Research Statement

Nick Feamster

Summary

My research focuses on developing tools, algorithms, and protocols that make the network easier to manage, more secure, and more available.

Nobody notices when the network works well, but everyone suffers when it doesn't. Thus, communications networks should be both secure and available. Network *security* has many facets, ranging from the ability to stop "unwanted traffic" (e.g., spam and denial-of-service attacks) to the ability to trace back attacks to their perpetrators ("accountability"). *Availability* means that the network must provide good performance for users whenever they want to use it—unfortunately, the increasing complexity of the network, coupled with hardware faults, software bugs, misconfigurations, and malice, make it difficult to achieve this goal. Unfortunately, these two important goals have also been among the most evasive. Breakthroughs require not only extensive domain knowledge, but also the ability to techniques from a wide range of areas, ranging from economics to machine learning. My work combines domain knowledge, extensive interactions with network operators, techniques from a wide range of disciplines, and—perhaps most importantly—the competence and tenacity to implement and deploy these systems in practice. This unique combination has allowed me to build one of the few networking research groups in the world that interacts directly with network operators to deploy fundamentally new systems and technologies in real-world networks.

I discover interesting and challenging practical problems through frequent discussions and meetings with network operators and people in industry. I then tackle these problems from first principles, develop new methods, and transfer these solutions back to practice in the form of working systems. I have tackled a variety of problems in network operations, ranging from real-time network diagnosis to stemming unwanted traffic like spam to architectures for fast failure recovery. Many people—most notably, operators "in the trenches"—are also working on these problems. Unfortunately, many of the people who have the domain knowledge that best equip them to solve these problems are busy with day-to-day operations, putting out fires as they arise but rarely taking time to think about fundamental changes to the network that might eradicate these problems. My research fills this niche. I first devise methods to understand the nature of the problem in practice. I tackle domain-specific problems with tools and techniques from other disciplines—ranging from machine learning to economics to program analysis—whose principles might provide insights into a new, previously undiscovered solution. I then devise a new approach or solution, and I transfer it to practice through implementation and deployment of real-world systems.

My research in this broad area is currently focusing on several themes: (1) Internet censorship and open access; (2) home and access networks; and (3) software defined networking. These themes, which I have been developing in the past several years since receiving tenure, build on the broader research themes I have developed on network security and operations. I first survey the new leadership roles that I have assumed in research, teaching, and service. Then, I discuss each of the new research themes I have developed since tenure and the impact that they have had on both other researchers and on industry.

Highlights In the Past Three Years

Research. Since I received tenure, I have expanded my research along three themes: Internet censorship and information control, home and access networks, and software defined networking. My work in these areas produced four *SIGCOMM* papers and more than $5M in research funding. My work in home networking has already received several awards, such as the IRTF Advanced Networking Research Prize and a Research Highlight in *Communications of the ACM*. Our home networking work is now being commercial-

ized, and our technologies and Huawei is now attempting to license our innovations in Software Defined Networking (SDN) from Georgia Tech. According to Google Scholar, my h-index is 38 and my citation count is nearly 6,000.

I briefly provide some specific highlights of my accomplishments since receiving tenure below:

- *Home and Access Networks.* We began our research studying the performance of home networks and mobile networks in June 2010, when we began studying the performance of DSL networks in France. Upon realizing that accurate measurements would require deploying infrastructure in the home router itself, we began developing BISmark (Broadband Internet Service Benchmark), custom router firmware which has now been deployed in more than 300 home networks in nearly 30 countries around the world. We also developed a version of this software that runs on Android phones and has been installed by more than 4,000 users in 130 countries. The testbed that we have developed is the first of its kind to study access and cellular networks, and our work characterizing broadband Internet performance has already produced two *SIGCOMM* papers, a *Communications of the ACM* journal article, and several other workshop papers. The work has won the IETF Advanced Networking Research Prize and an ACM *Communications of the ACM* research highlight. Our research has garnered more than $2M in funding from various funding agencies, as well as initial seed money for commercialization.
- *Censorship and Information Manipulation.* Although my first work on censorship circumvention dates back to my work on the Infranet system in 2002, we started this work again on a system called Collage, which appeared in the *USENIX Security Symposium* in 2012 and has received attention from multiple news outlets, including *The Economist*, *Slashdot*, and *Ars Technica*. Recently, I founded the *USENIX Workshop on Free and Open Communications on the Internet* and I successfully led a new large $3M NSF project (awarded 2012) on censorship measurement and circumvention. In 2011, I also received a $1.5M Google Focused Research Award (with co-PI Wenke Lee) on Internet Transparency.
- *Software Defined Networking.* Our work on software defined networking over the past three years has led to one *SIGCOMM* paper, a journal article in *IEEE Network Magazine*, and several workshop papers in the *ACM SIGCOMM Workshop on Hot Topics in Software Defined Networking (HotSDN)*, a workshop that I co-founded in 2012. Additionally, our work on event-driven network control is being licensed by Huawei; I have presented our work on event-driven network control at invited keynote presentations at the *Open Network Summit* (and industry forum that draws more than one thousand attendees), the *Internet Research Task Force*, and the *IEEE Conference on Network and Service Management*. This research has also appeared in many popular technical publications such as *Ars Technica*.

Teaching and Outreach. I have begun to disseminate the results of my research and teaching efforts through online media that can reach a much greater set of people. I have developed two blogs—one about my own research and one about research methods. I have also developed the first Massive Open Online Course (MOOC) on Software Defined Networking, which is currently being offered to a course of 33,000 students. I briefly describe these significant outreach efforts below:

- *Research Blog.* I have been blogging about various networking topics at `http://connectionmanagement.org`; each post receives more than 150 views. I also received two awards: The 2012 Hesburgh Teaching Fellows award from Georgia Tech, and the Bronze Anvil journalism award for an article I wrote on Internet censorship that appeared in the *Wall Street Journal* in 2011. This blog has had nearly 10,000 views and has more than 100 regular readers.
- *Research Methods Blog.* Professor Alex Gray and I have begun turning our course notes for CS 7001 (the Instruction to the Ph.D. course that we designed) into an online book, at `http://greatresearch.org`; scheduled completion for this book is Fall 2013, when I will next offer the course.

- *Coursera Course.* To help a larger number of people learn about the history of Software Defined Networking (SDN), I have developed a Massive Open Online Course (MOOC) on SDN, which is currently being offered to more than 30,000 students. I have been making video lectures on a variety of topics ranging from the history of SDN to its uses and applications, and I intend to use these videos as part of a "flipped classroom" seminar for students at Georgia Tech this fall.

Service. My external service has included acting as program committee co-chair for a major top-tier conference in my area (*USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, the poster and demo co-chair for the other major conference *ACM SIGCOMM*. I also founded two workshops in areas where I have focused my research recently: the *USENIX Workshop on Free and Open Communications on the Internet* (co-founded with Wenke Lee) and the *ACM SIGCOMM Worskhop on Hot Topics in Software Defined Networking (HotSDN)*. I elaborate on these events and others where I hold significant leadership roles below:

- *NSDI PC Co-Chair, SIGCOMM Poster/Demo Co-Chair.* One of my most significant recent service accomplishments was to serve as the program committee co-chair for *USENIX Networked Systems Design and Implementation* in 2013. The conference is the premier conference for networked systems research and just completed with the largest attendance in history (about 260 registered attendees).
- *Founder of two new workshops.* I founded the *ACM SIGCOMM Workshop on Hot Topics in Networking (HotSDN)*, which had about 150 participants in August 2012 and about 80 submitted papers; and the *USENIX Workshop on Free and Open Communications on the Internet (FOCI)*, which is now a regular workshop at the *USENIX Security Symposium* and regularly gets about 50 attendees.
- *Founder and Co-Chair of IRTF Working Group.* I was recently selected as one of the founding co-chairs of the Software Defined Networking Research Group at the Internet Research Task Force (IRTF), a division of the Internet Engineering Task Force (IETF).

In addition to external service, in recent years, I have continued my service to Georgia Tech by co-leading the CS Ph.D. visit weekend committee (including organizing a "College of Computing Research Day" for many of the years we have had the visit day, and serving on the committee to design the Master of Science degree based on massive open online courses (MOOCs).

In the remainder of the research statement, I outline my accomplishments and research themes in more detail. I first survey three research themes that are new since receiving tenure; I then discuss my more established research themes and my accomplishments in those areas. Finally, I discuss my plans moving forward.

New Theme 1: Home and Access Networks

Access networks (*i.e.*, cellular networks and home broadband networks) are proliferating: Over 90% of US households now have Internet access, and networks have become an essential part of every home. Video streaming already accounts for over 60% of the peak download bandwidth for the Internet; remote learning is flourishing, with Khan Academy alone delivering over 86 million videos; and within five years, Forrester Research expects 63 million Americans to telecommute from home. Bandwidth to the home is also growing rapidly: Huge investments by industry and government mean that over 60% of US homes have broadband access. Inside the home, 55% of traffic is delivered to game consoles, set-top boxes, smart TVs, and mobile devices. Further, cellular networks have become the predominant mode of Internet access for many people: For example, in Brazil, Russia, India, China, and Indonesia, there are 610 million Internet users, but 1.8 billion mobile-phone connections.

Towards providing better *transparency* to users concerning their Internet service, I am developing objective, independent third-party services for users that help them both determine whether their Internet service

provider or government is restricting access to certain content or services or degrading service for particular applications and gain access to information that they might not otherwise have access to. My research on Internet transparency is focusing on three areas: (1) the *performance* that they receive from their ISP; (2) *connectivity* to various Internet destinations; (3) the *information* that they can discover via search engines and social media.

To provide users better information about the performance that they are receiving, I started Project BISmark (`http://projectbismark.net`) in 2010; BISmark is a software platform for home routers. We have already used BISmark to develop a network measurement suite for access Internet service providers; our first paper on BISmark appeared in *ACM SIGCOMM* in 2011. With collaborators in programming languages and human-computer interaction, I am now exploring ways to use BISmark to simplify the management of home networks by applying some of the same network management principles that we have learned in our studies of transit and enterprise networks.

Impact. Our results from the initial BISmark study influenced the design and implementation of the performance measurements used by the Federal Communications Commission's study of broadband connectivity across the United States. The project has been featured in *Ars Technica* and *GigaOm* and has received over 20,000 signups from interested users. We have currently deployed BISmark routers in about 250 home networks around the world; it is also currently deployed on Google's Measurement Lab. To transition some of the technologies we are developing in research to practice, I participated in Georgia Tech's venture program, Flashpoint, to scale our efforts to a larger number of users and learn more about the problems faced by ISPs, content providers, and consumers. We also received an NSF Innovation Corps grant and Georgia Research Alliance grant to help us commercialize this technology.

More recently, we have been expanding our work on BISmark across developing countries and across a broader range of devices. For example, we recently completed a study with Research ICT Africa (RIA) to characterize fixed and mobile performance across South Africa; we are in the process of expanding this study to other countries in Africa. Second, we have developed a home network performance troubleshooting tool that helps users identify whether performance bottlenecks are within their home network or in the Internet service provider (ISP) network. The Federal Communications Commission (FCC) has recently agreed to back the deployment of our software in 4,000 home networks across the United States, and Comcast has also recently agreed to a trial deployment of this software.

Beyond the impact of the technology itself in industry, I have been developing the BISmark platform as an educational tool. In Summer 2011, I hosted a BISmark "summer camp" at Georgia Tech to help students become familiar with programming network applications on the OpenWrt router platform; the week-long event was attended by about twenty students and faculty members from across the United States, France, and Italy. I have incorporated much of the material into the graduate networking course at Georgia Tech, to give students hands on experience with developing and deploying a variety of network measurement tools. Through these activities, I aim to provide students both concrete exposure to problems and concepts in networking and a platform on which they can innovate.

Most Cited Publication (56 Citations)

> S. Sundaresan, W. de Donato, N. Feamster, R. Teixeira, S. Crawford, A. Pescape "Broadband Internet Performance: A View From the Gateway" *ACM SIGCOMM*, Toronto, Ontario, Canada. August 2011. **Winner of the IETF Advanced Networking Research Prize. Selected for Communications of the ACM Research Highlights.**

Representative Publication

> S. Sundaresan, W. de Donato, N. Feamster, R. Teixeira, S. Crawford, A. Pescape "Measuring Home Broadband Performance" *Communications of the ACM*, Volume 55, Number 9. September 2012.

New Theme 2: Internet Censorship and Information Manipulation

Free and open access to information and communications on the Internet is at risk: the Open Net Initiative reports that nearly 60 countries censor some access to information on the Internet. Similarly, ISPs can degrade network performance for certain subsets of users for some or all services. For example, some ISPs have been found to routinely block or throttle certain application traffic (e.g., BitTorrent); additionally, studies of access network performance in the United Kingdom and France have revealed that the level of performance that users achieve in their homes is sometimes as little as half of the rates that ISPs advertise to their users. Although it may not be feasible to always guarantee open, unfettered access to information, users should know when their access to information has been obstructed, restricted, or tampered with.

Second, I am actively developing techniques that help users gain access to information that they might not otherwise see, as a result of overt censorship. Ten years ago, I developed Infranet, a tool to circumvent Internet censorship that was both robust to blocking attempts and deniable—meaning that an adversary could not easily detect that a user was engaged in activities to circumvent censorship; the work won the Best Student Paper Award at the *USENIX Security Symposium* in 2002. Recent developments, such as the rise of user-generated content, have made it easier to deploy censorship circumvention systems, since sites that host user-generated content can be used as covert "drop sites" for messages; based on this insight, we designed and implemented Collage, a tool that allows users to circumvent censorship firewalls by building covert channels into user-generated content. Collage was presented at the *USENIX Security Symposium* in 2010; it has been downloaded hundreds of times and appeared in various news outlets including *Ars Technica*, *GigaOm*, and *Slashdot*.

One of the growing threats to free and open access to information in the coming years will be the emergence of "soft" forms of censorship, such as intentional performance degradation, the spread of propaganda through social media, and selective filtering or placement of search results. To defend against these threats, I have begun developing techniques to identify propagandistic behavior in social media and to allow users to compare their search results with a baseline set of search results assembled through crowdsourced measurements. We have developed tools such as Bobble (`http://bobble.gtisc.gatech.edu/` and Appu (`http://appu.gtnoise.net/`), both of which now have large groups of users, to help users track online information manipulation and privacy. Our work on search poisoning, whereby an attacker can affect the search results that a user sees by polluting a user's search history through cross-site request forgery (XSRF) attacks. This work will appear in the 2013 *USENIX Security Symposium*.

Most Cited Publication (97 citations)

> N. Feamster, M. Balazinska, G. Harfst, H. Balakrishnan, D. Karger. "Infranet: Circumventing Web Censorship and Surveillance" *Proceedings of the 11th USENIX Security Symposium*, San Francisco, CA, August 2002. **Best Student Paper Award.**

Representative Publication

> S. Burnett, N. Feamster, S. Vempala "Chipping Away at Censorship Firewalls with User-Generated Content" *USENIX Security Symposium*, Washington, DC. August 2010.

New Theme 3: Software Defined Networking

In 2002, Larry Peterson, Scott Shenker, and Jon Turner argued that networking research had "ossified", because researchers faced a huge deployment hurdle for deploying their research in production environments, and also because the large stakeholders had little incentive to allow disruptive innovation to take place. Their argument was essentially that, by "letting a thousand flowers bloom", multiple networking

technologies could be deployed in parallel, thereby providing researchers a path to innovation. The main research challenge was how to design and implement a virtual network infrastructure that supported this philosophy.

Towards solving this challenge, I began working on network virtualization during my postdoc at Princeton. Network virtualization allows multiple networks to operate in parallel on the same physical infrastructure. Although this concept is not new (commonly used Virtual Private Networks, or "VPNs", come to mind as a prominent real-world example of network virtualization), virtualizing all aspects of the network infrastructure—in particular, both the links *and* the routers themselves—holds great promise for enabling innovation. Jennifer Rexford and I wanted to implement a new network protocol we had designed at the end of my graduate career. Our plan was to use PlanetLab—a large testbed with virtualized servers distributed around the world—to do it. Unfortunately, we quickly realized that PlanetLab did not have the necessary functions to instantiate test *networks*; in particular, PlanetLab offered no functions for building virtual routers and links, and also had no support for forwarding traffic at high rates for virtual routers (e.g., every packet needed to be copied several times at each node, significantly slowing the packet forwarding rates). These shortcomings caused us to pursue a larger project to build such a testbed that would support the kinds of experiments that we wanted to run. With Andy Bavier and Larry Peterson, we built a Virtual Network Infrastructure (VINI), a testbed that allows researchers to build virtual networks. This work appeared in *ACM SIGCOMM* in 2006. The concepts behind virtual programmable networks, in concert with some of our earlier work on the Routing Control Platform (RCP) ultimately led to the advent of Software Defined Networking (SDN)

Since this initial work, I have focused on three aspects of software defined networking: (1) providing Internet connectivity and routing control to software defined networks; (2) designing very fast packet forwarding technologies for software defined networks; (3) designing better languages and control models for software defined networks.

A virtual network—either an experiment or a distributed "cloud" service—typically needs connectivity to the rest of the Internet so that users can actually exchange traffic with it. To provide such connectivity, and to give each virtual network direct control over how user traffic reaches it, I designed, implemented and deployed the Transit Portal, a software-defined controller for interdomain routing that provides individual virtual networks the illusion of having a direct, physical upstream connection to multiple Internet service providers. This work appeared in *USENIX Annual Technical Conference* in June 2010. We performed several research projects to follow up on this work, which used the Transit Portal to improve both the reliability and performance of cloud-hosted Internet services. This follow-up work has appeared in *ACM SIGCOMM* in 2012, and *ACM SIGMETRICS* in 2013.

The Transit Portal is also a cornerstone of the larger nationwide GENI effort (featured here, for example: `http://www.geni.net/?p=1682`). Our work on designing faster packet forwarding technologies for virtual networks started with the Trellis project, which moved packet forwarding for virtual networks into the kernel; although this work resulted only in a workshop publication, the software itself was adopted by University of Utah's Emulab, the most prominent emulation-based testbed for networking research. Our current efforts have focused on accelerating packet forwarding further by supporting custom packet forwarding for virtual networks in Field Programmable Gate Arrays (FPGAs); our work on SwitchBlade, a platform for rapidly developing and deploying custom forwarding engines in hardware for virtual networks, appeared at *ACM SIGCOMM* in August 2010.

Finally, I have been developing new control models and languages to support event-based control for software defined networks. I have focused on how better control models and languages can help solve three problems in network management: (1) enabling frequent changes to network conditions and state; (2) providing support for network configuration in a high-level language (including developing one of the first formal languages for software defined networks, Procera); and (3) providing better visibility and control over tasks for performing network diagnosis and troubleshooting. With my students, I built and deployed software defined networks in campus and home networks to demonstrate how SDN can improve common network management tasks. An early version of this work appears in the February 2013 issue of *IEEE*

Network Magazine.

Impact. The impact of this work thus far has been to support network experimentation for researchers; many other virtual network technologies and platforms have built on this work. Our work on virtual networks has been over nearly 500 times (the VINI paper has been cited more than 300 times, and our work describing a network architecture based around network virtualization has been cited over 200 times).

The Transit Portal is currently deployed in six locations—including a recent deployment in the Amsterdam Internet Exchange in May 2013—and I am using it in my courses to provide students with hands-on experience configuring networks of routers and connecting them to real BGP-speaking routers on the Internet. The course I have developed that uses this technology is likely serves as the first course where students can directly configure networks of routers that are connected to the global Internet. The Transit Portal is also being actively used in research and has supported many other research projects, including several projects at the University of Souther California and the University of Washington that have resulted in multiple independent research papers that have appeared at *ACM SIGCOMM*.

Most Cited Publication (428 citations)

> A. Bavier, N. Feamster, M. Huang, L. Peterson, J. Rexford. "In VINI Veritas: Realistic and Controlled Network Experimentation". *ACM SIGCOMM*, August 2006. Pisa, Italy.

Representative Publication

> V. Valancius, B. Ravi, N. Feamster, A. Snoeren "Quantifying the Benefits of Joint Content and Network Routing" *ACM SIGMETRICS* Pittsburgh, PA. June 2013.

Established Theme 1: Network Operations

A well-established of my work is *network operations*, which is what I call the field of designing networks so that they are easier to run and manage. Much of my work in this area has focused on fault detection and troubleshooting. Prior to my dissertation work, operators relied on detecting problems with networks "at runtime" on a live network. My dissertation work demonstrated that, in fact, many routing problems could be detected simply by examining the configuration of the routing protocols, before the configuration is even deployed. I applied techniques from static program analysis to routing configuration to help network operators catch mistakes and predict dynamic network behavior before the configurations are deployed on a live network, preventing costly and catastrophic network downtime.

Beyond predicting behavior and proactively detecting configuration faults, operators must understand the network's behavior *as it is running* (e.g., to detect equipment failures, attacks, or unplanned shifts in network traffic). Unfortunately, operators are drowning in heterogeneous data. To help operators better understand network faults "at runtime", I have applied unsupervised learning techniques to Internet routing data to help them efficiently mine the data for network events that require corrective action. This work appeared in *ACM SIGMETRICS* in 2007. My work has also applied statistical inference techniques to help network operators determine the answers to "what if" configuration questions in content distribution networks; we developed a system called "WISE" (What-If Scenario Evaluator) to help network operators determine the effects of configuration changes on network response time. A paper on this system appeared at *ACM SIGCOMM* in 2008 and is now used by operators and network designers at Google. A more mature version of this work that also describes deployment experiences at Google is in submission to *IEEE/ACM Transactions on Networking*.

Users of communications networks also face the potential of intentional performance degradation or manipulation by Internet Service Providers (ISPs); these problems are popularly referred to as "network neutrality violations". This transparency can help users determine whether their network is the cause of performance degradation, or whether performance problems that they are seeing are due to some other cause. With students, I designed, built, and deployed the *Network Neutrality Access Observatory (NANO)*, a system that aggregates measurements from end systems to help users and operators of edge networks infer when transit networks may be discriminating against certain types of traffic. This work appeared in *ACM SIGCOMM CoNext* in 2009, and we have deployed the system on Google's Measurement Lab (`http://www.measurementlab.net/`). More recently, we have been looking at methods for helping users diagnose general problems with access network performance and examining which factors have the most influence on access network performance.

I have developed new network protocols and architectures that improve availability and accountability in communications networks in the face of both faults and malice. Networks face the continual threat of failures and attacks that disrupt end-to-end connectivity. Prior to my work, one promising approach to improving connectivity involved routing traffic along multiple paths between two endpoints ("multipath routing"); despite the promise of this approach, previous approaches encountered two significant challenges: First, previous approaches for disseminating information about multiple paths through the network did not scale to large networks. Second, end systems had no way to signal to the network that an end-to-end path had failed or was providing inadequate performance. My research applied a new perspective to this problem: rather than simply routing traffic on one of multiple paths to a destination, allow traffic to switch paths at intermediate points en route to the destination, and allow end systems to signal to the network when it should attempt to use a different path to the destination using a small number of bits that can be carried in the traffic itself. This system, called *path splicing*, provides up to an exponential improvement in reliability for only a linear increase in the amount of state that each router in the network must store.

New research since tenure. Since receiving tenure, I have continued to work on tools and protocols that help operators configure their networks better. To better understand how network operators make changes to network configurations, we performed a study of the evolution of network configuration over five years

across two campus networks and have clustered these changes into common tasks, with an eye towards raising the level of abstraction for network configuration. This work appeared in the *ACM SIGCOMM Internet Measurement Conference* in 2011.

Second, we have been actively developing systems on top of the Internet routing infrastructure to help network operators optimize the performance of their applications that run in the network. We developed a system called PECAN which jointly optimizes content routing (*i.e.*, the mapping of clients to service replicas) and network routing (*i.e.*, the network-level paths between clients and their respective replicas). We discovered that jointly optimizing network and content routing can significantly improve performance over simply performing each operation independently; our results will appear in *ACM SIGMETRICS* 2013.

Finally, we have performed a data-driven econometric analysis that showed how a tiered pricing model can yield both higher profit margins for Internet service providers and greater consumer surplus for users. These results appeared in *ACM SIGCOMM* in 2011.

Impact. The foundation of this research theme comes from a system I built called called "rcc" (router configuration checker). This system was the centerpiece of my doctoral dissertation and has had significant impact in both research and industry. The work received the Best Paper Award at *ACM/USENIX Networked Systems Design and Implementation (NSDI)* in 2005 and has been used by hundreds of Internet Service Providers (ISPs) around the world to check their network configurations for errors.

The path splicing work resulted in a Sigma Xi undergraduate research award for Megan Elmore. The work was funded by Cisco, and they have considered the possibility of extending their existing multiple routing configuration (MRC) function to support path splicing. A more likely deployment scenario, however, may be the incorporation of path splicing into a network where network elements are more programmable. We have published an open-source implementation of path splicing on several programmable networking platforms.

Our work on tiered pricing was covered extensively in the media, including in *The Economist*.

Most Cited Publication (223 Citations)

> N. Feamster and H. Balakrishnan "Detecting BGP Configuration Faults with Static Analysis" *2nd Symposium on Networked Systems Design and Implementation (NSDI)*, Boston, MA, May 2005. **Best Paper Award.**

Representative Publication

> M. Tariq, A. Zeitoun, V. Valancius, N. Feamster, and M. Ammar. "Answering What-if Deployment and Configuration Questions with 'WISE': Techniques and Deployment Experience". *IEEE/ACM Transactions on Networking*, January 2013. (also appeared in *SIGCOMM* in August 2008)

New Representative Publication (Since Tenure)

> V. Valancius, C. Lumezanu, N. Feamster, R. Johari, V. Vazirani "How Many Tiers? Pricing in the Internet Transit Market" *ACM SIGCOMM* Toronto, Ontario, Canada. August 2011. **Appeared in multiple popular news venues, including** *The Economist*.

Established Theme 2: Network Security

My research explores the role that communications networks—in particular the network layer—can play in improving computer and communications security. This line of research began with my arrival at Georgia

Tech in 2006. A cornerstone of this research is a system that was published in August 2009 called "SNARE" (Spatio-temporal Network-level Automated Reputation Engine). This work appeared at the *USENIX Security Symposium*, a top-tier security conference. The main idea behind SNARE—and the key insight behind my research in spam filtering—is that spammers have different sending behavior than legitimate senders. Filters can distinguish spammers from legitimate senders by examining their *sending behavior* (i.e., how they send traffic), rather than what is in the messages themselves.

Prior to my research, conventional spam filters attempted to distinguish spam from legitimate email by looking at message contents: that is, they would look at the words or language used in the messages themselves and try to detect spam based on what the message said. This approach has become increasingly untenable, since spammers have begun to embed their messages in all sorts of media, ranging from images to PDFs to audio files to spreadsheets—by the time developers perfected their content filters for one type of medium, spammers moved onto the next. My line of work has taken an entirely different, but complementary approach: I look at features of the senders' *behavior* (e.g., the time of day they are sending, whether there are other "nearby" senders on the network, whether and how the sizes of the messages of the senders vary over time) to distinguish spamming behavior from legitimate email use. The method is harder for spammers to evade, it is more flexible because it can be deployed anywhere in the network, and it can work at much higher traffic rates than conventional approaches. This idea was first laid out in the initial award paper at *SIGCOMM* and finally realized in the SNARE paper from August 2009 at *USENIX Security*.

I have also worked on sweeping changes to the Internet architecture that could improve *accountability*, thus making it more difficult for malicious parties to operate unfettered in the first place. The current Internet architecture provides little to no accountability. Malicious end systems can conceal the source of their traffic ("spoofing"), and edge networks can provide false information about their reachability to various Internet destinations ("route hijacking"); both of these attacks make it difficult to track down perpetrators of attacks. Current approaches to solving these problems require manual configuration and operator vigilance, which make them weak and error-prone. Towards building networks that are inherently accountable, I have developed the Accountable Internet Protocol (AIP). One of my contributions to the design was to make the addresses in this protocol self-certifying, which forms the cornerstone of the basic design. I also demonstrated how to apply AIP to secure BGP, the Internet's interdomain routing protocol.

New research since tenure. I have continued my research in network security by studying how attackers use the underlying Internet infrastructure to achieve *agility*. In particular, we performed a study that explored the initial DNS behavior of spammers that appeared in the *ACM SIGCOMM Internet Measurement Conference* in 2011. We also performed a second study that explored how attackers use the Internet's interdomain routing protocol, Border Gateway Protocol (BGP) to evade detection when sending spam and performing other malicious activities. That study appeared in the *Passive and Active Measurement Conference* in 2011. Finally, we are exploring how to prevent data leaks from cloud-based Web applications, even when the applications themselves may be compromised. We have one preliminary paper in the *USENIX Workshop on Hot Topics in Cloud Computing (HotCloud '11)*.

Impact. My research in network security has had impact in research, in industry, and on the national level. My research on this topic has earned the Presidential Early Career Award for Scientists and Engineers (PECASE), a Sloan fellowship, and the Best Paper Award at *ACM SIGCOMM* (the premier computer networking conference). Aspects of my work have also been incorporated into commercial spam filtering products and Web mail clients at companies including Yahoo, Cisco/Ironport, and McAfee, as well as a project for the Department of Defense on high-speed network monitoring. My paper on understanding the network-level behavior of spammers—which won the Best Student Paper award at *SIGCOMM* in 2006—has been cited over 300 times since its initial publication in August 2006—it spawned a variety of high-impact follow-on work, including looking at network-level behavior not only to develop better spam filters, but also to detect botnets more effectively and defend against phishing attacks, click fraud, and other serious threats to the Internet infrastructure. I have also been working on similar approaches to help detect and dismantle the Internet's scam hosting infrastructure (e.g., Web sites that attempt to steal user passwords, money, and so forth). My initial paper on this topic ("Dynamics of Online Scam-Hosting Infrastructure")

won the Best Paper award at the *Passive and Active Measurement* conference in April 2009.

My work on SNARE has also garnered significant attention in industry. This work was featured in *Technology Review* and on Slashdot (a popular, high-traffic site for news in information technology). Several companies including Yahoo have incorporated the network-level features that SNARE identifies into its spam filters, and companies that develop spam filtering appliances, such as McAfee, are also using these features to improve the accuracy and performance of their spam filtering appliances.

AIP appeared in *ACM SIGCOMM* in 2008; an early version of the design also appeared in *ACM Workshop on Hot Topics in Networking (HotNets)*. I am incorporating a version of this technology into a working system and transferring them to practice. I am working with BBN on a DARPA project that will ultimately result in incorporating AIP's mechanisms into a military network protocol that allows attribution of traffic to sources (the details may ultimately be classified).

My impact on the broader field of cybersecurity goes beyond my own research. I am also having impact in the national arena in several ways. Last year, I was involved in setting the nation's agenda for cyber security, through multiple additional activities. First, I led a community-wide effort to develop a "wish list" document that describes the security community's needs for access to better data—ranging from network traffic data, to data about our country's infrastructure. This report was ultimately delivered to Tom Kalil, the deputy director for policy in the Office of Science and Technology Policy. Second, with program managers Karl Levitt and Lenore Zuck at NSF, I organized a community-wide, multi-agency workshop on "Security-Driven Architectures". The workshop included participants from computer science, with an eye towards setting a research agenda for developing more holistic approaches to computer security that consider *all* aspects of computer and communications systems, rather than just a single piece (like the network). Finally, my work on developing next-generation Internet protocols to improve accountability (which could eradicate spam in the first place), based on work that appeared at *ACM SIGCOMM* in 2008, was included in reports for the National Cyber Leap Year.

Our recent work on DNS and BGP reputation has been patented and implemented by Verisign, and is currently in use in several of their security products. Our recent work on detecting fraudulent voting on webmail messages has been implemented and deployed in Yahoo's webmail system.

Most Cited Publication (445 Citations)

> A. Ramachandran and N. Feamster. "Understanding the Network-Level Behavior of Spammers". *ACM SIGCOMM*, August 2006. Pisa, Italy. **Best Student Paper Award.**

Representative Publication

> S. Hao, N. Syed, N. Feamster, A. Gray and S. Krasser. "Detecting Spammers with SNARE: Spatio-temporal Network-level Automatic Reputation Engine". *USENIX Security Symposium*, August 2009. Montreal, Quebec, Canada.

New Representative Publication (Since Tenure)

> S. Hao, N. Feamster, R. Pandrangi "Monitoring the Initial DNS Behavior of Spammers" *ACM SIGCOMM Internet Measurement Conference*, November 2011. Berlin, Germany. **Resulted in two Verisign patents.**