**NEIL SPRING**

**Research Statement**

My research interests lie in the systems aspects of networking — designing networking protocols for robustness and efficiency, evaluating them with real workloads and topologies, and building systems that enhance the Internet. I first describe my dissertation work in understanding Internet structure and configuration. This work is a step toward my broader goal: to build protocols that make the Internet more robust to misbehavior, misconfiguration, and failure. I then describe my system to make network measurement tools easy to develop and deploy, allowing researchers to develop new tools and execute large-scale analyses. Next, I briefly describe my work in building tools using this framework to diagnose faults in network paths. Finally, I describe two future directions I plan to pursue: collaborative measurement and information-rich routing protocols.

## Understanding Internet Structure and Configuration

Without measured data about Internet topology and configuration, the consequences of failure are hard to predict and the effectiveness of new protocols is hard to evaluate. This global view of Internet topology and configuration does not exist: the Internet is administered by many service providers (ISPs) that have no incentive to share the detailed topologies needed to assemble a global view. Measuring these ISP topologies without cooperation is difficult because the Internet provides little explicit support for measurement; what few primitives exist may be blocked or rate-limited because they are not central to Internet operation. The Internet is also large, heterogeneous, and growing. My hypothesis was that despite these challenges we could discover enough information to answer operationally relevant questions. My work showed that this hypothesis is true: reasonably accurate network maps can be measured without help from service providers (ISPs), and these maps have allowed researchers to evaluate protocols on realistic topologies.

With colleagues, I built the Rocketfuel network mapping system to produce good maps of ISP networks. This work appeared in SIGCOMM'02, won the best student paper award, and is soon to appear in ACM/IEEE Transactions on Networking. My approach was to consolidate the information made available in the course of normal network operation, such as naming (DNS) and routing (BGP) information, to guide the active measurement of network paths and to help interpret the result. By carefully selecting which paths to measure, I was able to use public traceroute servers as measurement sources. Although public traceroute servers provide hundreds of distinct vantage points on the network, each is capable of conducting relatively few measurements compared to those of the dedicated servers used by previous efforts. Taking fewer measurements from more sites is a worthwhile tradeoff because a diversity of vantage points improves the fidelity of the resulting network map. With this approach I was able to take only 0.1% of the measurements required by a brute force approach and still capture an unprecedented view of ISP networks with seven times as much detail as prior mapping efforts. Rocketfuel maps have been validated by ISP operators and have been widely used by other researchers. These maps satisfied a real need of the research community; in fact, the last SIGCOMM conference featured three papers that used Rocketfuel maps.

Although topologies alone are useful because they determine which paths are available to traverse the network, understanding routing policy is just as important because it determines which paths are used. Routing policy is expressed at three levels: inter-domain, the selection of which ISP to use to reach a given destination; peering, the selection of where to deliver a packet to an adjacent ISP; and intra-domain, the selection of a path within an ISP network. With colleagues, I used the paths measured in Rocketfuel to infer routing policies at all levels, and then evaluated how these policies affect end-to-end Internet path latency. This work provided some of the first empirical data on ISP peering policies and appeared in SIGCOMM 2003.

To infer routing policy, I observed that a detailed map includes many paths *not* taken, and each alternate path represents a policy choice made by the network operator. For intra-domain policy, in which packets are routed along the shortest weighted path, my colleagues and I translated this observation into a system of linear constraints. The weight, or cost, of each link is a variable, and each relation states that the sum of the weights along chosen paths is less than the sum of the weights along every alternate. This constraint system may not be consistent, however, because measurements observe false paths that occur when routes change, and these false paths must not be allowed to corrupt the result. So that the constraint solver would find link weights consistent with the most measurements, we adapted the constraint hierarchy technique: each constraint is given an error value weighted by the number of times the path was observed. The constraint solver minimizes the sum of weighted error values. The resulting link weights showed that most ISPs route according to shortest distance, with adjustments to avoid bottlenecks.

This study also featured a router-level view of inter-domain routing, allowing us to estimate the performance implications of routing policy choices. Inter-domain routing is constrained primarily by bilateral commercial relationships. I found, contrary to expectation, that these constraints did not make paths significantly longer. Instead, the default route selection algorithm deserves most of the blame for inefficient routes because it is insensitive to performance. Routing policy previously thought to adversely constrain path selection often chooses better paths!

### Distributed Measurement Systems

My network mapping experiments showed me the importance of having many vantage points for measurement, but they also exposed limitations of standard tools. I wanted to allow novice researchers, including students in graduate courses, to design and test more robust and efficient measurement tools. This openness is important because creativity is not limited to those with experience. Because existing systems restrict the measurements that can be run, new measurement techniques are hard to validate on diverse paths. As a result, although the community has developed many measurement tools, these tools are not always trusted to produce the correct result. The technical obstacle is that measurement tools often require privileged access to the network interface to send and receive unconventional packets. This requirement makes it difficult to develop, test, and deploy even the slightest modification. In response, I built Scriptroute, a system for wide-area network measurement, to allow researchers to create, fix, and extend measurement tools and then widely deploy the result without violating security constraints. This work appeared in *Scriptroute: A public Internet measurement facility*, which received the best student paper award at USITS '03.

My insight was that the part of a network measurement tool that requires administrator privilege is not the part that tool writers are likely to change. The privileged mechanism can thus be implemented in an unchanging daemon that prevents the tool from sending "bad" (malformed or excessive) traffic or "sniffing" (receiving unrelated) traffic. The measurement logic can be implemented in an interpreted script so it is easy to understand and customize. This separation simplifies tool development, both because measurements no longer run as root and because they can be implemented in scripting languages that support high-level features for composition and extension.

While Scriptroute is useful for development alone, it also simplifies deployment, allowing researchers to execute customized measurements remotely without the trust of the site administrator. My approach is to restrict network measurements to safe behavior, both by limiting use of host resources and by filtering dangerous packets, so that any client can connect and request a measurement. I rejected as too inflexible the conventional approaches in which either the code of a measurement tool is audited or the user must be considered trustworthy. Either limits network measurement to the privileged few. This approach borrows

from recent techniques for safely executing untrusted code, and provides an example of how to support a class of anonymous networked applications.

Although only a year old, Scriptroute has proven successful. It is one of the first services to run consistently on the PlanetLab testbed, which hosts more than 200 Scriptroute servers. It has been used in graduate networking course projects and research at MIT, Michigan, Minnesota, Georgia Tech, UC San Diego, UC Santa Barbara, the Hong Kong Polytechnic University, and in a variety of published research here at the University of Washington.

### Internet Diagnosis

One application of measurement is to improve the diagnosis of performance faults along Internet paths. End to end failures in the Internet are common, but users are often at a loss to diagnose problems and determine who to contact. Although network management tools allow individual network operators to diagnose problems on their own networks, when a problem lies elsewhere, operators are in no better position than users to determine who is responsible. In *User-level Internet Path Diagnosis*, SOSP 2003, my colleagues and I found that surprising diagnosis capability is latent in the Internet. We built an Internet path diagnosis tool, named Tulip, built on Scriptroute and detects where pathologies such as packet loss, reordering, and significant queuing, occur down to the level of an individual link. Perhaps the most interesting result of this work is that we present an ideal yet practical architecture for Internet diagnosis derived from first principles. This ideal Internet diagnosis architecture can be approximated on the Internet today and suggests ways to improve the diagnostic capability of the Internet in the future.

### Future Work: Collaborative Measurement

Despite much progress in improving network measurements, we still lack the ability to completely understand the Internet's behavior, let alone predict how the Internet will evolve. Yet we know it will change: Industry is quickly building devices too small, cheap, or numerous to have their own user interfaces, and these devices will use the network to interact with users and the outside world. In designing Internet protocols, the diversity of devices and interests becomes important, and today we lack the tools to predict the effects of failure, flash crowds, and aggressive applications. I believe that measurement is the first step toward this understanding.

Isolated efforts in network measurement will not allow researchers to predict the efficacy of proposals because a complete picture of the Internet is outside the reach of any single research group. While a few efforts provide historical context or a global view, others are narrowly focused on a few properties of a few paths at a single point in time. These analyses have yet to be combined, so a vast space of properties and correlations remains unexplored.

To develop our understanding of the network, I plan to build a "measurement blackboard" to coordinate measurements. The goal is to do for network measurement research what the network simulator, ns, has done for simulation: to allow researchers to verify prior results and build upon prior analyses. My idea is to build a logically centralized but physically distributed data store in which different research groups can contribute incrementally updated tables of raw measurement data, as well as analyses that produce views that summarize the data. As each new link is discovered, other tools can immediately begin to measure its performance. Constructing and populating this database is a grand challenge I have described in my HotNets position paper, *Reverse-Engineering the Internet*.

This collaborative platform is exciting because it would change the way researchers study the Internet. We could compare results across time to track and predict Internet evolution. We could more easily validate our analyses, both by comparing results at different times or with different tools, and by comparing to

authoritative partial views of ISPs. We could look for correlations between properties that are relevant for design, for example, to tell what makes a "backbone" router different from an "access" router. Finally, we could change the way we evaluate new designs through simulation, avoiding artificial or random scenarios and instead choosing realism.

### Future Work: Information-Rich Routing Protocols

Internet measurement provides us with a new opportunity to rethink how protocols work. Most Internet protocols are naive about network state: they either measure relevant state from each source to each destination (as in congestion control) or ignore performance altogether (as in most routing protocols). The result is that systems are highly fragile and require operator tweaking to make them perform well. The more a network is hand-tuned, the less robust it is to failure. A different vision is to use data about how the network is working to make the network configure itself. For example, I intend to build a network routing protocol that satisfies both users and operators: it will provide the high performance, reliable paths that users expect, and the configurability that operators need. As I have found in my measurements, the complexity of the inter-domain routing protocol in practice is unjustified relative to how poorly it meets both goals. It is time to develop the framework of the next inter-domain Internet routing protocol in anticipation of the opportunity to deploy its redesign.

My routing protocol approach will have two features. First, it will make local decision criteria explicit — encoding the commercial relationships between customers and providers, noting links that should be avoided due to congestion or used only as backups, and promoting paths with good performance. This is challenging to design because although these relationships can be inferred by measurement, operators may be reluctant to make such details explicit. Second, it will use a measured knowledge of topology and past performance to choose good routes and minimize the propagation of invalid or transient routes. Researchers have noticed that as routes are withdrawn after a link has failed, alternate routes that share the same failed link are advertised, slowing convergence and ultimately subverting the effectiveness of redundant "backup" links to other ISPs. Using a measurement to test whether a route is valid before propagating it further can help, but the goal is not simply to never propagate invalid routing information, but also to always choose the best available path. The best path can only be recognized by composing a measured view of performance with explicit routing policy information. This explicit, measurement-reliant protocol is far from how inter-domain routes are chosen today, but within reach and ripe for exploration.