An Effort to Democratize Networking Research in the Era of AI/ML

Arpit Gupta
UC Santa Barbara
Santa Barbara, CA

Chris Mac-Stoker
NIKSUN Inc.
Princeton, NJ

Walter Willinger
NIKSUN Inc.
Princeton, NJ

ABSTRACT

A growing concern within today's networking community is that with the proliferation of Artificial Intelligence/Machine Learning (AI/ML) techniques, a lack of access to real-world production networks is putting academic researchers at a significant disadvantage. Indeed, compared to a select few research groups in industry that can leverage access to their global-scale production networks in their data-driven efforts to develop and evaluate learning models, academic researchers not only struggle to get their hands on real-world data sets but find it almost impossible to adequately train and assess their learning models under realistic conditions.

In this paper, we argue that when appropriately instrumented and properly managed, enterprise networks in the form of university or campus networks can serve as real-world production networks and can, because of their ubiquity, help create a more level playing field for academic researchers. Their various limitations notwithstanding, as real-world production networks, such enterprise networks can *(i)* serve as unique sources for some of the rich data that will enable these researchers to influence or advance the current state-of-the-art in AI/ML for networking and *(ii)* also function as much-needed test beds where newly developed AI/ML-based tools can be evaluated or "road-tested" prior to their actual deployment in the production network. We discuss new research challenges that arise from this proposed dual role of campus networks and comment on the opportunities our proposal affords for both academic and industry researchers to benefit from the advantages and limitations of their respective production environments in their common quest to advance the development and evaluation of AI/ML-based tools to the point where they can be deployed in practice.

1 INTRODUCTION

The networking area is currently facing a growing "digital divide" between those researchers that have access to plenty of data and those that suffer from a basic lack of data. On the one hand, we are experiencing the emergence of a select few research groups in industry that can leverage their global-scale production networks to *(i)* obtain the sort of data (typically proprietary in nature) that is needed for their data-driven efforts to develop their latest AI-/ML-based learning models (e.g., to perform real-time network automation tasks) and *(ii)* evaluate the resulting models under real-world conditions prior to their full-fledged deployment in a production network. On the other hand, unless they engage in close collaborations with their "data-rich" colleagues in industry (and after signing an NDA), academic networking researchers typically lack access to any type of real-world production networks and are thus seriously disadvantaged when it comes to obtaining adequate data for developing their learning models and evaluating them "in the wild".

At the same time, a closer look at the areas where AI/ML applications have generated excitement in recent years (e.g., computer vision, autonomous vehicles/driving) shows that their successes depended critically on having access to an abundance of publicly available (labeled) data. For example, the field of computer vision has benefited immensely from IMAGENET, a community-led effort that has resulted in the creation of a large database of hand-annotated images which in turn revolutionized the field of visual object recognition research [17]. By providing researchers with an opportunity to focus on developing new and better learning algorithms and spend less or no time on data collection and labeling, IMAGENET has fuelled as well as democratized research in the area of computer vision. Similarly, research on AI/ML applications for autonomous vehicles took off with the introduction of the LIDAR (short for light detection and ranging) and the recent industry-led efforts to open-source entire collections of

Arpit Gupta, Chris Mac-Stoker, and Walter Willinger

curated, cleaned and labelled datasets (e.g., Argo AI's ARGO-VERSE [16] or Adaptiv's NUSCENES [15]).[1] These datasets contain hundreds of thousands of LIDAR sweeps, millions of camera images and data from other sensors such as Sonars, GPS, radar, and accelerometers and are intended *"to enable researchers to study challenging urban driving situations using the full sensor suit of real self-driving cars"* [15]. This data can be used at once by these researchers in their data-driven efforts to develop and evaluate novel learning models for autonomous vehicles/driving; they no longer have to do the driving themselves to collect the necessary data.

In stark contrast, the field of networking has had no equivalent of IMAGENET and has not had the benefits of the analogue of a LIDAR or of comprehensive collections of datasets such as ARGOVERSE or NUSCENES. Instead, for many problems, researchers typically have to start from scratch, first conjecturing what the critical features for the problem at hand might be, then developing the tools or systems for collecting the necessary data, and finally extracting the identified features from the obtained data. As a result, these researchers often spend more time on designing and running experiments to collect the data needed for extracting the features required for the development of their learning models, with essentially no opportunities to evaluate the resulting models in settings that have any resemblance to a real-world production network (see also [27] and references therein).

To correct this situation and also achieve the goal of democratizing networking research in the era of AI/ML, we propose in this paper that academic network researchers start to consider and treat their university or campus networks as real-world production networks.[2] When adequately instrumented with state-of-the-art network monitoring and data collection capabilities and a set of the latest programmable devices, such production networks could be operated and managed in ways that go beyond their expected or traditional role (e.g., delivering IT services and resources to the larger university community). In particular, they could as well serve as the source of the data that populates an IMAGENET-like database or results in a ARGOVERSE- or NUSCENES-type comprehensive data store and could also function at once as a testbed where these researchers could evaluate and "road-test" their learning solutions under real-world conditions.

Figure 1 illustrates our proposed use of properly instrumented university or campus networks by academic researchers for developing and evaluating their learning models and provides an overview of the paper. In particular, the networks' proposed use as unique sources of rich network data
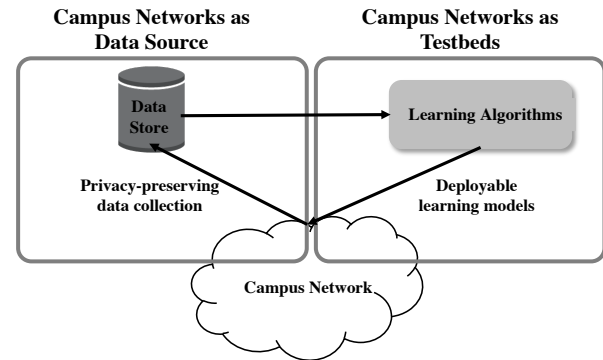

Figure 1: Paper in a nutshell: Treating a campus network as both a data source (*i.e.*, Section 3) and a testbed (*i.e.*, Section 4) for networking research in the era of AI/ML.

has to be complemented with renewed attention to the topic of data privacy (Section 3). Moreover, their intended use as premier testbeds requires novel ideas for breaking down existing barriers to road-testing learning models (Section 4). We also address some of the main limitations of the proposed approach and discuss in Section 5 how they can, in fact, be leveraged to benefit networking researchers in academia and industry alike so as to close the mentioned digital divide.

2 KEY IMPEDIMENTS

Constant pressure to operate and manage their networks in ever-more efficient and cost-effective ways has resulted in network operators looking towards network automation as a viable solution to their problems. At the same time, network automation depends critically on the ability to develop purposefully-designed learning algorithms to perform specific tasks by themselves, without any assistance from the operator. Take the case of a typical enterprise where automating a given network management task generally involves performing a particular action upon detecting a certain network event. For example, the network event in question could be a DDoS attack in the form of a DNS amplification attack on the enterprise and the corresponding action could be "drop attack traffic on ingress if confidence in detection is at least 90%."

In general, such automation tasks require three basic activities: *(i)* sensing the network to obtain the data required for the automatic inference of the network event in question, *(ii)* using the acquired data as input to a learning model that alerts on detecting the network event, and *(iii)* upon detection, performing the required mitigative actions. Moreover, the allocation of compute resources that are available in the network for performing any of these activities for a given task (e.g., data plane, control plane, cloud) will depend on how fast and with what accuracy that task has to be performed.

---

[1]It is not clear what exactly incentivized these companies to open-source previously proprietary data, and why now and not earlier.

[2]Existing research and education networks can also be viewed as real-world production networks as long as they carry a significant amount of Internet-facing traffic.

For example, while modern data plane technologies are critical for enabling the real-time detection and mitigation of task-specific network events, they are currently not capable of supporting this capability at scale; *i.e.*, executing hundreds or thousands of such tasks concurrently and in real time.

At the same time, the described combination of basic activities makes network automation a very promising but challenging research area for both leveraging recent advances in AI/ML (for the development of automated inference techniques to accomplish activity *(ii)*) and experimenting with the latest programmable data plane technologies (for supporting flexible network measurements and applying programmatic controls in support of activity *(i)* and *(iii)*, resp.). It is also an area that highlights at once the main obstacles that have stymied much of academic research on network automation to date.

The data problem. Without the benefits of analogues of IMAGENET or ARGOVERSE and NUSCENES, networking researches have had no choice but obtain their desired data in a "bottom-up" manner; that is, collecting much of the data in an ad-hoc fashion, one special-purpose measurement experiment at a time, often as an after-thought, with hardly any useful metadata and generally with little (if any) attention given to data-cleaning issues. As a result, publicly available network measurements in support of network automation tasks are rare, not necessarily representative, often a by-product of some other measurement activities, and typically of only very limited use for the task at hand. Even worse, labelled data that is key to applying some of the existing AI/ML techniques to network-specific problems is largely non-existent.

Note that this described bottom-up approach to data collection in networking is in stark contrast to the "top-down" approach used by today's researchers in areas such as computer vision. With IMAGENET at their disposal, these researchers can treat feature engineering as a first-class citizen. That is, they typically start by examining "ground truth" in the form of full-fledged images and then use all their expertise and imagination when deciding on the most promising features for a given task. Importantly, the subsequent steps (e.g., feature extraction and model training) and possible iterations of this process can take full advantage of the already available datasets – no new measurement experiments and/or data collection efforts are required.

The evaluation problem. As the success stories of AI/ML mount and AI/ML is increasingly used as an integral part of more and more real-word systems (e.g., for medical diagnosis, decision making in the legal system), the in-depth evaluation of newly developed learning models has become a top priority. For example, while in areas such as computer vision, researchers have began evaluating their learning algorithms with an eye towards bias or fairness [13, 19], in the area of autonomous vehicles/driving, the learning models are increasingly being evaluated with respect to their safety and robustness [14, 25]. In the networking area, the researchers' efforts to evaluate their learning models are already jeopardized by the amount of time they have to spend on running experiments and collecting data. More importantly, their validation efforts are largely meant to convince network operators who look for tools they can trust. In short, for those operators, what matters is an ability to explain or interpret learning algorithms and models. However, due to the much-maligned black-box nature of existing AI/ML techniques, today's network operators have a general distrust of learning models and are opposed to deploying untrustworthy tools in their production network, thus limiting the researchers' evaluations of such tools to testbeds that have little (if any) resemblance with the production network where these tools are supposedly automating tasks that used to be performed by a human operator.

3 CAMPUS NETWORKS AS DATA SOURCES

A look at research areas such as computer vision and autonomous vehicles/driving shows that the development of AI/ML-based learning models has benefited enormously from access to rich collections of datasets such as IMAGENET or ARGOVERSE/NUSCENES that effectively showcase the full power of feature engineering. In fact, the recent successes in applying AI/ML in these fields are proof of the power of the feature engineering-centric top-down approach that lets researchers focus more fully on the actual development and subsequent evaluation of their new learning models and largely relieves them from the burden of collecting the necessary data themselves.

However, while the IMAGENET or ARGOVERSE/NUSCENES analogies are helpful for comparing ongoing AI/ML-related efforts in networking to those in computer vision and autonomous vehicles, they quickly break when stretched too far. For example, the type of data that we propose to collect from a university's network is much richer and more diverse than, for example, the images in IMAGENET. However, the biggest differentiator derives from the data privacy aspect of most network data. In fact, an important aspect of our proposal is that the data collected from a campus network is only meant for "internal" use; that is, the data is only available to the researchers of that university and it not intended to be available to or shared with third parties (e.g., researchers from different universities).

One main goal of this proposal is to ensure that networking researchers, especially in academia, are afforded similar opportunities as their counterparts in areas such as computer vision or autonomous vehicle. To this end, we argue that the

networking research community has to come up with means for creating and populating new types of data stores that contain the most comprehensive collections of network data and that leverage the latest network monitoring technologies.

Proposal (Part 1): *To address the long-standing data problem that has plagued academic networking researchers, we propose that these researchers consider and treat their university or campus network as a real-world production network and view and use it to its fullest extent as a source of realistic network data of unprecedented quality and quantity.*

Why campus networks? As far as real-world enterprise networks are concerned, university or campus networks are unique in the sense that they often strike a balance between scale, complexity, and diversity. Typically small to moderate in size, they carry traffic for a sizable university population, and the traffic is real in that it is generated by a range of actual applications and services from a variety of different real-world users. These networks are operated and managed by a professional IT organization that includes network operators, engineers and security analysts. Among many other responsibilities, these IT professionals support a variety of research, academic and administrative systems, services, and resources; are in charge of operating the university's voice, (wired and WiFi) network, and computing infrastructures; deal with identity and access management and security; and are responsible for protecting the university's information and the privacy of its users.

As sources of potentially valuable intellectual property, as locations of various national laboratories, or as administrators of student applications and records, universities experience their fair share of cyber attacks. However, many of them are usually unaware of most of these attacks and only find out after a breach occurred [23]. University networks are also prone to network faults and outages and experience performance issues in their infrastructures or with their offered services (e.g., see [2]). In particular, there is a need to be able to pinpoint performance problems and notify the service or cloud provider(s) in case the root cause is not internal to the campus network.

Limitations. Without even attempting to put together a comprehensive list of limitations that the use of campus networks as data sources imposes, the following are some of the most obvious ones. For one, with respect to aspects such as scale (e.g., size, traffic volume, end user population), workload (e.g., type of applications or services) and economics (e.g., business model, peering relationships), university networks are clearly not comparable to the real-world production networks of large enterprises, large ISPs such as AT&T and Verizon, or global-scale content or cloud providers such as

Google, Microsoft, Facebook, Amazon, Alibaba, etc. In particular, campus networks are not meant to help with problems in the data center space or with different aspects of wireless networks (e.g., spectrum allocation).

As a result, the data collected from a typical university network will not be representative for any of these companies' production networks, the complexity of most network management tasks for university networks pales in comparison to what is needed for operating the domestic- or global-scale network infrastructures of these companies, and when it comes to the degree of heterogeneity of deployed hardware, installed software, available tools and manpower, most campus networks have to be rightly viewed as little more than toy examples when compared to the real-world production networks of the large Internet companies.

4 CAMPUS NETWORKS AS TESTBEDS

Given that making mistakes or wrong decisions in a production network has real-world consequences (e.g., disruption of critical services, dis-satisfied users, poorly performing applications, loss of revenues), network operators are extremely averse to deploying untested tools or experimental hardware.[3] At the same time, academic network researchers have in general only very limited options (e.g., simulations, emulations) when it comes to evaluating their newly developed AI/ML-based network automation tools. Importantly, they typically lack access to any real-world production network, and as a result, they have made little to no progress in the past with respect to convincing network operators of their tools' correctness, robustness, and safety or explaining to them how their tools' learning models work and why they output a certain result and not something else. A second goal of our proposal is to suggest a viable option for networking researchers to gain access to actual production networks.

Proposal (Part 2): *To address the ever-present evaluation problem that has vexed academic networking researcher in the past, we propose that these researchers should consider and treat their university or campus network as a real-world production network and view and use it to its fullest extent as a testbed for extensively evaluating their newly developed learning models and road-testing their newly designed AI/ML-based network automation tools.*

Why campus networks? In view of our stated objective of democratizing networking research in the era of AI/ML, the proposed use of university networks as real-world testbeds is especially fitting because these proposed production networks are ubiquitous (*i.e.*, every university has one); they can be instrumented with the necessary networking technologies (*i.e.*,

---

[3]Interestingly, the companies behind ARGOVERSE and NUSCENES do not discuss any road-map for integrating new tools developed by third-party researchers into their solutions.

these technologies are already available today); transforming them into testbeds can be achieved incrementally (e.g., use of configuration management tools such as Puppet [7]); and realizing them is well within the budgetary constraints of a typical university's networking research group and/or IT organization. In short, every academic networking researcher can in theory gain access to a real-world production network and start using it as data source and testbed.

Moreover, in theory, there is no reason why every university's IT organization could not operate its university's network so as to perform the proposed dual role; that is, function as traditional service provider (e.g., delivering IT services and resources to the larger university community) and, at the same time, operate the network for the benefits of the university's networking researchers in particular and its larger research community in general (e.g., as data source and real-world test bed). However, ensuring that the task of running the campus network "as a lab" to advance research and support innovation becomes part of any university's IT organization's mission statement will require university-wide efforts and novel incentives for the involved parties (e.g., students, faculty, IT organization, administration). We argue that with the proliferation of AI/ML, it is high time for universities to engage their various constituents in such efforts.

Limitations. In addition to the already mentioned limitations (see Section 3), the unavailability of modern programmable devices in production settings in existing campus networks is another limiting factor. A few research groups are currently experimenting with programmable devices, but we are not aware of any university IT organizations that are actively using or experimenting with them in their production networks.

Deploying a researcher's AI/ML-based tools in an university's production network requires production-quality software and clearly-defined guidelines that spell out the responsibilities for both the researcher and the IT organization. However, not only is the current academic environment ill-suited for researchers to write production-quality code, but a general lack of interactions and synergies between a university's networking research groups on the one hand and its IT organization on the other hand makes it difficult to even start entertaining the idea of some sort of support contract between the "vendor" (*i.e.* the researcher who wrote the tool's code) and the "customer" (*i.e.* the IT organization that uses the tool).

Just as our proposed use of campus networks as real-world testbeds (and as sources for data from actual production networks) has its limitations, so do alternative approaches. For example, attempts to incentivize the large commercial service providers to allow academic researchers limited and carefully-controlled use of their production networks as data sources or testbeds have been largely futile in the past. Other efforts that start more testbed-oriented and have a more production-like

setting as their goal (e.g., Massachusetts Open Cloud (MOC) project [3] run the risk of failing to attract sufficient commercial traffic onto their platform. In view of such alternatives, we argue that the common goal of advancing and broadening networking research in the era of AI/ML is best served by *also* exploring the use of the ubiquitous real-world production networks owned and run by universities.

5 NEW OPPORTUNITIES

The listed limitations in Sections 3 and 4 may be viewed as strong ammunition against using campus networks as data sources and/or real-world testbeds for advancing a viable research agenda in the area of AI/ML-based networking. However, we argue that precisely because of these limitations, leveraging these real-world toy production networks creates unique opportunities for both academic and industry researchers and lets them both benefit from the advantages as well as limitations of their respective production settings.

From network monitoring to data stores. Because the overall traffic that a campus network exchanges with its upstream provider(s) is typically in the 10-20 Gbps range, these networks are ideal for experimenting with the latest network monitoring technologies in ways that would be impractical for the large commercial service provider networks that have to handle way larger traffic volumes on a daily basis. In particular, we propose to instrument campus networks with monitoring solutions that can perform enterprise-wide, continuous, lossless, full packet capture at scale; *i.e.*, every packet that enters or leaves the enterprise is collected (packets that stay inside the enterprise can also be part of this solution), with full payload, with no sampling, and at link speeds of up to 100 Gbps or higher. This technology already exists [4], a number of enterprises in the private and public sectors have already adopted such commercial solutions, and a typical campus network (e.g., a 10 Gbps upstream connection, data storage requirements of the order of a week) can deploy this technology today for a few $100K.[4]

A critical feature of these latest network monitoring solutions is that they also create and populate a data store that contains not only all the raw packet-level data but also complementary data from other available sensors or sources (e.g., server logs, firewall rules, configuration files, events) as well as an extensive set of "on-the-fly" generated metadata. Moreover, all the stored data is not only cleaned, curated, time-synchronized and (where possible) labelled, but also linked and indexed to provide fast and flexible search capabilities. Comprising a single platform for collecting, storing, indexing, mining, and visualizing network data, a university network's data store that results from this proposal becomes the single

---

[4]This cost increases proportionally with the size and number of the upstream links and the duration of data retention.

source of all campus network-related data. That is, everything seen "on the wire" is collected and stored and augmented with all different types of complementary network-related information and metadata.

Also note that because of the small-moderate size of typical campus networks, the data volume in the resulting data stores is manageable and storage costs are in general within the budgetary constraints of an university's IT organization. Because they invite exploring problems "in the small" that may be difficult or impossible to study "in the large", these data stores could become promising venues for possible industry-university collaborations. For example, a campus network-based study may identify precisely-defined problem-specific small subsets of data that are amenable for continuous collection even in a large production network where a more full-fledged data collection would be infeasible.

Revisiting data privacy. To start, our proposal creates new responsibilities for the university's IT organization. As sole owner of all the university's information, it is responsible for safeguarding the resulting data store, protecting user privacy, deciding on what data can/should not be collected and/or stored (and in what form), and arbitrating what data can or cannot be made available to which of the university's many different constituents.

Mainly because of data privacy, the data stores that result from our proposed use of campus networks as data sources are not likely to ever become publicly available. Each data store is only assumed to be available for 'internal" use (see Section 3). In fact, our proposal neither assumes nor relies on these data stores ever being open-sourced and instead advances the notion of open-sourcing the learning algorithms that university researchers will develop using their own campus network's data store. In fact, using such open-sourced learning algorithms and training them with data from some other campus networks (each with its own data store) suggests a viable path for tackling the much-debated reproducibility problem in science in the era of AI/ML [20]. Being able to evaluate the resulting learning models (one per considered campus network) and comparing their performance across these various production networks may increase the overall confidence in newly designed learning algorithms. It may even peek the interest of industry researchers to experiment with them in their own production networks (assuming they practial value in the algorithms) and also may tempt operators to consider such reproducibility as an item on a possible future road-map for "road-testing" AI/ML-based tools.

More generally, it is our hope that this proposal will motivate a renewed engagement between academic researchers, IT organizations, the administration (e.g., Office of General Counsel, Office of Audit and Compliance), and the user community at various universities. In particular, if a university
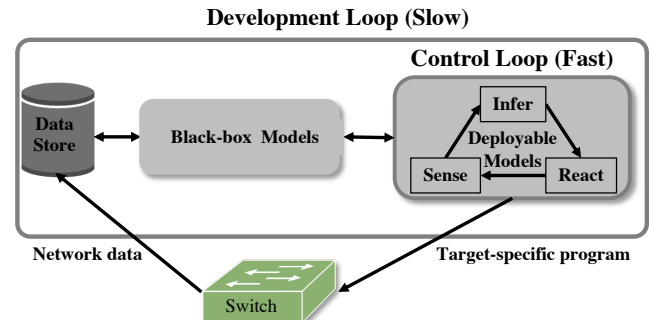

Figure 2: Road to deployment: A (slow, offline) development loop leverages the campus network (both as a data source and a testbed) to obtain a deployable learning model that performs the (fast, online) control loop capable of sensing, inferring, and reacting in real time to a specific network event.

can convince the users of its network that even though traffic is monitored, user data is guaranteed to be only used for improving the network's security and performance and not for any kind of monetization by third-parties, there may be an opportunity to finally get data privacy "right", and what better places to accomplish this than in university networks!

A roadmap for "road-testing" AI/ML-based tools.

For learning models developed for certain network automation tasks, a possible road to deployment entails four distinct steps and is illustrated in Figure 2: *(i)* Develop a learning model in an offline fashion, fully utilizing the available data store information, unconstrained by time and compute resources needed, and leveraging traditional and typically complex and heavyweight black-box learning models; *(ii)* leverage recent developments on the topic of Explainable Artificial Intelligence (XAI) [8–10, 21] and replace the learning model in *(i)* with a *deployable* learning model (*i.e.*, a learning model that is explainable or interpretable, lightweight and closely approximates the original model); *(iii)* compile the deployable learning model from *(ii)* into a target-specific program (e.g., P4 [11]) and configure the programmable switches (e.g., Barefoot Tofino [1]); *(iv)* explain to the network operator how a given deployable learning model works and accomplishes the task it was designed to perform.

To illustrate the need for opening up the black box [18] (step *(ii)*) and turning it into a "white box" (step *(iv)*), envision, for example, a learning model developed for a specific network automation task that could be routinely queried for the list of pieces of evidence that the model used to arrive at its decisions. If after going through such a list, the operators would conclude that based on the presented evidence, they would have made the same decision, their level of trust in

the learning model would increase. That trust level would increase even more if the provided evidence describes scenarios that the operators only add to their knowledge base de-facto; *i.e.*, after recognizing the correctness of the model's decisions for scenarios they were unaware of—a learning model that teaches operators things they know they didn't know!

Innovation in education. With the enormous interest in AI/ML across the sciences and the recent flurry of data science initiatives and associated openings of data science institutes at the different universities, we argue in this proposal that the time may have come for academic institutions in general and the universities' CS or EE departments in particular to change their time-tested approach to teaching and research. For example, using campus networks as real-world production networks and leveraging them for populating data stores with network data of unprecedented quality and quantity provides unique opportunities for education (e.g., new and innovative data science curricula for CS students) and research (e.g., easy and much-needed access to real-world data in support of the CS researchers' interest AI/ML-based methods).

Our proposal also highlights the importance for modern academic institutions to radically change the ways a university's IT organization, its research and teaching faculty, and its student body interact and can benefit from one another. For example, for a university's IT organization and its CS or EE departments, it may make sense to join forces and contemplate hiring a small staff of professional developers and engineers to oversee and coordinate all activities related to operating the campus network "as a lab." Such an arrangement has the potential of facilitating internal tech transfer (e.g., the IT organization testing and deploying tools developed by the researchers), completely revamping entire programs of studies (e.g., programming, networking, cyber security), and revolutionizing the training of the next generation of researchers and engineers (e.g., "hands-on" and data-driven teaching and learning network management-related problems in an actual production network). The workforce produced by such radically-overhauled academic institutions promises to have the skill set needed to advance AI/ML-based networking for the benfits of both academia and industry.

6 RELATED WORK

Two recent survey papers [12, 27] provide an excellent account of the past use of and recent advances in AI/ML for networking and also allude to some of the challenges for networking research in the era of AI/ML discussed in this paper.

There have been numerous efforts in the past that have advocated for a democratization of the field of networking. An early such effort that predates the age of AI/ML and also today's growing digital divide between academic and industry

researchers was the "Network Oracle" proposal [22]. Our proposal has much in common with this early effort, but is both less ambitious and more practical. Instead of a system for the Internet's core, we propose to focus squarely on the Internet's edge by leveraging to the fullest possible extent university or campus networks. At the same time, with the necessary technologies already available (e.g., scalable network monitoring and data collection solutions, programmable data planes, scalable platforms for processing distributed streaming data), our proposal can be implemented today and at a reasonable cost.

Similarly, our two-part proposal has much overlap with and builds on the ideas of past work such as [26] (*i.e.*, use of campus networks for deploying OpenFlow), ongoing efforts such as the "learning-and-deployment" platform Puffer [6, 28], and the recent P4 Campus initiative for developing production-level P4 programs to run in real campus networks [5]. Our proposal generalizes these learning and/or deployment environments by leveraging a university or campus network in its entirety and to its fullest extent; that is, as source for an abundance of rich network data (in support of "learning") and as real-world testbed for experimenting with and evaluating new AI/ML-based tools (in support of "deployment").

7 CONCLUSION

Concerned about a growing chasm between the "haves" and "have-nots" [24], we outline a vision for democratizing networking research in the era of AI/ML. In particular, we propose that academic networking researchers should start leveraging campus networks for their research efforts. As real-world production networks, these settings become invaluable sources for network data of unprecedented quality and quantity that promises to transform how AI/ML will be viewed and used in networking. At the same time, these settings can also be used as unique testbeds where AI/ML-based solutions can be road-tested under realistic conditions. We posit that the proposed re-purposing of university networks has enormous potential, and the best way to realize this potential is for networking researchers at universities to go and start talking to their IT colleagues.

REFERENCES

[1] 2019. Barefoot's Tofino. (2019). https://www.barefootnetworks.com/technology/

[2] 2019. Georgia Tech Technology Services Organization: System Notices. (2019). https://support.cc.gatech.edu/alerts

[3] 2019. MASS OPEN CLOUD: An Open Cloud Exchange Public Cloud. (2019). https://massopen.cloud/about/

[4] 2019. NIKSUN Supreme Eagle. (2019). https://www.niksun.com/c/1/ds/NIKSUN_datasheet_Supreme_eagle.pdf

[5] 2019. P4 Campus: P4 Applications For Campus Networks. (2019). https://p4campus.cs.princeton.edu

[6] 2019. Puffer. (2019). https://puffer.stanford.edu

[7] 2019. Puppet. (2019). https://puppet.com/

[8] O. Bastani, C. Kim, and H. Bastani. 2017. Interpretability via Model Extraction. In *FAT/ML Workshop*. https://obastani.github.io/docs/fatml17.pdf.

[9] O. Bastani, C. Kim, and H. Bastani. 2019. Interpreting Blackbox Models via Model Extraction. https://arxiv.org/abs/1705.08504.

[10] O. Bastani, Y. Pu, and A. Solar-Lezama. 2018. Verifiable Reinforcement Learning via Policy Extraction. In *NeurIPS*.

[11] P. Bosshart, D. Daly, G. Gibb, M. Izzard, N. McKeown, J. Rexford, C. Schlesinger, D. Talayco, A. Vahdat, G. Varghese, and D. Walker. 2014. P4: Programming Protocol-independent Packet Processors. *ACM SIGCOMM* (2014).

[12] R. Boutaba, M. Salahuddin, N. Limam, S. Ayoubi, N. Shahriar, F. Estrada-Solano, and O. Caicedo-Rendon. 2018. A Comprehensive Survey on Machine Learning for Networking: Evolution, Applications and Research Opportunities. In *Journal of Internet Services and Applications*.

[13] J. Buolamwini and T. Gebru. 2018. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In *FAT*.

[14] S. Burton, L. Gauerhof, and C. Heinzemann. 2017. Making the Case for Safety of Machine Learning in Highly Automated Driving. In *SAFE-COMP*.

[15] H. Caesar, V. Bankiti, A. H. Lang, S. Vora, V. E. Liong, Q. Xu, A. Krishnan, Y. Pan, G. Baldan, and O. Beijbom. 2019. nuScenes: A multimodal dataset for autonomous driving. *arXiv preprint arXiv:1903.11027* (2019).

[16] M.-F. Chang, J. W. Lambert, P. Sangkloy, J. Singh, S. Bak, A. Hartnett, D. Wang, P. Carr, S. Lucey, D. Ramanan, and J. Hays. 2019. Argoverse: 3D Tracking and Forecasting with Rich Maps. In *Conference on Computer Vision and Pattern Recognition (CVPR)*.

[17] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei-Fei. 2009. Imagenet: A large-scale hierarchical image database. In *IEEE Conference on Computer Vision and Pattern Recognition*.

[18] A. Dethise, M. Canini, and S. Kandula. 2019. Cracking Open the Black Box: What Observations Can Tell Us About Reinforcement Learning Agents. In *ACM SIGCOMM NetAI Workshop*.

[19] C. Dwork, M. Hardt, T. Pitassi, O. Reingold, and R.Zemel. 2012. Fairness through awareness. In *ITCS*.

[20] P. Gosh. 2019. AAAS: Machine learning 'causing science crisis'. https://www.bbc.com/news/science-environment-47267081.

[21] D. Gunning. 2017. DARPA Program on Explainable Artificial Intelligence (XAI). https://www.darpa.mil/program/explainable-artificial-intelligence.

[22] J. M. Hellerstein, V. Paxson, L. Peterson, T. Rosco, S. Shenker, and D. Wetherall. 2005. The Network Oracle. In *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering*. bni-acwiki.icsi.berkeley.edu/pubs/networking/networkoracle05.pdf.

[23] S. Ikeda. 2019. Recent Hacks Show That Even Tech-Savvy Universities Are Still Very Vulnerable to Cyber Attacks. https://www.cpomagazine.com/cyber-security/recent-hacks-show-that-even-tech-savvy-universities-are-still-very-vulnerable-to-cyber-attacks/.

[24] S. Lohr. 2019. At Tech's Leading Edge, Worry About a Concentration of Power. (2019). https://www.nytimes.com/2019/09/26/technology/ai-computer-expense.html

[25] X. Ma, K. Driggs-Campbell, and M. Kochenderfer. 2019. Improved Robustness and Safety for Autonomous Vehicle Control with Adversarial Reinforcement Learning. In *https://arxiv.org/pdf/1903.03642.pdf*.

[26] N. McKeown, G. Parulkar, T. Anderson, L. Peterson, H. Balakrishnan, J. Rexford, S. Shenker, and J. Turner. 2008. OpenFlow: Enabling Innovation in Campus Networks. In *ACM SIGCOMM Computer Communication Review 38(2)*. http://ccr.sigcomm.org/online/files/p69-v38n2n-mckeown.pdf.

[27] M. Wang, Y. Cui, X. Wang, S. Xiao, and J. Jiang. 2017. Machine Learning for Networking: Workflow, Advances and Opportunities. In *IEEE Network*.

[28] F. Y. Yan, H. Ayers, C. Zhu, S. Fouladi, J. Hong, K. Zhang, P. Levis, and K. Winstein. 2019. Continual learning improves Internet video streaming. https://puffer.stanford.edu/static/puffer/documents/puffer-arxiv.pdf.