Justine Sherry | Previous Research. The Internet today is a black box; it does not provide any information about where a packet travels between source to destination. For most uses, this design is sufficient. In most communications, a user only cares that data arrives at the destination, not how it arrives there. However, this is not appropriate for all uses. Many applications in industry, finance, and government require the Internet to be both reliable and fast. When packets are frequently lost or delayed, one of these users might wish to pinpoint the location of the loss in order to register a complaint with the network operator responsible. Unfortunately, the basic interface to the network does not provide any insight into internal network structure to resolve this problem. My research focuses on improving visibility into the internals of the network.

My undergraduate work at the University of Washington addressed internal network visibility in the context of active Internet measurements. I sent specially crafted packets into the network to illicit responses that reveal properties about the network topology. With topology measurements, any user can develop a model of the internal structure of routers in the network. This can support a variety of applications, such as diagnosing where problems lie within the Internet and performing estimates of Internet latencies. Previous work in this area relied on a limited set of tools, including traceroute and ping.

My thesis work proposed including the IP Prespecified Timestamp option, an obscure IP option that was previously ignored, in the toolset of active measurement techniques. Prespecified Timestamps are an option built-in to every IP packet. They allow the sender to specify up to four IP addresses from which to request timestamps. Each router along the path checks the first unstamped address field. If it owns that address, it will provide a timestamp. My work showed that timestamps were a useful tool to investigate the Internet topology. To this end, I worked on several applications of prespecified timestamps, which contributed to two conference publications [1, 2] as well as a conference poster [3] and an invited talk at another university [4]. My thesis on IP prespecified timestamps [5] was honored as best senior thesis in my department, and the Computing Research Association named me Outstanding Undergraduate Researcher 2010 for this work. I will discuss here two projects I undertook involving IP timestamps: Reverse traceroute, and IP alias resolution.

Reverse traceroute. When I started my work with prespecified timestamps, they were already used in a limited fashion for Reverse traceroute, a project being developed at my university. This project was a large collaboration, led by graduate student Ethan Katz-Basset and involving several students and faculty. Traceroute is a common tool that identifies the routers a packet encounters along its path when transmitted from source to destination. However, a peculiarity of Internet routing is that this source to destination ('forward') path is often very different from the destination to source ('reverse') path, and traceroute cannot reveal this reverse path. Understanding the asymmetry of forward and reverse paths is extremely important for communication; interacting with a client involves both receiving and sending traffic, and a problem with either transmission means a failure to connect to the client. This can be disastrous for Internet users who depend on connectivity, such as web commerce businesses. To address this, we designed Reverse traceroute.

My primary contribution to the project was to design an algorithm for identifying reverse path hops using IP timestamp, as one of three mechanisms used by Reverse traceroute to discover the full reverse path. A number of 'guess' hops are generated from the known topology, and timestamp probes are sent along the round-trip path requesting stamps from the destination, followed by the guessed IP address. If a stamp is received from the guessed address, this serves as confirmation that the addresses indeed was on the Reverse path the packet traversed. The basic challenge I addressed was that a variety of router vendors implemented timestamp support

very differently from each other. After performing extensive experiments over a large sampling of routers, I was able to identify the basic behaviors routers exhibited in response to my timestamp probes. One idiosyncracy was that some routers would provide timestamps for addresses that were not their own, introducing the threat of false positives. Another problem was that some destinations would fail to provide timestamps, which limited coverage. My new algorithm overcomes both of those limitations while still maintaining accuracy and efficiency.

In addition to developing the timestamp technique, I also collaborated with the other undergraduates in evaluating the accuracy of the complete system, and independently designed, built, and continue to maintain an 'atlas' of traceroutes to perform constant monitoring of Internet paths that are useful for Reverse traceroute. Our work on Reverse traceroute appeared at Network Systems Design and Implementation 2010 [1], where it was honored as best paper.

Alias Resolution. IP alias resolution is an issue that arises frequently in Internet measurement. Routers have multiple IP addresses, but multiple measurements of the same machine can reveal different IP addresses. Thus, it is necessary to identify IP addresses that belong to the same machine, or are *aliases* of each other, to understand the router-level structure of the Internet. Without proper alias data, topology measurements might lead to misdiagnosis of network failures or inaccurate latency estimates. To address this, I led a project to design an alias resolution technique which made use of timestamp measurements. As with my work in Reverse traceroute, I needed to develop classifications for how different routers would respond to requests for timestamps from their alias pairs. This was particularly challenging because I had only limited ground-truth alias data, which I needed to save for my evaluation. Thus, I had to develop my alias classifications on imperfect data generated from other techniques and heuristics. I discovered several different behaviors and to address them, I developed three distinct techniques for the identification of alias pairs. These techniques work by sending specific timestamp requests that combine timestamp requests for two addresses suspected to belong to the same router. Using my analysis, I can often classify the two addresses as aliases if I receive multiple timestamps in the router's reply, depending on the format of the packet and the timestamp values received. I evaluated my technique on a limited ground-truth dataset, and found it to be accurate. Further, I found that my technique complimented existing techniques, in that the majority of aliases I discovered using timestamps were unidentifiable using any other alias resolution tools. Thus, my technique provides strong gains in our ability to generate accurate topologies for the performance-critical Internet applications that depend on them. I recently presented this work at the Internet Measurement Conference 2010 [2].

Future Work. Since starting graduate school, I have continued thinking about how the Internet should expose its internals to users from a new perspective. My measurement work focused on what we could learn about network internals without modifying the basic interface to the Internet, relying on indirect inferences to understand the network topology. Lately, I am considering how to redesign the interface to the network, to provide for direct interaction with internal features of the network. My attached proposal describes this work. In the future, I intend to continue to pursue networking research and hope to complete my PhD in 2016.

[1] E. Katz-Bassett, H. V. Madhyastha, V. Adhikari, C. Scott, J. Sherry, P. van Wesep, T. Anderson, and A. Krishnamurthy, "Reverse traceroute," in USENIX NSDI, 2010.

[2] J. Sherry, E. Katz-Bassett, M. Pimenova, H. V. Madhyastha, T. Anderson, and A. Krishnamurthy, "Resolving IP aliases with prespecified timestamps," in ACM SIGCOMM IMC, 2010.

[3] J. Sherry, E. Katz-Bassett, H. V. Madhyastha, M. Pimenova, A. Krishnamurthy, and T. Anderson, "Internet measurements with prespecified timestamps." Poster at USENIX NSDI, 2010.

[4] J. Sherry, "The Internet measurement toolbox." Invited Talk at University of Puget Sound, Tacoma, Washington, April 2010.

[5] J. Sherry, "Applications of the IP timestamp option to Internet measurement." University of Washington, Department of Computer Science and Engineering, March 2010.