How's My Network? A Java Approach to Home Network Measurement

Alan Ritacco, Craig Wills, and Mark Claypool
Computer Science Department
Worcester Polytechnic Institute
Worcester, MA 01609, USA
{Alan.Ritacco, cew, claypool}@wpi.edu

***Abstract*—**
As part of a project to develop a user-centered network measurement platform that limits impediments to participation, this work focuses on using the execution of a signed Java applet for home network measurement. We have developed a Java applet tool to understand the capabilities of such a tool for measuring characteristics of a user's network environment from the browser. This paper reports on the capabilities of the tool, the measurement methodology employed, and initial results obtained for a set of residential users employing the tool. Despite the sandbox-type restrictions in Java, the results include information about the configuration of the user's testing machine, wireless connectivity of the testing machine, available upload and download throughput, DNS performance and the number and type of devices on the user's network.

I. Introduction

Traditionally, Internet measurement has been done from points in the network infrastructure or from well-connected research labs and universities. However, with the dramatic growth in Internet access from residences and out in public, often hidden behind Network Address Translation (NAT) boxes, the old measurement paradigm increasingly excludes the performance vantage points seen by the majority of Internet users. The size of this cadre of "invisible" Internet users is increasing as public wireless networking becomes more commonplace and home networking spreads further through the developing world.

The need for a new network measurement paradigm focusing on where users live and their specific interactions with the Internet has already been recognized. One outcome of the Community-Oriented Network Measurement Infrastructure (CONMI) Workshop Report was that "@home-style measurement" is needed to increase the number of Internet vantage points [5]. Desirable outcomes from a recent NSF Computing Infrastructure session on testing for the new Internet [13] include better representation of the user population, non-Linux performance tests and a "SETI@home" type mechanism for networking. Previous work [4] laments the widening gap between measurements for the visible and largely invisible portions of the Internet community motivating the need for "attractors" to provide incentives for user participation in measurement.

While existing network measurement platforms have several desirable features, they do not satisfy these needs. Platforms such as PlanetLab [3], [11], [17] and Archipelago [2], [8] provide flexibility for researchers in choosing metrics to collect, but their platform nodes are permanent, immobile and within a dedicated infrastructure. Alternative platforms such as NETI@home [15], DIMES [14] and DipZoom [18] allow measurements from any node in the Internet, but the scopes of their measurements are limited with currently little incentive for the general populace to participate. Finally, Gomez [7] and a variety of "speedtest" services [6], [16] include limited incentives for broader user participation, but are not designed to inform network research.

The work presented in this paper is part of a larger project to develop a *user-centered* network measurement platform, called *How's My Network (HMN)*, which provides incentives via games and feedback on application performance, while limiting impediments so that the public perceives benefits in participation. Our initial work has focused on what performance measures can be obtained via a Web browser, which is a low-impediment platform for a wide variety of users. One of our projects examined network performance measures obtainable via JavaScript and Flash [9], while this work focuses on using the execution of a signed Java applet for home network (HN) measurement.

We have developed a HMN Java applet tool to better understand the capabilities of Java applets for measuring characteristics of a user's network environment from within a Web browser. This paper reports on the capabilities of the tool, the measurement methodology employed and initial results obtained for a set of residential users that used the tool. The results include information about the configuration of the user's testing machine, wireless connectivity of the testing machine, available upload and download throughput, DNS performance and the number and type of devices on the user's network. This work is important because it demonstrates that a wealth of information about the home network environment can be obtained via the ubiquitous Web browser and the work establishes a basis for understanding long-term trends in this domain.

This work makes a number of contributions for home network measurement:

1) demonstration of the viability of the Web browser for obtaining network performance information;
2) discovery of information about wired versus wireless

connectivity of home machines;
3) the ability to learn about the number and types of networked devices on residential networks;
4) using JDIG, a Java-based DNS tool we developed, the ability to measure DNS performance obtained by users in a home network;
5) the ability to integrate upload and download throughput results, comparable to existing speedtest services, with other measures; and
6) a core measurement component of a future user-centric network measurement platform with incentives for user participation via feedback on applications and servers of interest to users.

The rest of this paper is organized as follows: Section II outlines research questions of interest for this work; Section III describes our testing framework; Section IV describes the methodology used for obtaining the information; Section V provides details on the study; Section VI presents the results; and Sections VII and VIII conclude with directions for future work and a summary of this work, respectively.

II. Research Questions

The activities associated with the HMN platform can answer a number of research questions about home networks. The following list enumerates some of these questions that focused on in this work.

1) What is the nature of the home machine in which tests are run?
2) Do home machines use wireless connections and, if so, what can be learned about their wireless profiles?
3) What is the performance of networked applications running in a home environment?
4) What is the performance of DNS and what is its influence on application performance?
5) What is the nature of the home environment? What are the number and types of network devices in the home network?

III. Testing Framework

A testing framework tool was developed to allow for testing modules to be easily added. The specific testing modules used in this work are described in the following section. In addition, a custom client/server environment was designed to capture results from each test and store the results at the server for later analysis.

With the HMN-testing framework, our testing suite was provided to users via a self-signed Java applet. This approach allows a user's Web browser to execute our testing via the Java Runtime Environment (JRE). The signed applet provides range of access beyond the traditional Web browser sandbox level access and control, but is still constrained by the JRE. This approach is beneficial for users as it allows participation while not requiring the installation of any additional software on a user's machine.

The HMN testing suite is comprised of a simple graphical user interface with a single "Start" button along with a refresh time. Once the application is running it is set by default to repeat execution every five minutes. Re-execution of the tests has shown to be valuable for providing both longer-term information for the user and in for data in our repository. All data is stored based on the Internet Protocol (IP) address, although the use of cookies in the future can help correlate multiple tests and to anonymize results when made available to others. During the loading of the HMN applet a security certificate requests the user to accept the applications digital signature (the signed applet). Accepting the certificate allows the applet to perform operations outside of the sandbox. A user who does not accept the certificate will run the applet in the browser sandbox limiting its capabilities to learning some information about the testing machine as well as the upload and download throughput to the origin server.

IV. Methodology

The HMN applet used in this work consists of five modules that each obtain distinct types of information about a user's testing machine and networked environment. Each module is designed to obtain information about a research question posed in Section II. These modules are run in phases as described in the following.

A. Test Configuration

Configuration information is first gathered about the machine performing the test. This information includes the type of browser and operating system, the internal and external IPs employed by the machine including whether the machine resides on a non-routable network (behind a NAT box) [12], and address of the primary DNS server. Local system property information such as Java version, paths, architecture type, CPU type and processor speed are also obtained.

B. Wireless Connectivity

Two types of information about a user's wireless connectivity are obtained. First, by querying the network configuration information the applet is able to determine whether the machine is networked via a wired or wireless connection. Second, when available, the applet can obtain the types of wireless network profiles employed by a user. Although not available for this set of tests, a future module is being developed to obtain the number and signal strength of wireless access points in range of a test machine.

C. Upload/Download Throughput

As a measure for comparison with other testing tools another model included a module measures upload and download throughput over TCP between the testing client and our origin server. This test provides a baseline for comparison with existing tools. Future modules will include similar tests to non-origin servers, which are allowed from a signed applet, where the chosen servers can be customized based on user preferences.

D. DNS Performance

DNS performance continues to be an important, if overlooked, aspect of service provided to home users. In order to test DNS performance we created a Java-based tool, called *JDIG*, with an interface similar to the public domain *dig* tool. Our tool can run as a standalone Java application, but for our work it is packaged as part of a module. JDIG performs a variety of DNS tests including the round-trip time (RTT) for obtaining a cached DNS entry. The tool also measures the average DNS RTT for a random set of *.edu* servers, the average DNS RTT for a set of popular servers [1] and the RTT to obtain a top-level domain (TLD) and generic TLD (gTLD) entry.

E. Local Network Environments

The final module in our set of tests determines information about the number and types of devices on the local network of the user's testing machine. This module is only invoked on networks with non-routable IP addresses that are behind a NAT box. The first step performed by this test is to determine the number of active devices on the network. It does so by issuing an ICMP request for the 255 IP addresses obtained by varying the low-order byte of the testing machine's IP address. A thread-based parallel can completes in 10-20 seconds. While not all active devices reply to the ICMP request an underlying ARP request causes a reply for each valid IP address with the device's corresponding MAC address.

Once the scan of IP addresses is complete, the list and count of active devices is obtained by consulting the ARP table on the test machine. The type of each device is determined in two ways. A manufacturer of each device is obtained by matching the device MAC address with ranges assigned to manufacturers as done in *nmap* [10]. This approach works to determine special-purpose devices such as printers or game consoles. For general-purpose computers selected ports are scanned to fingerprint the type of operating system the machine is likely using.

V. Study

The Java testing applet resides on a quad-core server with 8GB of RAM running Linux located on WPI's campus network. The applet is downloaded via an Apache Web server. The server is also used for logging and throughput experiments. As experiments could be run at any time, a timestamp was created on the server for each applet result.

Once the applet was deployed, users on and off campus were invited to participate in testing. A total of 50 users (based on unique IP addresses) participated in the December 2008 timeframe. Using reverse DNS mappings each IP address was classified according to a commercial company, an educational institution or an ISP known to provide service to residences and public hot spots. Because our immediate focus is on residential and public users, tests from commercial and educational sites are not reported in this paper. Thus, the results from 36 residential and public hot spot users are analyzed in this work. Based on the reverse DNS names all of these users are in the northeastern U.S. and can be classified into four ISPs, as shown in Table I. Two of these ISPs are known to provide cable modem service, one provides DSL and one provides fiber optic service (FIOS).

TABLE I
ISPs of Home User Tests Participating in Study

Provider	# Users	# Sessions	# Tests
Cable1	12	13	106
Cable2	12	25	109
DSL	6	9	23
FIOS	6	12	33
Total	36	59	271

The third column in Table I shows 59 unique sessions performed by our 36 users where additional sessions occur when the same user initiates the testing applet more than once. Finally, because the applet automatically re-executes its test after a five-minute sleep period, multiple tests are run within a session if the user allows the applet to remain active. Table I shows that a total of 271 tests were performed by our set of residential users.

All users accepted the digital certificate of the applet so in all cases it executed with signed applet privileges allowing for the full-range of data collection described in Section IV. A small number of data collection errors occurred because of insufficient local privileges even when using a signed Java Applet. These errors were specific to security privileges required by Windows-based operating systems, and occurred on non-residential networks so do not impact the results reported in this paper. All phases typically take on the order of 40 seconds to execute within a user's browser.

VI. Results

This section reports the results obtained by residential users in our study set for each of the five modules described in Section IV.

A. Testing Configuration

A summary of the testing configuration results for our 36 users are shown in Table II. More than 94% of HMN residential users ran a Windows-based OS and of these users 45% were running Windows Vista. 61% of these users employed Internet Explorer as their browser while the remaining 39% of users employed Firefox. A small number of users with Linux-based systems ran our HMN tests and all of these users employed Firefox as their Web browser.

TABLE II
Testing Configuration Highlights for Residential Users

94% of users run a Windows-based Operating System
45% of Windows-based OS are Windows Vista
61% of users run tests with Internet Explorer
39% of users run tests with Firefox
100% of users have a non-routable internal IP address
97% of users have a DHCP-assigned external IP address
47% of users have a non-routable primary DNS server

The internal (used by the testing machine) and external (used by the access point) IP addresses used by the residential

testing platforms were examined. All home users running HMN had a non-routable internal IP address for their machine, meaning that the access point was using NAT. Based on examination of the reverse DNS name, almost all external IP addresses (97%) of were assigned by the ISP using DHCP with only 3% of HMN users with a static IP address

In looking at the DNS configuration, for 47% of HN users, the primary DNS server resided on a router/switch in the HN. Based on experience, these servers typically do not cache results, but simply "pass through" DNS requests to a caching DNS server managed by the ISP. All of the FIOS users employed such a DNS cache with mixed usage by users of the other ISPs.

B. Wireless Connectivity

Accessing the local network configuration information of the testing machine shows 38% of users ran from a wireless connected machine, while the remaining 62% used a wired PC.

Examining clients wireless caches shows that 56% of the testing machines have attached to at least one wireless network at some point in time. These wireless users have connected to five or more wireless networks at some time. These network types are in the range of: home, business, resort, and hot spot WiFi locations. Our set of users have attached to a total of 96 unique wireless networks. From the data, the most popular wireless networks are from Xerox, Cisco, and DLINK. Fifty-five unique types of hardware manufacturers used, with some overlap due to the convergence of MAC address space.

C. Upload/Download

As a measure of a user's connection performance upload and download throughput is determined to the server at WPI. Figure 1 shows a scatter plot of upload/download characteristics for each of the 36 HN users where each point is the average of all throughput tests for the given user.

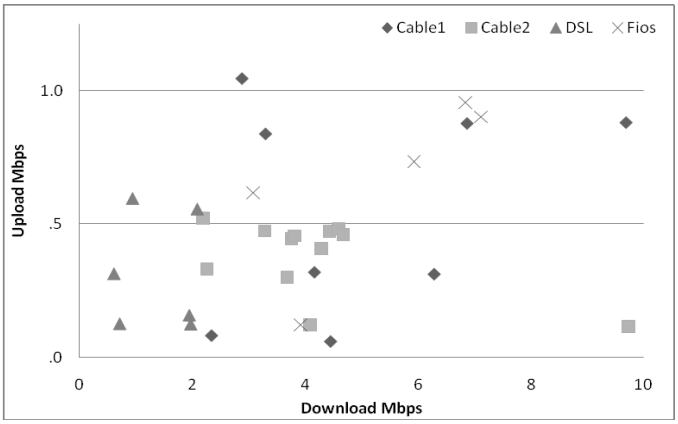


Fig. 1. Scatter Plot of Average Upload and Download for Users by ISP

Each point in Figure 1 is characterized by the service provider from Table I. Most of our HMN users fell in the 5Mb or less category for download and upload combined properties. The fastest upload/download throughputs were those using Cable1 and FIOS service providers. The distribution in Figure 1 shows how users with the same ISP have similar properties. All HN users have an asymmetric Internet connection where the download throughput is more than the upload throughput.

Figures 2 and 3 show a cumulative distribution function (CDF) for the download and upload throughputs of all 271 tests by our 36 users. The results in Figure 2 show variation amongst the download throughput, with DSL providing the lowest download throughput while FIOS and Cable2 providing higher download throughput and Cable1 providing the highest download throughput in our tests. On the other hand, the upload throughput values in Figure 3 show less distribution, with DSL and Cable2 users never receiving more than 0.5 Mbps in upload throughput. This clearly is a situation where users pay for download speeds and get nominal and/or obligatory upload speeds. ISPs can oversubscribe data lines by allowing lower bandwidth for upload than download.

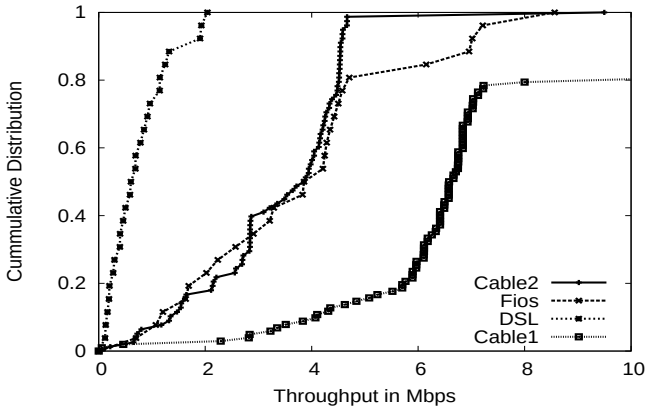


Fig. 2. CDF of All Download Throughput Tests by ISP

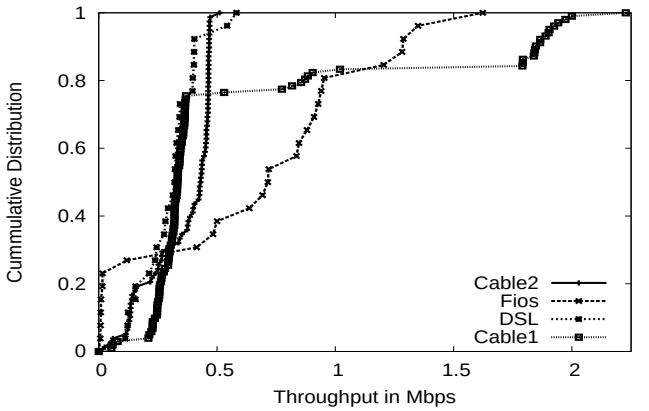
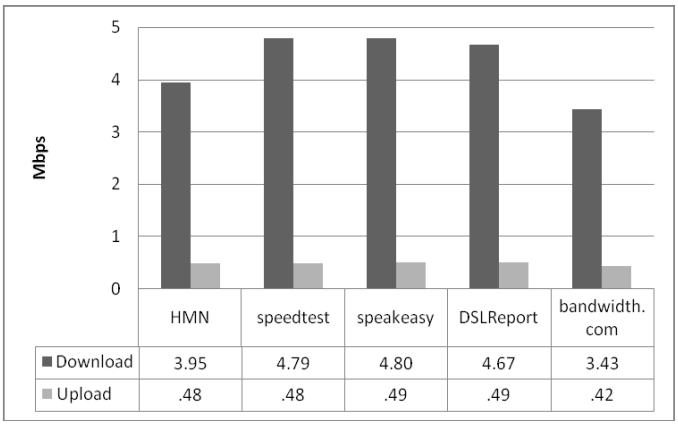


Fig. 3. CDF of All Upload Throughput Tests by ISP

Figure 4 shows the performance of popular speed testing services versus our HMN tests. The following speed testing services: speedtest.com, speakeasy.com, DSLreports.com, and bandwidth.com. Figure 4 shows the representative upload and download throughput obtained for each service, each run from

the same HN computer and cable provider. The HN system used has a known speed of 5Mb download and 512Kb upload. In each case, the nearest server was chosen for each of the speed testing services. The HMN results are similar to those of other speed testing services. While this is a sampling of data for one home network, similar results were found for other home network tests.

The amount of data each sent through for the download and upload measurements was examined using sniffer traces and other data analysis. HMN sent 1MB for download and 512KB for upload, speedtest.net sent 4.5MB for download, and 460KB for upload, DSLreports.com sent over 6MB of data for download, and over 800KB of data for download, bandwidth.com sent4MB for download and 550 KB for upload.



	HMN	speedtest	speakeasy	DSLReport	bandwidth.com
Download	3.95	4.79	4.80	4.67	3.43
Upload	.48	.48	.49	.49	.42

Fig. 4. Throughput of HMN vs. Popular Speed Testing Services

D. DNS Performance

The availability of the JDIG tool within our test suite allows us to test DNS performance obtained by each of our users. Since our focus is on the DNS performance provided by the local DNS server of the ISP, the JDIG tool does not use OS resolver routines and therefore bypasses any OS-specific DNS caches, such as are present on Windows-based Operating System machines. As described in Section IV four types of DNS performance tests are conducted:

The first DNS test retrieved the A record for a server name then did a subsequent retrieval for the same name to measure the lookup time for a cached entry. Even for cases where the local network access point was configured as the primary DNS server, these requests are still passed through to a local DNS server of the ISP, which is caching the results of previous queries. Figure 5 shows the RTT results for cached queries of the 59 sessions in Table I. The results in the figure show that the median lookup time for three of the ISPs is on the order of 20ms, although users of the Cable1 ISP typical provide the worst cached DNS performance. Worst case results are on the order of 150ms.

The DNS performance for a set of unlikely used servers is examined next. The servers are compiled from a list of 4000+

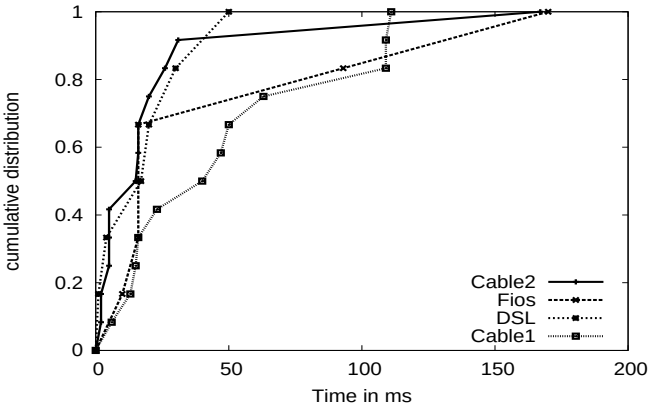


Fig. 5. CDF of DNS Cached Entry RTT per ISP

.edu sites. From this list, 25 random DNS requests are used as part of each user test. The average RTT is determined for the first test within each session with a CDF of these averaged results shown in Figure 6. As expected, these results show much higher RTTs than for the cached results of Figure 5 with median values between 100 and 150ms for the ISPs. 10-20% of the average values are over 200ms indicating much larger individual lookup times.

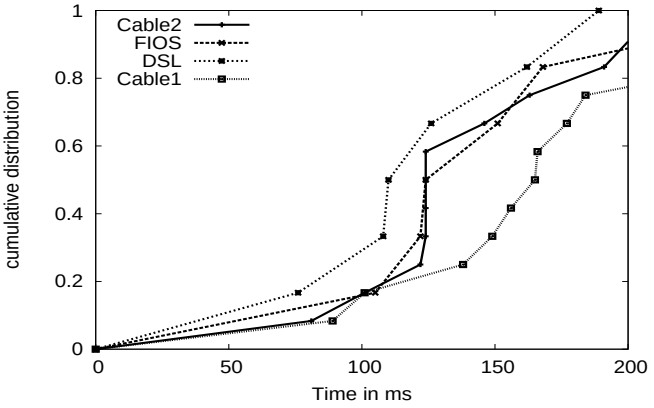


Fig. 6. CDF of Average DNS RTT for 25 Random DNS Queries per ISP

The DNS lookup times for 100 popular Web sites [1] was examined, with a CDF of the average for these results shown in Figure 7. Figure 7 indicates that many of these entries were already cached on the DNS servers as the median times are near the values for cached entries shown in Figure 5. Despite the relatively low lookup time in most cases, Figure 7 shows a small number of cases where the average across 100 servers is still large, again indicating some much larger individual lookup times. These results at the upper end require further study when all individual DNS results are recorded.

The final set of tests examined the performance of the local DNS servers to look up top-level, such as .com, and second-level, such as wpi.edu, domain names. These tests were conducted by generating invalid first- and second-level domain names that force a lookup to a root and a gTLD domain

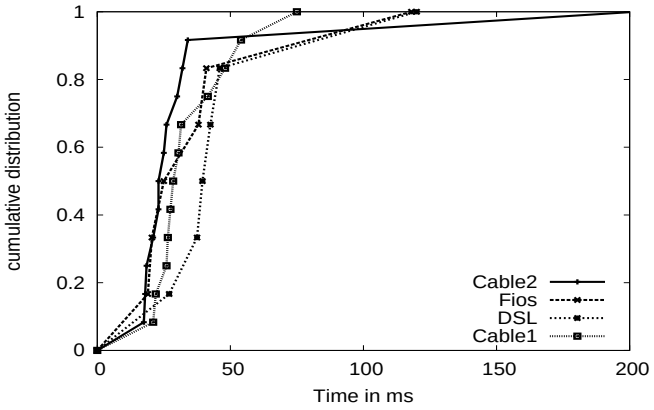


Fig. 7. CDF of Average DNS RTT for Top 100 Queries per ISP

server. Figures 8 and 9 show CDF results for first- and second-level domain requests. These results show that the RTT from the client to the TLD and gTLD DNS servers is less than 100ms in most cases. While it is expected that TLD servers are not frequently queried, the gTLD servers must be queried for each new domain name that is encountered so performance is important.

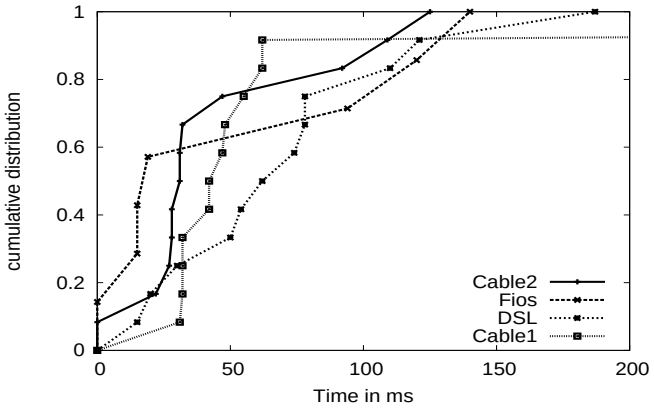


Fig. 8. CDF of First-Level Domain RTT per ISP

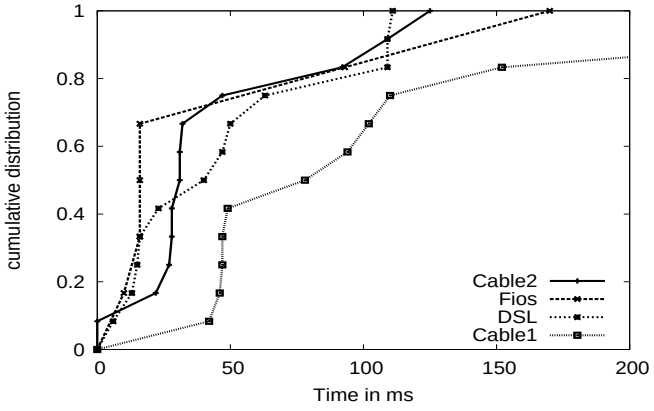


Fig. 9. CDF of Second-Level Domain RTT per ISP

E. Local Network Environments

The last set of results use the methodology described in Section IV to determine the number and type of devices on the 36 residential networks in our study. Home users have an average of three active devices. The typical scan returned the following devices: PC, router (where Cisco is the most popular brand), and a broadband modem (again where Cisco is the most popular brand.) Other HN devices detected ranged from gaming consoles (Nintendo Wii, PS3, etc.), video recording boxes (TiVo, Slingbox, etc), printers, hardware-based routers and switches (Cisco, 3Com), along with Windows and Linux-based PCs. Figure 10 shows the distribution of the most popular system types found (based on OS (Windows/Linux), and networking hardware).

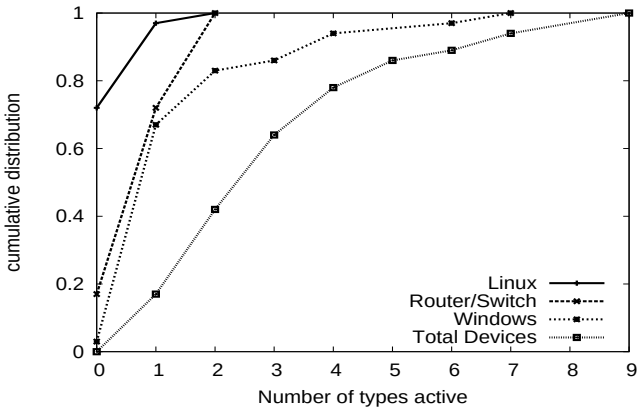


Fig. 10. CDF of host types found during HMN scans

Table III shows the same data about devices as a percentage of the total number of devices and the total number of users. The results show that 52% of the devices are machines running a Windows OS with such a device present in 97% of networks for our users. Smaller percentages were found for specialized devices such as game consoles, digital-video recorders (i.e. Tivo) and printers, although these numbers are likely conservative as devices need to be active at the time of the tests in order to be detected.

TABLE III
DEVICES OF HOME USERS PARTICIPATING IN STUDY

Device Type	% Devices	% Users
Windows Machine	52	97
Network Device	33	83
Linux Machine	9	28
Game Console	2	6
Tivo	2	6
Printer	2	6

VII. FUTURE WORK

While the methodology developed and the results obtained show promise, they indicate a number of directions for future work both in terms of methodology and approaches for improving home network performance.

Having a broad representation of home users participating as network measurement points is the primary goal of our overall project. While the work described in this paper is primarily about what measurements can be obtained, the next part of our project seeks to put results in the context of their impact on applications of interest to users. A level of trust needs to be built with users as the use of signed applets not only allows a range of network measures, but can legitimately raise security concerns on the part of users.

In terms of methodology, baseline testing can be extended to not only include upload and download throughput to our the origin server, but also to allow tests to include other third-party servers that may be of particular interest to the user performing the test. Thus, rather than a generic throughput test, throughput to a particular server can be tailored for the user, thus serving as an incentive for user participation. Baseline tests are also to be extended to include RTT measurements, which were not explicitly measured in our initial set of tests.

Studies of DNS performance provided by ISPs using our JDIG tool will be extended. Recording the RTT of individual DNS lookups will allow not only better understanding the average, but often more importantly, the worst-case DNS performance.

An ongoing area of interest is to understand the wireless network configuration of public and residential users. This work was able to obtain wireless connectivity and profile information, but additional work has begun for the applet to obtain the number and signal strength of wireless access points in the range of a test machine.

A graphing software module for the Java HMN applet is planned. This feature will provide visual representation of data from previous users and current data for the user in real-time. Better visualization of results will also be an attraction for users, particularly as the test repeats itself over time.

A feature to assist with HMN discovery and repair procedure will be implemented for various aspects of networking and systems-level performance enhancements. These enhancements can be as simple as changing registry or file settings, or suggesting different hardware scenarios for networking advantages.

The results of our initial DNS testing work also point to potential improvements in home DNS performance. A local DNS server that caches results could potentially improve performance for servers regularly visited by users. Such a local DNS server could also provide a pre-fetch cache service, as needed when the TTL expires, for the most popular entries of a given environment. This enhancement would be a simple setup for the server and actually reduce the amount of outgoing traffic as the system could maintain a link connectivity status and during non-peak times the updates would occur, thus saving local and external throughput.

VIII. Summary

In this work, we have built and demonstrated the capabilities of using a signed Java applet for network measurement. This approach is particularly appealing for home network measurement as a signed Java applet can be run on a user's machine without permanent installation of any software. In combination with user-oriented feedback this approach can broaden the set of users employing such a tool and allow researchers to gain valuable insight into home network environments.

Our initial work has successfully deployed an applet and used it to gain information about the configuration of a user's testing machine, wireless connectivity, available upload and download throughput, DNS performance and the number and type of devices on the user's network.

Results from an initial set of users both provide data about wireless connectivity of home network environments as well as the number and type of devices on these networks. The results also allow comparison of upload/download throughput and client DNS performance.

References

[1] Alexa top 500 sites. http://www.alexa.com/site/ds/top_sites.

[2] Archipelago measurement infrastructure. `http://www.caida.org/projects/ark/`.

[3] A. Bavier, L. Peterson, M. Wawrzoniak, S. Karlin, T. Spalink, T. Roscoe, D. Culler, B. Chun, and M. Bowman. Operating system support for planetary-scale network services. In *USENIX Symposium on NSDI*, San Francisco, CA, USA, Mar. 2004.

[4] M. Casado, T. Garfinkel, W. Cui, V. Paxson, and S. Savage. Opportunistic measurement: Extracting insight from spurious traffic. In *Proceedings of the Fourth Workshop on Hot Topics in Networks*, College Park, MD USA, November 2005.

[5] k. claffy, M. Crovella, T. Friedman, C. Shannon, and N. Spring. Community-oriented network measurement infrastructure (CONMI) report. *SIGCOMM CCR*, 36(2), 2006.

[6] Broadband reports.com speed test. `http://www.dslreports.com/stest`.

[7] Gomez peer community. `http://www.porivo.com/`.

[8] Y. Hyun. The Archipelago measurement infrastructure. In *Proceedings of the CAIDA-WIDE Workshop*, Nov 2006.

[9] A. Janc, C. E. Wills, and M. Claypool. Network performance evaluation within the web browser sandbox, February 2009. Submitted for publication.

[10] Nmap - free security scanner for network exploration & security audits. http://nmap.org.

[11] L. Peterson, A. Bavier, M. E. Fiuczynski, and S. Muir. Experiences building PlanetLab. In *Proceedings of the 7th USENIX SOSDI*, Seattle, WA USA, November 2006.

[12] Y. Rekhter, B. Moskowitz, D. Karrenberg, G. de Groot, and E. Lear. Address allocation for private internets, February 1996. RFC 1918.

[13] H. Schulzrine. The new Internet – goals, testing and infrastructural needs. In *Breakout Session of the NSF Computing Research Infrastructure PI Meeting*, Boston, MA USA, June 2007.

[14] Y. Shavitt and E. Shir. DIMES: let the Internet measure itself, 2005. `http://www.arxiv.org/abs/cs/0506099v1`.

[15] C. R. Simpson, Jr., D. Reddy, and G. F. Riley. Empirical models of TCP and UDP end-user network traffic from NETI@home data analysis. In *In PADS*, pages 166–174, May 2006.

[16] Speedtest.net. `http://www.speedtest.net/`.

[17] N. Spring, L. Peterson, A. Bavier, and V. Pai. Using planetlab for network research: myths, realities, and best practices. *SIGOPS Oper. Syst. Rev.*, 40(1):17–24, 2006.

[18] Z. Wen, S. Triukose, and M. Rabinovich. Facilitating Focused Internet Measurements. In *Proceedings of the ACM SIGMETRICS*, New York, NY, USA, June 2007. ACM Press.