOCs), Prof. Nick Feamster brings particularly extensive experience in developing engaging content for online settings.

3.2 Formative Research to Elicit Desired Skills and Develop Case Studies

Consistent with the proposal's core goal of developing new methods and curricula for re-skilling mid-career security professionals in cybersecurity applications reliant on AI and data, we aim to deeply engage American companies in our curriculum development process. Through semi-structured interviews of managers and technologists at both large and small companies and organizations, we will identify the following three classes of information that will enable our curriculum design to respond to current and future needs in industry in the United States. In addition to their expertise in AI and cybersecurity, all three PIs have extensive experience in conducting robust and reliable human-subjects research that incorporates best practices from the human-computer interaction (HCI) sub-field. We expect to interview professionals from at

least ten United States companies and organizations.

Understanding foundational knowledge needs. We will identify foundational knowledge in AI and data-driven methods that mid-career security professionals currently lack. Most of the target learners will have studied cybersecurity prior to the widespread deployment of data-driven-techniques. We will use the lessons learned to pinpoint both emphasis areas and potential gaps in our "Foundations of Machine Learning and Data Science" class. Because interview participants may not be experts in foundational AI, ML, and Data Science techniques, we will design our protocol to uncover both "known unknowns" and "unknown unknowns" in assessments of colleagues' skills. We will do so by asking about specific techniques not just by name, but by purpose.

Identification of application areas. We will elicit real-world challenges in data-driven cybersecurity. We will ask about the participant's attempted and desired applications of AI, ML, and data-driven methods to specific cybersecurity problems within their organization. We will use specific cybersecurity applications and workflows participants discuss as part of the "Data-Driven Network and Computer Security" course, augmenting the cybersecurity applications the co-PIs have researched or seen discussed in the literature. We will also use findings of the practical nuances and societal implications of applying these methods in "Ethics, Fairness, Responsibility, and Transparency in Data-Driven Cybersecurity." For instance, a nuanced transparency consideration for bot detection is that while greater transparency could increase trust in the system, yet also help attackers defeat the system. Similarly, a fairness consideration for ML detection of fake account registrations is that false-positive rates may differ across cultural groups or regions of the U.S.

Case studies. We will use this process to create concrete case studies that we will further develop and integrate into the hands-on course materials (Section 3.3) we will be open-sourcing. The case studies will: (1) reflect real-world cybersecurity applications; (2) involve state-of-the-art AI and ML techniques; and (3) highlight real-world challenges in deployability and ethics. Consider the following example. Fraudulent authentication attempts are a major cause of data breaches. To combat those attempts, companies can collect vast auxiliary data about a web authentication attempt, including the browser user agent, timing data, client IP address, latency, keystrokes, stored cookies, and more. Thus, data-driven methods are a natural fit. Companies like Microsoft are exploring how best to use this data to block fraudulent log-in attempts (Goal 1). Modeling this auxiliary data is challenging for myriad reasons, including that some data may be missing/spoofed, data like user-agent strings requires careful pre-processing, and only a small fraction of log-in attempts will be labeled as fraudulent or not. Building a model by naively running an off-the-shelf classifier will therefore be insufficient, enabling the case study to provide a scaffolded introduction to modern ML and data science techniques (Goal 2). It also introduces crucial ethical considerations (Goal 3). The classifier will run as part of the log-in process, so classification latency at scale is of the utmost importance. Furthermore, false positives will incur human and monetary costs of calls to a Security Help Desk and the potential loss of frustrated customers, which the case study will cover. Finally, some seemingly suspicious data (e.g., authentication attempts for different accounts from the same device) might be caused by legitimate sharing of devices, which is associated with users' socio-economic status, thus raising fairness concerns.

3.3 Toolkits and Datasets

Toolkits for applying machine learning to real-world datasets. As part of these courses and instructional materials, we will develop toolkits that give students experience in applying a variety of machine learning concepts to real-world datasets in security. Each course lecture above will be accompanied by hands-on material—likely in the form of a Jupyter notebook—that will provide students with hands-on experience with each concept in the courses. While this material will include the case studies developed in Section 3.2, they will be substantially augmented with examples from the literature on data-driven cybersecurity, including the PIs' own experience using such techniques in research and in practice.

For example, in the course on foundations of machine learning and data science, topics will include fundamental machine learning concepts such as linear regression, logistic regression, tree-based models, and so forth. Each course module will be accompanied with a Jupyter notebook that allows students to gain hands-on experience with the material. PI Feamster has developed a preliminary version of some of these notebooks and made them publicly available, at `https://github.com/noise-lab/ml-networking/`.

Subsequent courses will also incorporate and include similar notebooks.

Datasets for hands-on training. An important aspect of providing students with hands-on training is not only the notebooks that we provide them but also real-world datasets on which to apply the concepts that they learn, through conceptual material and practical programming. In particular, a critical aspect of applying machine learning techniques to network security problems involves a firm grasp of the steps in a data science pipeline that *precede* the model training itself, including data cleaning and validation, imputing missing values, checking the dataset for systematic measurement errors, and so forth. Students thus need access to real-world datasets, in the form of network traces (e.g., packet traces, IPFIX logs). Unfortunately, such real-world datasets are often considered proprietary, particularly when concerning security incidents. In particular, convincing industry partners to share these types of datasets can be prohibitive, as these datasets often contain private information, such as traffic volumes on different parts of a private network, mixes of application traffic (e.g., how much traffic on a given network may be from a particular application such a streaming video, sources and nature of attack traffic), and sensitive information such as Domain Name System (DNS) queries and responses, which can in many cases reveal sensitive information such as the websites that a user is visiting, or the devices that a user owns (or is using).

Towards providing students with real-world and realistic datasets for these hands-on activities, our outcomes will involve both *acquiring* and *generating* datasets that capture real-world cybersecurity settings.

- **Dataset acquisition.** For dataset acquisition, we will rely on relationships with the University Information Technology Services (ITS), industry partners (where possible and appropriate), and our ability to generate real-world datasets in the Internet of Things Laboratory at the Center for Data and Computing (CDAC) at the University of Chicago [30]. At the University of Chicago, the network collection infrastructure that we have developed in partnership with ITS will provide us access to network traffic data for network-connected building infrastructure across campus, providing a data source for allowing students to detect a wide range of outlier events in real-world environments. The IoT Lab allows us to collect data from real devices in controlled environments, allowing us to create datasets for synthetic anomalies, as well as labeled datasets that correspond to various activities that correlate with security events (e.g., network intrusion, physical intrusion).
- **Dataset generation.** An alternative to acquiring datasets is generating them. A challenge with dataset generation, of course, is doing so in a way that preserves the realistic properties of a real-world dataset. Fortunately, recent techniques in Generative Adversarial Networks (GAN) are beginning to enable this type of functionality [18]. These emerging technologies enable organizations to share *models* that can generate traffic according to specific distributions, where the models themselves are trained on real-world datasets but the resulting generated datasets do not contain private or proprietary information. We intend to build on existing research in this area to generate representative traffic models that can be included in educational toolkits for training students using synthetic datasets that nonetheless share properties from real-world networks.

3.4 Augmented teaching materials, customized to individual students

A clear advantage of online teaching is that it can easily be scaled up to a massive audience for broader participation. Once we develop the course materials (including the online videos, assignments, toolkits, and datasets described above) covering the essential topics, we will then assemble them in a progressive, optimized sequence, bearing in mind of the mid-career security professionals as the target audience. While it is possible to commit to a fixed course sequence upfront, we will make the default sequence of course modules configurable, allowing certain levels of curriculum customization for individual students, both prior to and during their enrollment of the program. This is especially convenient in the context of online teaching: in contrast to traditional hour-long lectures, the highly-customizable format allows us to further optimize the students' learning experience with automated, fine-grained adjustments of online curriculum.

Concretely, we will first identify the collection of course materials as either mandatory (e.g., ML models as fundamental building blocks for the subsequent security applications) or optional (e.g., alternative analytical tools for the analysis of ML models, certain supplemental case studies in security, and historical anecdotes). To maintain a basic structure of the default curriculum, we will manually organize and fix the mandatory course modules, while keeping a maximal flexibility in the presence and ordering of the optional course materials. We will build upon recent advances in *intelligent tutoring systems* [2, 17] and machine teaching [39] to automate the augmentation of the teaching materials *online*. In particular, co-PI Chen

has developed principled and practical machine teaching approaches for a variety of real-world teaching applications, including teaching with automated explanations [6,20], adaptive teaching for forgetful learners [16]. We will leverage and extend these techniques, and adaptively adjust the course materials as they progress through the course. Furthermore, by analyzing what activities worked for different types of students in the past, we will apply this knowledge to instruct new students.

4 Research Questions

Here, we lay out several fundamental research questions we aim to investigate in data-driven cybersecurity education. These research questions span across the online curriculum design process, ranging from: (1) the acquisition and generation of customized datasets for the course modules, (2) the semi-automated design and optimization of personalized online curriculum, to (3) the systematic evaluation of the course materials developed in the project.

4.1 Dataset Acquisition and Generation

A critical challenge in data-driven cybersecurity education is the lack of proper datasets. As described in Section 3, a significant outcome of our work will involve the production of datasets that can be used for courses in cybersecurity education. The application of machine learning to cybersecurity often involves automated detection and classification problems, including the detection of denial of service attacks, botnets, malware, spam, phishing, compromised devices, and so forth. In many cases, the detection problem relies on access to some type of labeled *network traffic trace*. The most raw form of this dataset typically comprises packet captures (sometimes called "pcaps").

Unfortunately, while these datasets are absolutely critical for network monitoring and security, they are also often extremely difficult to release to the public—they are effectively a wiretap, revealing the complete set of activities that took place on any given network. Because a significant fraction of network traffic still remains unencrypted, this data can contain extremely sensitive information, but even the metadata, such as traffic volumes and Internet destinations, can reveal sensitive information. Beyond the sensitive nature of packet capture data itself, assigning *labels* to this data is also challenging—unlike conventional datasets in machine learning (e.g., ImageNet), the labeling task is not amenable to crowdsourcing. As such, datasets containing labeled packet captures of attacks are exceedingly rare.

Generation of high fidelity synthetic network traces. An important aspect of our research will thus involve developing ways to *generate* synthetic traffic traces that contain representative types of network activity (including attacks) without publishing sensitive information. Towards this end, we will investigate the application of generative models for creating synthetic datasets in the cybersecurity domain, tailored for different course modules. Promising techniques in this area include the application of *Generative Adversarial Networks*, as described in Section 3.

Privacy-preserving transformations. Another potential approach may involve acquiring and suitably anonymizing network data from external sources (e.g., companies, the university network). Although in general, anonymization can be a difficult endeavor, many cybersecurity models and algorithms depend only on access to various derived statistics, including various measures of traffic volume over time, packet header information, and so forth [34]. In many cases, it may be possible to suitably sanitize or transform packet captures in ways that preserve the ability to perform machine learning tasks such as outlier or anomaly detection without exposing private information. One potential approach to this problem that we are exploring is through the use of dimensionality reduction techniques (e.g., principal components analysis), combined with additive noise on the resulting transformation. The noise has the effect of obscuring the transformation in such a way that makes it difficult to recover the original data before transformation, without distorting the high-energy components of the traces that are most useful for applications of machine learning, such as outlier detection (as can be used in the detection of various types of attacks, including denial of service and intrusions).

4.2 Optimizing the Course Modules Online for Personalized Teaching

Given the collection of course modules, a natural research question is how to properly assemble them in an online curriculum to optimize the students' learning experience. The sequence of course modules should be tailored to the abilities and progress of each student, in order to keep them engaged and avoid confusion and dropout. Unfortunately, although the demand for massive open online courses (MOOC)

has spurred the recent development of intelligent tutoring systems, the lack of educational data in the intersection between machine learning and cybersecurity makes it challenging to apply existing data-driven automated tutoring systems which are often data-hungry. It thus remains a central challenge in machine learning and educational research, to construct principled approaches for optimizing the sequence of course materials online for personalized teaching.

Establishing computational models of human learning in data-driven curriculum design. Student modeling for optimizing human learning is a complex task that has a rich history in psychometrics, machine learning, and educational data mining. For example, it has been well established in psychology that in the context of human learning, students' knowledge decays rapidly without reconsolidation [12]. Unfortunately, traditional online teaching systems do not employ these findings in practice, and thus may lead to undesirable outcomes such as rapid forgetting [27]. As another example, Bayesian Knowledge Tracing [9] is an algorithm used in many intelligent tutoring systems to model each student's mastery of the knowledge being tutored, and how the knowledge evolves when the student is presented with new course materials. To ensure that students can grasp and retain the new concepts learned over the online sessions, we aim to explore the following two problems: (i) how should we model and infer students' learning dynamics in an *online* fashion, and (ii) which metrics should be considered when optimizing student learning?

As concrete steps towards addressing the above questions, we consider the *spacing effect* [24, 26] and the *lagging effect* of human learning [29] from the cognitive science and psychology literature, with the goal to maximize the students' learning *progress* and *retention*. The spacing and lagging effects suggest that spaced repetition of the course materials, as well as inter-study intervals between study sessions leads to better long-term retention. These results have inspired many computational models of human learning, including the half-life regression (HLR) model as was deployed in Duolingo for language learning [25]. We will adopt the trainable HLR model to model the students' memory [22, 23, 32, 33], and apply data-driven approaches on existing educational datasets from the MOOC platforms to extract a prior model on the learning dynamics for human students.

Applying machine teaching to optimize the sequence of course modules online. Once establishing a prior model for human learning, one needs to calibrate the learner's model online as data flows, and adaptively adjust the subsequent course materials to optimize students' performance. In our online teaching setup, such adaptation is done through interacting with the human students via problem assignments, quizzes, and mini course projects. Equipped with the students' feedback on their learning performance, we can cast online curriculum design as an adaptive optimization problem, where the goal is to maximize the learning outcome subject to constraints on the bandwidth of human attention. Such problem has been formally studied in the literature as *adaptive machine teaching* [5, 7, 15, 19, 28], where the adaptive teacher (i.e., online course scheduler), after delivering each course module, uses the feedback from the student to guide the teaching policy by integrating the existing teaching outcomes into the optimization process.

To facilitate automation of the scheduling of the course modules, we will leverage existing tools in machine teaching, and investigate the problems inherent to adaptive online teaching, in terms of fairness (to inadvertent discrimination), privacy, and robustness to adversaries trying to game the tutoring system. In prior work, Co-PI Chen [16] investigated the problem of automated tutoring from the perspective of discrete optimization and introduce a novel *algorithmic machine teaching* framework for teaching multiple concepts online. The algorithm allows the design of repetition schedules for different memory models, while being adaptive to the underlying learning dynamics of the student. We will utilize and extend these models to handle more complex target concepts, as well as deal with real-world privacy and fairness constraints.

4.3 Evaluation of Course Materials and Learning Objectives

Rigorous evaluation of our educational materials throughout the project will ensure the programs advance our learning objectives for mid-career security professionals and improve educational outcomes. While they are experienced teaching ML, data, and cybersecurity courses, none of the co-PIs is formally trained in pedagogy or program evaluation. To that end, we will leverage internal expertise from the Outlier Research & Evaluation team [31], a leading educational evaluation unit in UChicago's STEM Education program. The Outlier team includes scientists, teachers, and researchers who develop innovative approaches to evaluation in STEM and CS education. The Outlier team has a broad range of experience evaluating programs in schools, adult professional development settings, and online learning. Education, a leading research center

dedicated to improving mathematics and science education. The co-PIs will meet monthly with the Outlier team. throughout the grant timeline to achieve these goals.

The Outlier team will design and conduct formative and summative evaluations of the curriculum materials and student learning outcomes. Each online module or course will have learning objectives specific to mid-career security professionals that can be assessed through interviews, surveys, and online engagement metrics. The team will focus on three frameworks within the Outlier group's toolkit. First, they will leverage the Implementation Framework to organize and identify the innovations and goals of re-skilling mid-career professionals. specific component might involve measuring the number and types of deployment considerations referenced in the model-selection process during a given case study. Similarly, they will use the Capacity Framework to identify ways for sustaining the specific educational objectives (e.g., careful consideration of how attackers can bias inputs to data-driven models) when the learner has returned to their company. Finally, they will use the Sustainability Framework to steer toward fundamental lessons in data-driven cybersecurity that are agnostic to change in ML's state of the art. Outlier will establish data collection timelines, samples and data collection processes in collaboration with project Principal Investigators.

5 Effects of COVID-19 on Project

As our project involves the development of an online curriculum, we do not expect our project to be adversely affected or delayed by COVID-19, beyond the normal amounts of productivity reduction that have resulted from the inability of our team to meet together in person.

In fact, we view the current situation as an *opportunity* to pilot and deploy this new online curriculum, given that a large fraction of the population now relies on remote learning and education. Furthermore, a significant segment of the population is also re-skilling given the changing structure of today's economy.

With the continuing presence of the global pandemic, the effects on jobs are increasingly clear—some jobs have disappeared and certain jobs may never come back. This evolution of the economy will demand re-skilling more than ever. The curriculum we plan to develop is designed to be the first of its kind to provide individuals with the opportunity to gain practical, hands-on experience with machine learning in a growth sector of the economy. Towards this end, we plan to coordinate with our office of professional education to make the materials we develop broadly available, in a variety of modular formats.

6 Results from Prior NSF Support

Chen: Co-PI Chen has no prior NSF support.

Feamster: CPS: Medium: Detecting and Controlling Unwanted Data Flows in the Internet of Things *1739809* (PI: Feamster); 10/1/18–9/30/22, $875,000. This project develops technologies that ensure that IoT smart devices remain secure and protect user privacy. **Intellectual Merit:** This project advances the theory and practice of network traffic analysis and anomaly detection to secure IoT deployments. The project has resulted in publications at top-tier conferences [1, 3, 8, 10, 11, 13, 14, 21, 38], including studies of how home network traffic can expose users to new privacy risks, new methods for modeling privacy in smart homes, and a broader understanding of how devices collect private user data. They also include broader research on IoT security and privacy. **Broader Impacts:** The research has produced open-source software (IoT Inspector) to help smart home users better understand how devices collect and share private data and the largest labeled dataset of smart home device traffic with data from about 10,000 homes.

Ur: FMitF: Collaborative Research: User-Centered Verification and Repair of Trigger-Action Programs *1837120* (PI: Ur); 9/1/18–8/31/22, $999,998. This project develops formal methods that underpin data-driven user interfaces to help end-users more accurately express their intent in trigger-action programming (TAP). **Intellectual Merit:** We are developing novel methods that tightly couple formal methods with the design of user interfaces. We are developing graphical interfaces that elicit desired properties from users, automatically translating them to LTL and synthesizing programs [35]. We also developed a system that uses SAT-solving to synthesize potential programs based on traces of observed behaviors [36]. The project has already led to papers at UbiComp '20 (IMWUT) [36], ICSE '19 [35], and CHI '19 [4], as well as a CHI '20 poster [37]. **Broader Impacts:** Interfaces based on TAP are widely deployed (e.g., IFTTT, Microsoft Flow, Mozilla's Things Gateway). The research has produced open-source software (AutoTap, Trace2TAP [36]).

References

[1] G. Acar, D. Huang, F. Li, A. Narayanan, and N. Feamster. Web-based Attacks to Discover and Control Local IoT Devices. In *ACM SIGCOMM Workshop on Internet of Things Security and Privacy*, Budapest, Hungary, Aug. 2018.

[2] V. Aleven, E. A. McLaughlin, R. A. Glenn, and K. R. Koedinger. Instruction based on adaptive learning technologies. *Handbook of research on learning and instruction*, pages 522–560, 2016.

[3] N. Apthorpe, D. R. Yan Shvartzshnaider, and N. Feamster. Discovering Smart Home IoT Privacy Norms using Contextual Integrity. In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, Singapore, Oct. 2018.

[4] W. Brackenbury, A. Deora, J. Ritchey, J. Vallee, W. He, G. Wang, M. L. Littman, and B. Ur. How users interpret bugs in trigger-action programming. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 2019.

[5] M. Cakmak and M. Lopes. Algorithmic and human teaching of sequential decision tasks. In *AAAI*, 2012.

[6] Y. Chen, O. M. Aodha, S. Su, P. Perona, and Y. Yue. Near-optimal machine teaching via explanatory teaching sets. In *Proc. International Conference on Artificial Intelligence and Statistics (AISTATS)*, April 2018.

[7] Y. Chen, A. Singla, O. Mac Aodha, P. Perona, and Y. Yue. Understanding the role of adaptivity in machine teaching: The case of version space learners. In *Advances in Neural Information Processing Systems*, pages 1476–1486, 2018.

[8] G. Chu, N. Apthorpe, and N. Feamster. Security and Privacy Analyses of Internet of Things Children's Toys. *IEEE Internet of Things Journal (IOT-J)*, Nov. 2018.

[9] A. T. Corbett and J. R. Anderson. Knowledge tracing: Modeling the acquisition of procedural knowledge. *User modeling and user-adapted interaction*, 1994.

[10] T. Datta, N. Apthorpe, and N. Feamster. A Developer-Friendly Library for Smart Home IoT Privacy-Preserving Traffic Obfuscation. In *ACM SIGCOMM Workshop on Internet of Things Security and Privacy*, Budapest, Hungary, Aug. 2018.

[11] R. Doshi, N. Apthorpe, and N. Feamster. Machine Learning DDoS Detection for Consumer Internet of Things Devices. In *IEEE Security and Privacy Deep Learning and Security Workshop (DLS)*, San Francisco, CA, May 2018.

[12] H. Ebbinghaus. *Über das gedächtnis: untersuchungen zur experimentellen psychologie*. Duncker & Humblot, 1885.

[13] N. Feamster. Mitigating the Increasing Risks of an Insecure Internet of Things. *Colorado Tech Law Journal (CTLJ)*, Aug. 2018.

[14] M. X. Ferreira, D. Y. Huang, T. Chattopadhyay, N. Feamster, and S. M. Weinberg. Selling a Single Item with Negative Externalities. In *International World Wide Web Conference (WWW)*, San Francisco, CA, May 2019.

[15] L. Haug, S. Tschiatschek, and A. Singla. Teaching inverse reinforcement learners via features and demonstrations. In *Advances in Neural Information Processing Systems*, pages 8464–8473, 2018.

[16] A. Hunziker, Y. Chen, O. Mac Aodha, M. G. Rodriguez, A. Krause, P. Perona, Y. Yue, and A. Singla. Teaching multiple concepts to a forgetful learner. In *Advances in Neural Information Processing Systems*, pages 4050–4060, 2019.

[17] K. R. Koedinger, E. Brunskill, R. S. Baker, E. A. McLaughlin, and J. Stamper. New potentials for data-driven intelligent tutoring system development and optimization. *AI Magazine*, 34(3):27–41, 2013.

[18] Z. Lin, A. Jain, C. Wang, G. Fanti, and V. Sekar. Generating High-fidelity, Synthetic Time Series Datasets with DoppelGANger. *arXiv preprint arXiv:1909.13403*, 2019.

[19] W. Liu, B. Dai, A. Humayun, C. Tay, C. Yu, L. B. Smith, J. M. Rehg, and L. Song. Iterative machine teaching. In *ICML*, pages 2149–2158, 2017.

[20] O. Mac Aodha, S. Su, Y. Chen, P. Perona, and Y. Yue. Teaching categories to human learners with visual explanations. In *Proceedings of CVPR*, 2018.

[21] H. M. Moghaddam, G. Acar, B. Burgess, A. Mathur, D. Y. Huang, N. Feamster, E. W. Felten, P. Mittal, and A. Narayanan. Watching You Watch: The Tracking Ecosystem of Over-the-Top TV Streaming Devices. In *ACM Conference on Computer and Communications Security (CCS)*, Nov. 2019.

[22] P. I. Pavlik Jr and J. R. Anderson. Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect. *Cognitive Science*, 29(4):559–586, 2005.

[23] D. C. Rubin and A. E. Wenzel. One hundred years of forgetting: A quantitative description of retention. *Psychological review*, 1996.

[24] G. Rubin-Rabson. Studies in the psychology of memorizing piano music: Ii. a comparison of massed and distributed practice. *Journal of Educational Psychology*, 31(4):270, 1940.

[25] B. Settles and B. Meeder. A trainable spaced repetition model for language learning. In *ACL*, volume 1, pages 1848–1858, 2016.

[26] A. L. Simmons. Distributed practice and procedural memory consolidation in musicians' skill learning. *Journal of Research in Music Education*, 59(4):357–368, 2012.

[27] B. Subirana, A. Bagiati, and S. Sarma. On the forgetting of college academice: at "ebbinghaus speed"? Technical report, Center for Brains, Minds and Machines (CBMM), 2017.

[28] S. Tschiatschek, A. Ghosh, L. Haug, R. Devidze, and A. Singla. Learner-aware teaching: Inverse reinforcement learning with preferences and constraints. In *Advances in Neural Information Processing Systems*, 2019.

[29] O. J. Tzeng. Stimulus meaningfulness, encoding variability, and the spacing effect. *Journal of Experimental Psychology*, 99(2):162–166, 1973.

[30] University of Chicago Center for Data and Computing (CDAC) Internet of Things Laboratory. `https://cdac.uchicago.edu/iot-lab/`.

[31] University of Chicago STEM Education. Outlier research & evaluation, 2020. `http://outlier.uchicago.edu/`.

[32] M. M. Walsh, K. A. Gluck, G. Gunzelmann, T. Jastrzembski, M. Krusmark, J. I. Myung, M. A. Pitt, and R. Zhou. Mechanisms underlying the spacing effect in learning: A comparison of three computational models. *Journal of Experimental Psychology: General*, 147(9):1325, 2018.

[33] T. D. Wickens. Measuring the time course of retention. *Evolution, Progress, and Reflections on the 30th Anniversary of the Atkinson-shiffrin Model*, 1999.

[34] K. Yang, S. Kpotufe, and N. Feamster. A Comparative Study of Network Traffic Representations for Novelty Detection. *arXiv preprint arXiv:2006.16993*, 2020.

[35] L. Zhang, W. He, J. Martinez, N. Brackenbury, S. Lu, and B. Ur. Autotap: synthesizing and repairing trigger-action programs using ltl properties. In *Proceedings of the 41st International Conference on Software Engineering*. IEEE Press, 2019.

[36] L. Zhang, W. He, O. Morkved, V. Zhao, M. L. Littman, S. Lu, and B. Ur. Trace2TAP: Synthesizing trigger-action programs from traces of behavior. In *Proceedings of the of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT / UbiComp)*. ACM, 2020.

[37] V. Zhao, L. Zhang, B. Wang, S. Lu, and B. Ur. Visualizing differences to improve end-user understanding of trigger-action programs. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, 2020.

[38] S. Zheng, N. Apthorpe, M. Chetty, and N. Feamster. User Perceptions of Smart Home IoT Privacy. In *ACM Conference on Computer Supported Cooperative Work (CSCW)*, Jersey City, NJ, Nov. 2018.

[39] X. Zhu. Machine teaching: An inverse problem to machine learning and an approach toward optimal education. In *AAAI*, pages 4083–4087, 2015.