Justine Sherry

Statement of Purpose

As an undergraduate, I have had the opportunity to work on research in Internet Measurement, working towards more comprehensive and more accurate maps and models of the connections that make up the Internet. Moving forward, I want to continue to explore issues related to the Internet and network layer. I believe that an academic environment is the right place for me to explore these issues, because it will allow me to continue to study and learn as a student while enabling me to make active research contributions to the problems I find so consequential. To this end, I am applying to UC Berkeley for graduate school in Computer Science for Fall 2010. My experiences as an undergraduate at the University of Washington have prepared me for such a goal, both in developing my passion as an academic and in providing me with the background required to succeed.

Research Background

The emphasis of my undergraduate research has been on including the the IP Prespecified Timestamp option in the toolbox of Internet measurement utilities. For this work, I have been honored as the Computing Research Association's Female Outstanding Undergraduate Researcher, 2010.

Prespecified Timestamps, an option built-in to every IP packet, 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 in the form of milliseconds since midnight UTC. While traditionally ignored for presumed limited coverage and inconsistent behavior, I have demonstrated that over 25% of routers will support timestamps and extensively documented the router-specific, distinct implementations of timestamp support. Furthermore, I argue that measurements gathered with this type of probe have several unique characteristics that set them apart from other tools, providing for new means to investigate basic issues that face effective Internet measurement.

Demonstrating its utility, I have worked on several applications of prespecified timestamps. First, I have improved its usage within the Reverse Traceroute system. Next, I developed arguments for confirming IP alias pairs using timestamp measurements. Currently, I am exploring using the timestamp literal values to measure one-way link latencies of individual links. As I move closer to graduating, this complete work will constitute my senior thesis, fulfilling my the requirements for full college honors within my bachelor's degree in Computer Science.

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. Reverse Traceroute seeks to identify reverse paths (destination to source), an issue that has remained difficult despite the ease of discovering forward paths (source to destination) since the invention of simple traceroute in the early 80s. Timestamps are one of several tools used to measure reverse hops within Reverse Traceroute. However, the rudimentary initial implementation was unsuccessful for several relatively common cases, thus limiting it's effectiveness. Using my knowledge of the various timestamp behaviors, I was able to make several recommendations to improve the Reverse Traceroute algorithm with regard to timestamps. Most significant among the limitations were unsolicited timestamps in easily identifiable cases, introducing a threat of falsely inferred links. Another limitation was that the initial technique required two routers to respond to timestamps, limiting coverage when the first machine was unresponsive, despite the fact that second address might still provide timestamp values. The new algorithm overcomes both of those limitations, and also includes other optimizations which decrease the number of required probes. My contributions were included in the final work, which is currently under submission for publication.

Alias Resolution. Making use of multiple address requests in timestamp probes, I looked towards IP alias resolution as a potential application for timestamps. IP alias resolution is a frequent problem for those who may have information about two different IP addresses, but don't know whether the addresses actually belongs to two different machines, or one machine with two addresses. I discovered several distinct behaviors which routers demonstrated when requested for timestamps from two of their addresses in the same probe. Subsequently, I came up with three distinct techniques for the identification of alias pairs, each of which addressed a different implementation of the timestamp option. I then generated a new dataset of aliases, and both validated and compared against the leading alias resolution technique. I identified thousands of aliases that were previously undiscovered. As the next generation of alias datasets are moving to combining several techniques, I believe that timestamps are an important step towards a comprehensive alias resolution system.

One-Way Delay. I am currently invistigating using timestamp measurements to correctly identify the one-way latency of a single backbone link. This can potentially be acheived by sending a timestamp request which traverses the link, specifying timestamps from each machine on either end of the connection. While this work is just beginning, interesting challenges to the project are the variety of responses that routers provide when encountering a packet in transit, and independent clock skews for each router.

Future Goals

Continuing with my studies, I intend to explore more broadly the many issues that face the Internet. How can we locate and diagnose faults in the network from the edge of the system? Can we ensure that the system is being used fairly by all competing parties? What does "fair" even mean? How can we use models of the Internet as a whole to optimize decisions made by end-hosts utilizing the network? Can we make upgrades or interventions to the network without detrimenting the applications that depend on it? Questions of this sort are appealing to me because they are challenging technical problems, and because they have great impact on the entire network and it's countless users worldwide.

UC Berkeley, with it's exceptional research in Networking, would be an ideal environment for me to pursue these interests. With my background and aspirations, I believe I will be able to make valuable contributions to the program.

Thank you for considering my application.