EmNet: Satisfying The Individual User Through Empathic Home Networks

J. Scott Miller John R. Lange Peter A. Dinda
{jeffrey-miller,jarusl,pdinda}@northwestern.edu
Northwestern University

Abstract—**We consider optimizing the control of the wide-area link of home routers based on the needs of *individual users* instead of assuming a *canonical user.* A careful user study clearly demonstrates that measured end-user satisfaction with a given set of home network conditions is highly variable—user perception and opinion of acceptable network performance is very different from user to user. To exploit this fact we design, implement, and evaluate a prototype system, EmNet, that incorporates direct user feedback from a simple user interface layered over existing web content. This feedback is used to dynamically configure a weighted fair queuing (WFQ) scheduler on the wide-area link. We evaluate EmNet in terms of the measured satisfaction of end-users, and in terms of the bandwidth required. We compare EmNet with an uncontrolled link (the common case today), as well as with statically configured WFQ scheduling. On average, EmNet is able to increase overall user satisfaction by 20% over the uncontrolled network and by 12% over static WFQ. EmNet does so by only increasing the average application bandwidth by 6% over the static WFQ scheduler.**

I. INTRODUCTION

Home networks are challenging and complex environments [1]. Hosts and devices within the home network talk to the Internet via a broadband router, where the upstream link typically has much lower bandwidth than either the home network or the rest of the path through the Internet [2], [3]. The size of home networks in terms of devices and users is growing, especially given the increased usage of wireless. Home users are also increasingly running applications that use many long-lived and high-bandwidth flows alongside their interactive applications that strain the bottleneck upstream link. Section II further describes the issues in home networks; the core problem we address here is: *How should we schedule the upstream link of a home network in order to provide a high degree of user satisfaction for interactive applications while providing excellent service for the other applications?*

We investigated an approach to solving this problem that uses direct user feedback to optimize for the individual users and exploit the variation among them. It has been demonstrated in contexts outside of networks, such as CPU scheduling and power management, that actual measured user satisfaction with any given operating point exhibits considerable variability across users [4], [5], [6]. This introduces the question: *Does user satisfaction with network performance for interactive applications also exhibit this variation?* Our results suggest the answer is yes.

Measuring *individual user satisfaction* online and then using such measurements directly in the control process makes it possible to exploit this variation to the mutual benefit of users and systems. This result has been demonstrated through user studies in systems-level control as diverse as power management, scheduling, and remote display systems [7], [8], [9], [10]. Further, it is becoming increasingly clear that such measurement can be done with minimal intrusion on the user through the use of biometric information [11], [12], [13]. This introduces the question: *Does it make sense to leverage direct user feedback from individual users in controlling home networks?* Our results suggest that it does, and we have developed a prototype for doing so.

Our work started with a carefully constructed user study in which participants used a range of web-based and other interactive applications while we varied the nature and degree of non-interactive cross-traffic. Our participants represented a broad spectrum of individuals at Northwestern University. Section III describes our study in detail.

Although our study includes numerous qualitative and quantitative contributions, one result is clearly the most important: There is a large degree of variation in expressed user satisfaction for identical application/cross-traffic scenarios. This holds across the wide range of application/cross-traffic scenarios we studied. It is an inescapable conclusion that users react in distinct and highly individual ways to cross-traffic. It is not the case that a per-application utility model or other aggregate captures this variation.

Optimizing a home network by using an objective function that does not express this variation among individual users misses two opportunities. First, if the objective function is based on a canonical user who exhibits the mean or median user satisfaction, it is likely to result in a significant number of interactive users being far more dissatisfied than they need to be. Second, if the objective function is based on a canonical user who exhibits satisfaction at a deviation or more above the mean/median, it is likely to result in many users being significantly over-provisioned, leading to lower performance for non-interactive flows than is necessary.

Having found opportunity in this variation across individual users, we next designed and implemented a system, EmNet, that provides a mechanism for individual user feedback for web-based and other interactive applications. The system, which is described in detail in Section IV, schedules the outgoing link using weighted fair queuing [14], [15]. Non-interactive traffic is given a specific weight, and the weight of each user's traffic has upper and lower bounds. The user can at any time adjust a slider control that is overlaid on top of their current web page or desktop environment. By raising the slider, the user increases the weight of her traffic, but at a cost. The slider control displays the current monetary cost of the user's bandwidth usage and weight/provisioning.

We evaluated EmNet through an extensive user study on a different set of participants from a wide range of backgrounds. Our evaluation is presented in detail in Section V. The most important results are the following.

- EmNet increased the average user satisfaction by 20% over an unprovisioned network and by 12% over a statically provisioned network.
- EmNet's increase in satisfaction came at the cost of only a 6% increase in application bandwidth compared to a statically provisioned network.

These results suggest that EmNet successfully allows individual users to personalize their home network performance, trading off perceived satisfaction and cost. This personalization is beneficial both for the individual users and for the network as a whole.

The lessons learned in our study of user satisfaction with home network performance, and from our EmNet system, suggest that measurements of individual user satisfaction, or other individual user feedback, can be effectively incorporated into network processing, providing significant benefits.

II. Challenges of home networks

We characterize a home network as a small collection of end-hosts connected by a high-speed wired or wireless network and sharing a residential broadband connection (e.g., DSL, Cable, etc.). Home networks are ubiquitous: in 2007, the Consumer Electronics Association reported that 30% of households in the United States own home networking equipment [16]. The Pew Internet and American Life Project found that in 2008, 55% of adult Americans have broadband Internet access at home [17]. It is safe to say that the quality of the application-level experience of millions of Internet users is in some way influenced by the performance of a home network.

HomeMaestro: HomeMaestro [1] seeks to mitigate home network contention by using a distributed approach to identify competing flows and allocate network resources fairly to end-hosts. To motivate the need for special intervention, the authors conduct a study of several representative home networks, collecting network traces along with user accounts of dissatisfaction. Surprisingly, the households report daily issues with the network across a range of applications. Their system, HomeMaestro, is able to provide approximately fair allocations according to a set of application level weights. The setting of the application weights is left as an open question and the author's acknowledge that these weights are likely to be user and application dependent.

We demonstrate one possible way such weights can be derived directly from end-user feedback.

OneClick: Recently, Tu et. al. [18], [19] described OneClick, a system designed to inexpensively measure *mean* satisfaction level, the Mean Opinion Score concept from various ITU recommendations on digital audio and video. The system allows users to easily signal dissatisfaction with network performance. This feedback is then used to construct bounds on network-level quality of service metrics required for acceptable application-level performance. The approach can be applied to applications in which it is difficult to quantify the effects of the network on performance. While the authors' analysis point out some of the challenges of collecting user satisfaction levels, the individual user feedback is largely ignored.

Our work characterizes the extent to which *variation* across users can be seen in user satisfaction with home network conditions. We also demonstrate how to leverage this variation in controlling the home network to improve overall satisfaction.

III. Variation in user satisfaction

Given the results in other domains mentioned in Section I, a natural question is whether actual measured user satisfaction given particular network conditions also exhibits significant variation. To answer this question in the context of home networks, we conducted a rigorous user study in which a wide range of users participated. Each participant was asked to use and rate their satisfaction with a set of network applications on an emulated home network where a variety of cross-traffic scenarios were applied. We found the question could easily be answered in the affirmative: Variation in user satisfaction is large. Furthermore, this variation is not explained by variation in application-level QoS measures.

A. Instrumentation

We created a user interface that can be transparently applied to existing web applications.[1] We used a web proxy to modify any existing pages by injecting our own instrumentation before those pages are sent to the client. The user need only point her existing web-browser to the proxy server in order to receive the instrumented interface. The code added to the client page exists in a single JavaScript closure and, aside from the UI elements added to the page's document object model, shares no global state with other code on the page. The UI communicates with the proxy server using asynchronous callbacks to well-known control URLs that are intercepted by the proxy server.

We use this instrumentation in the study described in this section and in the study described in Section V. For the present study, we created a simple user interface that asks the user to give their satisfaction with the network performance during the preceding 30 seconds. The prompt is placed over the current web document on top of an obscuring layer that prevents the user from browsing while the prompt is visible. We have control over when the prompt is presented via an API exposed by the proxy server. The prompt does not disappear until the user has selected a satisfaction level.

[1] Further implementation specifics of our work can be found elsewhere [20].

Application	Familiarity	% Usage
Peer-to-peer downloading	5.05	83%
Web browsing	6.74	100%
Voice over IP	4.32	67%
Streaming video	5.84	94%
Interactive web apps	6.53	94%
Instant messaging	6.63	100%

Fig. 1. Application familiarity and usage information for our study population. Application familiarity was rated on a 1–7 scale, with 1 meaning "Never Used/Not Familiar" and 7 meaning "Very Familiar". We report usage as the percentage of respondents who use the application at least monthly.

B. Testbed

Our testbed is designed to emulate present-day home-networking configurations with the additional abilities to (a) add instrumentation as described in the previous section, and (b) add cross-traffic.

We emulate a broadband router using a layer 2 Ethernet bridge configured to match the bandwidth characteristics of a typical home broadband connection. The bandwidth on the bridge router is restricted to 3 Mbps downstream and 796 Kbps upstream (i.e., DSL speeds).

We emulate the home user's LAN environment using a 1 gigabit switch which is uplinked to the bridge's internal NIC. Another 1 gigabit switch is placed between the bridge and the Internet link for reasons we will discuss shortly. The uplink port on the external switch was connected to the Northwestern University network for Internet connectivity.

The LAN environment consists of 3 machines: (1) a client machine used by the test subject, (2) a web proxy server that implements instrumentation as described in the previous section, and (3) a cross-traffic generator used to provide configurable background traffic. The client machine is configured to use the web proxy.

The cross-traffic in our testbed is created using a client machine on the internal network and a server machine connected to the switch on the external network. A local cross-traffic server allows us to accurately repeat specific cross-traffic scenarios. The cross-traffic is generated using a modified version of IPerf [21] that provides support for parallel connections, time based traffic generation, and specific bandwidth usage.

The client's operating system is Windows XP and all others run RedHat Enterprise Linux ES 4.7. On the client, browser caching is disabled so that fetches always reach the web proxy.

C. Users, applications, and scenarios

We recruited subjects from the general population within Northwestern University. Formal approval by our Institutional Review Board allowed us to advertise widely across the university using a combination of fliers and email advertisements. Subjects were paid $20 for their time.

The study[2] involved 20 participants including 13 men and 6 women (one participant did not choose to answer the optional demographic questions). Each participant filled out a short survey measuring their use and familiarity with using various applications on a home network. Overall, 19 of the 20 participants have experience using a home network (11 report connecting using a Cable Modem while 8 report DSL). We present a detailed view of the participants experiences in Figure 1. The participants frequently use a wide-range of applications on home networks.

[2]Formal study documents available at http://empathicsystems.org/emnet.

Name	Direction	No. of Flows	Bandwidth
Up.1	Upload	1	100Kb
Up.2	Upload	1	300Kb
Up.3	Upload	1	500Kb
Up.4	Upload	1	768Kb
Up.5	Upload	4	100Kb/con
Up.6	Upload	4	300Kb/con
Up.7	Upload	4	500Kb/con
Down.1	Download	1	100Kb
Down.2	Download	1	500Kb
Down.3	Download	1	1Mb
Down.4	Download	1	3Mb
Down.5	Download	4	100Kb/con
Down.6	Download	4	500Kb/con
Down.7	Download	4	1Mb/con
Mixed.1	Download	4	1Mb/con,
	Upload	4	200Kb/con
Mixed.2	Download	8	1Mb/con,
	Upload	8	200Kb/con

Fig. 2. Cross-traffic scenarios.

Each of our subjects used three different network applications, as explained below.

Wikipedia: Participants are given a set of thirty questions to answer using Wikipedia [22]. The task is representative of common web-browsing, where traffic is aperiodic and bursty.

Image labeler: Participants play the Google Image Labeler game [23], a two-player web-based game in which players assign labels to images found on the Web. Players must agree on a label for an image without any form of direct communication to progress. The application generates small, asynchronous web requests that communicate game state.

Streaming video: Participants watch a streaming video. The video is compressed using the MPEG-4 codec such that the resulting bandwidth averages 2.9Mbps. Both the client and server are running the Video LAN Client (VLC) [24], version 0.8.5. The video is streamed over UDP using the MPEG-TS encapsulation method. The stream is unbuffered, such that packet loss introduces immediate effects that vary depending on the amount of data lost.

We consider 16 different cross-traffic scenarios, which are presented in Figure 2. All scenarios are bulk transfers, that is, we do not explore the effect of competing interactive traffic.

D. Study design

On arrival, each subject began by reading and signing a consent form. After this, they filled out an introductory questionnaire including demographic information and network application / home networking familiarity. Next, they were given a constrained period of time to familiarize themselves with the study setup. They then proceeded to the first application task, chosen at random. They were given a written description and a constrained period to read it. Then, the task would begin. As they worked on the task, each cross-traffic scenario would be applied in a random order for a fixed period of time. At the end of this period, the user was prompted for their satisfaction level. This process repeated for the other two application tasks. A final debriefing then occurred.

E. Results

Figure 3 presents Box plots of the prompted user satisfaction, one graph per application, with the graphs broken down further by cross-traffic scenario. The bold lines indicate the medians, while the boxes extend from the 25^{th} to 75^{th} percentiles.

It is abundantly clear that the data shows that for every application/scenario pairing, there exists substantial variation in satisfaction across users. This variation for the most part swamps the aggregate differences between the scenarios and across applications. The video application (Figure 3(c)) shows the *least* per-scenario variation among the different applications. However, even here the variation among users within each scenario is on par with the variation of averages among the scenarios.

One question the reader may have is whether we are observing the effects of different performance induced by network characteristics on the broader Internet that are not under our control. This is not the case. We measure and analyze the response time at the client side and at the network side (from the proxy's request to the ultimate web server) along with the packet loss rate across the bridge. We find that the vast majority of application performance variation is under our control and that the variation in satisfaction cannot be attributed to variation in the broader Internet.

Our study shows that users have a wide range of expectations of application-level performance. This implies that an optimal network provisioning strategy–that is, one that maximizes the aggregate user satisfaction–is unlikely to be an even or static allocation of network resources.

IV. EmNet system design

Based on the results from our initial study, we designed a system, EmNet, that is capable of shaping network traffic based on individual user satisfaction. We target EmNet specifically at web-based applications but believe that it can easily be extended to other application environments. EmNet presents the user with a throttle overlay that also notes the cost of the current throttle setting. The user can change the throttle setting at any point. The throttle setting is then an input to the link provisioning algorithm running on the edge router. The throttle setting can be associated with a set of one or more flows. Our link provisioning algorithm uses the throttle inputs of the different flow sets and network-wide parameters to derive weights for WFQ scheduling the Internet link.

A. Architecture

EmNet is designed to be fully implemented inside a commodity broadband router. Its architecture is shown in Figure 4. Conceptually, the system can be separated into four components: (1) a user satisfaction sensor, (2) a proxy server that injects the user interface into the user experience and tracks managed traffic, (3) a policy controller that uses a provisioning algorithm to configure the network control policies, and (4) a set of network control mechanisms that implement the control policies.

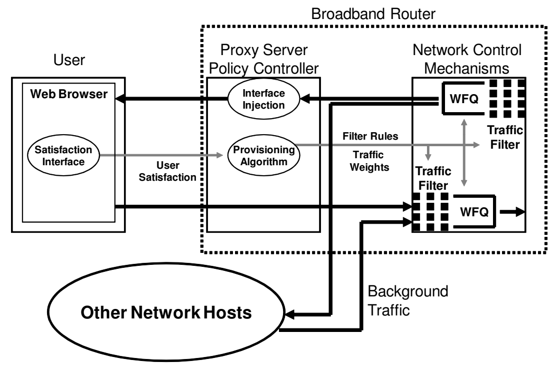


Fig. 4. The EmNet Architecture

B. Satisfaction sensor

The purpose of the satisfaction sensor is to provide periodic measurements of the individual user satisfaction in real time. These sensor readings are used to derive inputs for the policy controller by determining what needs to be optimized to maximize the satisfaction.

EmNet's satisfaction sensor is an on-screen throttle combined with a cost display. We require the user to explicitly gauge their own level of satisfaction and derive the performance values themselves. We can readily incorporate other sensors and are investigating adding the minimally intrusive sensors our group has developed (as noted in the Section I).

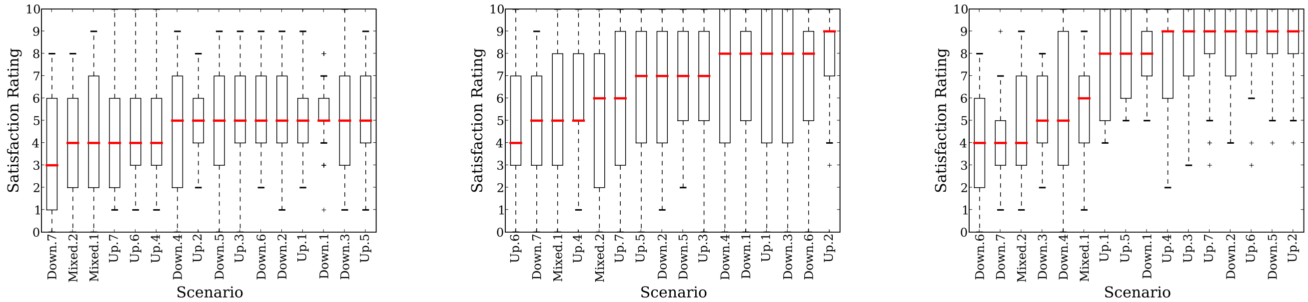
C. Proxy server

The proxy server is responsible for tracking network flows and associating them with the appropriate satisfaction sensor readings. Many network applications operate with multiple simultaneously open network connections, each of which contributes to the overall performance. EmNet assumes that user satisfaction is dependent on overall application performance. Specifically, EmNet targets web applications that can use multiple connections to multiple servers to complete a request and, because of the increased usage of AJAX, applications that use multiple persistent connections.

We refer to each group of connections that the proxy server tracks as a $FlowSet$. Each $FlowSet$ is a collection of network connections that is associated with a single satisfaction reading/performance value. Each $FlowSet$ consists of all the network connections being used by a single application. Each $FlowSet$ is treated as a single unit and the flows within it are operated on in aggregate by EmNet. The proxy server is responsible for creating descriptions of each $FlowSet$ that are used as packet classifiers by the network control mechanism.

D. Policy controller

The policy controller determines the policy that will be implemented by the network control mechanism. The policy controller is responsible for ensuring that the link bandwidth is



(a) Wikipedia (b) Image Labeler (c) Video

Fig. 3. Prompted user satisfaction under different cross-traffic scenarios sorted by increasing median satisfaction. Notice the considerable variation for each scenario, which often swamps aggregate differences between scenarios.

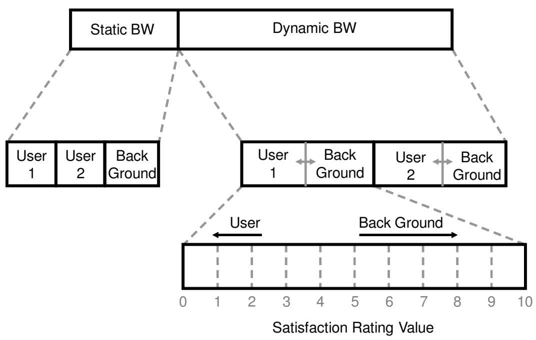


Fig. 5. Policy controller Algorithm

allocated such that user satisfaction is maximized while overall network performance is not significantly degraded. The inputs to the policy controller are the performance values that were previously calculated from the satisfaction of each user. EmNet assumes that not all traffic belongs to interactive applications. We refer to traffic for which there is no associated feedback as *background traffic*. The goal of the policy controller is to find the optimal tradeoff between user satisfaction and the amount of bandwidth allocated for the background traffic.

For EmNet to be useful in a multi-user context, there must be mechanisms in place to not only stop a single user from monopolizing the network but to also encourage honest reports of satisfaction from all users. To address this issue we rely on two components. First, we developed an algorithm for network provisioning that prevents bandwidth monopolization by misbehaving users. Second, we developed a method of associating a real world cost with lower satisfaction ratings.

1) Preventing link monopolization: The algorithm we use for driving the control policy ensures that no single user or group of users has the power to completely shut off the link for the other users. We do this by partitioning the link such that every user is guaranteed a portion of the link bandwidth. As we noted previously, EmNet groups network flows into one of two types: traffic that is associated with a satisfaction rating (a $FlowSet$) and everything else. The output of our algorithm is a partitioning of the network link such that each $FlowSet$ is allocated a portion of the link based on user feedback, while the background traffic is given the remainder. This partitioning is represented as a percentage of the total link bandwidth.

Figure 5 shows how EmNet partitions the link bandwidth between $FlowSet$s and the background traffic. At the highest level the bandwidth is split into two segments: one that is statically allocated and another that is dynamically allocated. The statically allocated segment is fairly partitioned between each $FlowSet$ and the background traffic. This bandwidth is reserved and is the minimum amount of bandwidth that is always available. The remainder of the available bandwidth is allocated dynamically to the $FlowSet$s based on the users' satisfaction. To prevent competition between users, the dynamically allocated segment is partitioned fairly between $FlowSet$s. However, each user is able to dynamically choose the amount of their dynamic partition that is actually reserved for their particular $FlowSet$. The leftover bandwidth in each $FlowSet$'s segment is allocated to the background traffic. The actual amount of reserved dynamic bandwidth is determined by the satisfaction rating collected from the user.

2) Cost function: While our control algorithm prevents any single network user from monopolizing the link bandwidth for themselves, it does nothing to prevent users from being dishonest in reporting their satisfaction and thus maximizing their bandwidth. Ideally, there would be a mechanism which can accurately detect a user's true satisfaction. In practice, we rely on an outside force that provides an incentive for the user to be honest and exerts a downward pressure on the user's satisfaction rating and performance value. This force takes the form of a performance cost function. By associating a given cost to the performance that a user requests, our system can encourage users to be honest in their satisfaction rating.

The choice of cost function is dependent on the local policy of the network using EmNet. In the context of the home network, this might take the form of a head of household setting per-user limits on network usage with access incurring a cost that grows in proportion to the weight. Perhaps no cost function is required if social pressure is an adequate mediator. A public wireless network may impose a financial cost on usage, once again scaled according to the weight selection.

E. Network control mechanism

The network control mechanism uses the results of both the proxy server and the policy controller to configure a network scheduler. EmNet uses weighted fair queuing as the mechanism for controlling network performance. All incoming and outgoing traffic is placed into the specific queues and then processed according to the queues' configured weights. By modifying the queue weights and altering the queues that a network connection uses, the network controller is able to optimize specific connections or groups of connections.

The queue weights are set by transforming the output percentages from the policy controller (the percentages of the network bandwidth available to each $FlowSet$ as well as the background traffic) into a set of weights that are assigned to the appropriate queue. Because the partitioning of the link is implemented using weighted fair queuing, any time a $FlowSet$ is not sending or receiving network traffic its queue is simply ignored. This has the effect of dropping that weight from the overall fairness calculation resulting in a larger bandwidth share for the remaining $FlowSet$s as well as the background traffic. Implementing bandwidth partitions on top of WFQ allows our provisioning algorithm to aggressively impose limits on network bandwidth without worrying whether those limits will cause underutilization of the available bandwidth.

F. Prototype implementation

In order to study the feasibility and effectiveness of EmNet, we built a prototype system that could be evaluated on our testbed (described in Section III-B). Much of our implementation used components developed and used during the initial user study. As a result, our implementation is not a monolithic system running inside an emulated broadband router but a distributed system with components at various locations in the network. There are no technical obstacles to building a system that runs on a commodity broadband router.

1) User interface: The EmNet implementation does not directly measure user satisfaction, but instead relies on the user to choose a preferred performance value, $Perf_{FlowSet}$, based on their satisfaction. We collect this value from the user by using techniques discussed in Section III-A. The web proxy adds a slider control that determines the user's allocation. The slider's scale is linear, and the initial slider value was set to the middle position. The interface also includes a display of the current cost value computed by the policy controller, displayed slightly above the slider. The slider setting is sampled every second and the cost is updated every 3 seconds.

2) Cost function: We implement a simple cost function that accumulates a price as the user makes use of the bottleneck link. The user effectively "pays" for each byte being transferred. We created a simple server application that monitors all traffic associated with each $FlowSet$, counting the number of bytes transferred. This value is multiplied by the current $Perf_{FlowSet}$ value—the higher the user's performance value, the dearer each byte is. Because the cost is proportional to the amount of data transferred, it tends to bias applications that consume more bandwidth and thus negatively effect our evaluation of EmNet. To limit the effects of this, EmNet multiplies the derived cost by a scaling factor. We tuned this parameter for each application such that the total cost of using the application during the study was within some range salient to the participants (around $10).

3) Proxy server: The EmNet implementation uses the proxy server implementation discussed in Section III-A. We extended the proxy to calculate the appropriate values for the cost function and send those values to the user. The policy controller is also implemented inside the proxy server.

4) Network controller: The network controller was implemented on top of a FreeBSD Server running DummyNET configured with an Ethernet bridge as in the first study. We used the FreeBSD Weighted Fair Queue implementation and relied on the kernel firewall to handle packet classification. Our configuration used a pair of rate limited upload and download *pipes* that were fed by multiple WFQs. Each FreeBSD WFQ corresponded to one WFQ from the EmNet design.

V. EVALUATION

To evaluate our EmNet prototype, we conducted a second user study with a different set of participants. The participants repeated the same tasks as in the study of Section III, but they were able to use EmNet to change their network performance. Our goal was to determine if users are able to use EmNet to increase their satisfaction with the network performance.

A. Study design

The study we conducted to evaluate EmNet was a modified version of the previous study, described in detail in Section III. We now summarize the differences.

Testbed: The testbed of the previous study, described in Section III-B, was reused but with EmNet running on it. This implies a significant change to the data path, in that the bridge server now runs FreeBSD and DummyNet, instead of Linux and IPRoute2.

Users, applications, and scenarios: The study was done using a new set of 18 participants drawn from the same university population as the initial study, using the same recruitment methods. Each participant was given the same set of survey questions used in the first study. 16 participants reported experience using a home network. The per-application familiarity and usage questions yielded results similar to those of first study (see Section III-C) so we omit them here.

The application tasks were identical to the previous study.

A subset of the cross-traffic scenarios from the earlier study was chosen. This subset contains those scenarios that had significant effects on median user satisfaction in the first study. A scenario with no cross-traffic was also included.

EmNet configuration: The EmNet implementation requires the setting of global parameters, as described in Section IV-D. The size of the statically allocated segment is 10% of the link capacity. Additionally, the initial performance value is set to the middle of the scale shown to the user.

Study design: The overall study design was identical to that described in Section III-D, with several differences summarized here.

First, in addition to the prompt for satisfaction after each network scenario, users were also given the slider interface. Second, each user was exposed to a given network scenario

Application	Strategy 1	Mean	Strategy 2	Mean	p-value
Wikipedia	Observation	4.15	Static WFQ	6.31	< 0.001
	Observation	4.15	EmNet	6.21	< 0.001
	Static WFQ	6.31	EmNet	6.21	0.544
Image Labeler	Observation	5.75	Static WFQ	7.37	< 0.001
	Observation	5.75	EmNet	7.73	< 0.001
	Static WFQ	7.37	EmNet	7.73	0.051
Video	Observation	5.96	Static WFQ	3.30	< 0.001
	Observation	5.96	EmNet	5.11	0.015
	Static WFQ	3.30	EmNet	5.11	< 0.001
Overall	Observation	5.29	Static WFQ	5.66	0.142
	Observation	5.29	EmNet	6.35	< 0.001
	Static WFQ	5.66	EmNet	6.35	< 0.001

Fig. 9. Mean satisfaction for each network provisioning strategy and two-tailed p-value for each pair. We find that EmNet significantly increases user satisfaction over both the uncontrolled (Observation) and statically provisioned cases.

for 60 seconds (as opposed to 30 seconds in the earlier study). The purpose was to give them time to find a satisfactory setting of the EmNet slider control.

The users were told that moving the slider up increased performance while moving it down decreased performance. To provide an incentive for the users to move the slider down, the cost penalty was described to them and the cost display was always visible as part of the interface. The users were given the goal of minimizing cost while maintaining a satisfactory level of network performance.

Each user experienced every network scenario twice; once where the slider control actually controlled their network performance, and once where the slider control was visible but had no effect. In the latter case, the network was configured with equal weights for both the user-generated and background traffic, resulting in a fair-share allocation. The ordering of tasks, scenarios, and network control policies was randomized.

B. Results

We consider the bandwidth used, the prompted user satisfaction, and the users' chosen performance values. Figure 6 shows the average download bandwidth used by the applications with EmNet and static WFQ, grouped by scenario. The whiskers represent the overall standard deviation, not the confidence interval for the mean. Figure 7 shows the average user satisfaction for EmNet, static WFQ, and for the earlier observational study. Finally, Figure 8 shows the distribution of the performance values ($Perf_{FlowSet}$) chosen by the users.

We also compare the significance of the changes in user satisfaction between each of the three network provisioning strategies. To compare both the static WFQ and EmNet to the data from the observational study, we apply a t-test assuming unequal variance to the satisfaction ratings for the set of scenarios common to both studies.[3] To compare the static WFQ and EmNet satisfaction values, we use a paired t-test. We report the p-values of the two-tailed test for each pair of strategies in Figure 9.

1) Wikipedia: In terms of bandwidth consumed, the static WFQ and EmNet results are similar. As seen in Figure 6(a) Wikipedia consumed a moderate amount of bandwidth for both, but with high variability across the users. The variation in application bandwidth that we see in scenarios where there is significant contention for download bandwidth is generally slightly higher with EmNet than with static WFQ.

Figure 7(a) shows the prompted user satisfaction numbers for Wikipedia. Note that unlike the bandwidth numbers, here we also compare with the observational study results. The overall shape of the results is similar across the cross-traffic scenarios. The average satisfaction levels for static WFQ and EmNet are comparable, and both are significantly higher than those in the observational study ($p < 0.001$). On average, static WFQ and EmNet users are $\sim 51\%$ more satisfied than the observation study participants. Users reported a slightly higher mean satisfaction using WFQ as compared to static EmNet, though this difference is not significant.

When using EmNet, many users did not change their setting from the default. Figure 8(a) shows the distribution of performance values ($Perf_{FlowSet}$) that users chose for each cross-traffic scenario. For each scenario the median value was near or slightly above 5, which would result in performance near that of static WFQ.

Taken together, the results show that the addition of network control, such as static WFQ or EmNet, dramatically increases satisfaction with Wikipedia. With such control, cross-traffic has little effect on the satisfaction. The average bandwidth consumed by Wikipedia with network control is largely independent of the amount of cross-traffic, suggesting that the necessary bandwidth was almost always available to the user. This in part explains the similarity between the static WFQ and EmNet satisfaction results.

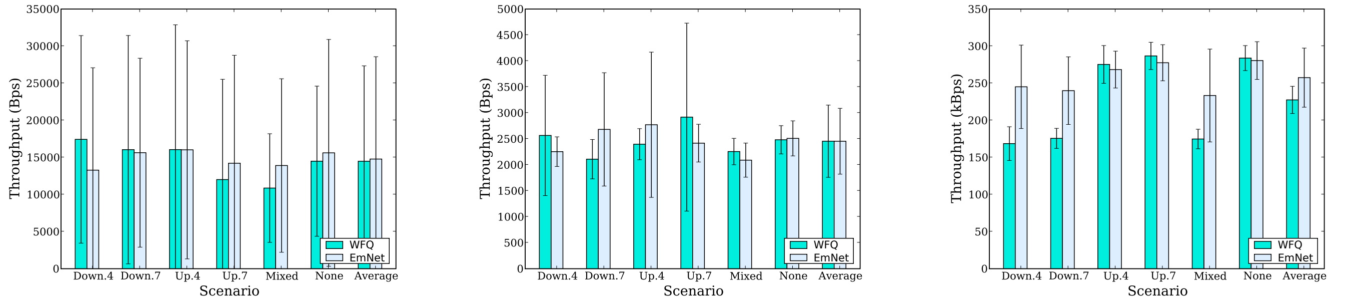
2) Image labeler: Figure 6(b) shows that the required bandwidth was fairly small. While the variation in bandwidth is not insignificant, it is generally much less than Wikipedia. Because the game is interactive this suggests that performance was strongly dependent on latency rather than available bandwidth. As with Wikipedia, the average download bandwidth for static WFQ and EmNet changes little across cross-traffic scenarios.

Average satisfaction increases dramatically going from the observational study to both static WFQ and EmNet ($p < 0.001$). There is also a slight increase in user satisfaction using EmNet as compared to static WFQ ($p = 0.051$). The satisfaction measurements are given in Figure 7(b). WFQ and EmNet increase average satisfaction $\sim 31\%$ compared to the observation study. However, due to the very small demand for bandwidth and the significant amount of interaction required, static WFQ and EmNet are likely improving satisfaction by providing latency bounds.

Compared to Wikipedia or Video, users of Image Labeler are far more likely to *decrease* their $Perf_{FlowSet}$ values. The distributions of $Perf_{FlowSet}$ values are shown in Figure 8(b). Adequate performance can be had at a lower setting due to low bandwidth demands. However, the median *is* 5—we had expected users would choose even lower $Perf_{FlowSet}$ values.

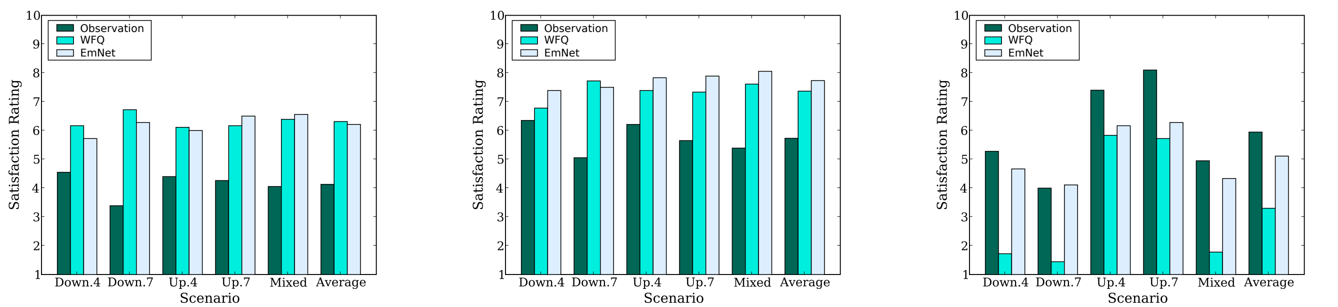
Intuitively, highly interactive low bandwidth applications can deliver very high satisfaction when provided with guaranteed latency bounds. Both static WFQ and EmNet provide

[3]Comparisons between the observational study and the evaluation study should take into consideration their differences, which are summarized in Section V-A.



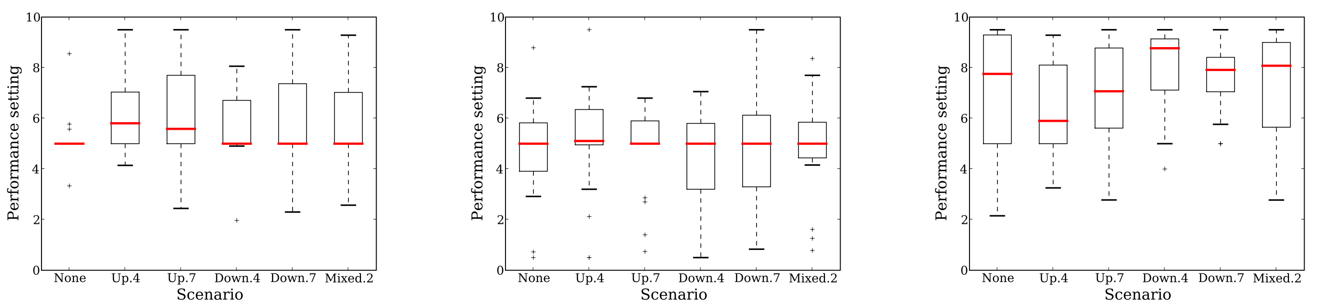
(a) Wikipedia (b) Image Labeler (c) Video

Fig. 6. Average downstream throughput under each scenario and system. Error bars indicate the standard deviation of the throughput.



(a) Wikipedia (b) Image Labeler (c) Video

Fig. 7. Average satisfaction level under each scenario and system.



(a) Wikipedia (b) Image Labeler (c) Video

Fig. 8. Variation in user-selected performance values ($Perf_{FlowSet}$) under EmNet for different cross-traffic scenarios.

this, while the configuration of the observation study does not. Additionally, the results also show that EmNet provides a consistent if small increase in average satisfaction over the static WFQ configuration.

3) Video: Unlike Wikipedia and Image Labeler, the Video task uses a large amount of download bandwidth. The bandwidth for Video is given in Figure 6(c). Unlike the earlier tasks, the video stream bandwidth is clearly different across the various cross-traffic scenarios, while at the same time showing much less relative variance. Because the video stream was sent over UDP it was less susceptible to upload cross-traffic. The bandwidth for static WFQ and EmNet is similar when there is no contention on the download link, but with contention EmNet provides considerably more bandwidth to Video than static WFQ. With download contention, EmNet provides $\sim$ 25% more bandwidth than static WFQ. That is, the users are able to demand more bandwidth for Video with EmNet.

Video shows the most dramatic differences in satisfaction and the largest differences between static WFQ and EmNet. Figure 7 shows the satisfaction results, showing that both EmNet and static WFQ result in lower mean user satisfaction as compared to the observation study. The difference in mean user satisfaction between static WFQ and the observational study is 2.66 ($p < 0.001$) and the difference between EmNet and the observational study is 0.85 ($p = 0.015$). Examining the packet traces collected in our study yields an explanation: In the observation study the UDP-based video stream was free to compete with the TCP cross-traffic, resulting in the congestion control algorithm being activated for the TCP flows. This substantially reduced the throughput of the cross-traffic flows, allowing the video stream to monopolize the link bandwidth. With static WFQ the video stream was limited to 50% of the link bandwidth, which resulted in a very large decrease in the video quality that was immediately noticeable to users.

Users reported significantly higher satisfaction with EmNet than with static WFQ ($p < 0.001$). With EmNet, users are able to increase the bandwidth allocated to video, resulting in large increases in satisfaction. This increase in satisfaction is purchased with only a tiny increase in application bandwidth (compare Figure 7(c) to the application bandwidth results in Figure 6(c)). Note that it is probably not possible for EmNet to achieve the same satisfaction results as in the observational study, nor should it be—performance there was due to choking the competing TCP flows.

Figure 8 shows the distributions of the users' $Perf_{FlowSet}$ values. As we might expect, most users moved the slider up to improve performance—the medians are all greater than 5. Further, the median $Perf_{FlowSet}$ values for the cross-traffic scenarios with download contention were the highest. It is interesting to note that even when there is no cross-traffic, the median value remains high, implying that the users are reluctant to decrease their performance.

4) General Results: Comparing the overall performance of all three network control configurations, we find that EmNet increases average user satisfaction by 20% compared to a configuration without network control ($p < 0.001$). Further, EmNet increases average satisfaction by 12% compared to static WFQ ($p < 0.001$). Finally, EmNet achieves these increases in satisfaction with only a 6% increase in application bandwidth compared to static WFQ.

Surprisingly, we find that, overall, users are reluctant to decrease their $Perf_{FlowSet}$ value once it has been raised. While the per-byte cost displayed in the user interface of EmNet is intended to encourage users to decrease their setting as circumstances permit, it is not clear that this cost was a powerful enough incentive. Of course, in the study the cost reflects no real world monetary value. We speculate that a real cost that is expected in a deployment would do better. Alternatively, pressure from other users in the home network could act as an additional cost function. We hope to explore both of these in our future work.

VI. Conclusions

We introduced a new method of optimizing home network broadband connections by using individual user satisfaction as an input. We conducted an observational user study that demonstrated that user satisfaction shows a high degree of variance, meaning that each user's perception of network performance is very different. Optimizing for a canonical user is not sensible. We designed and implemented EmNet, a system that optimizes a home network broadband connection based on measurements of individual user satisfaction. We evaluated EmNet in a second user study that demonstrated that individualized optimizations can considerably improve user satisfaction with low resource cost. On average EmNet is capable of increasing user satisfaction by 20% over an uncontrolled link and by 12% over a simple static configuration using weighted fair queuing. It does so by increasing average application bandwidth by only about 6%.

References

[1] E. Athanasopoulos, et al., "Homemaestro: Order from chaos in home networks," in *Proceedings of ACM SIGCOMM (demo)*, August 2008, also see Microsoft Research Technical Report MSR-TR-2008-84.

[2] M. Dischinger, et al., "Characterizing residential broadband networks," in *Proceedings of the 7th ACM SIGCOMM / USENIX Internet Measurement Conference (IMC)*, October 2007.

[3] K. Lakshminarayanan and V. N. Padmanabhan, "Some findings on the network performance of broadband hosts," in *Proceedings of the 3rd ACM SIGCOMM / USENIX Internet Measurement Conference (IMC)*, October 2003.

[4] A. Gupta, B. Lin, and P. A. Dinda, "Measuring and understanding user comfort with resource borrowing," in *Proceedings of the 13th IEEE International Symposium on High Performance Distributed Computing (HPDC 2004)*, June 2004.

[5] A. Mallik, et al., "User-driven frequency scaling," *Computer Architecture Letters*, vol. 5, no. 2, 2006.

[6] P. Dinda, et al., "The user in experimental computer systems research," in *Proceedings of the Workshop on Experimental Computer Science (ExpCS 2007)*, June 2007.

[7] A. Shye, et al., "Learning and leveraging the relationship between architectural-level measurements and individual user satisfaction," in *Proceedings of the 35th International Symposium on Computer Architecture (ISCA)*, June 2008.

[8] A. Mallik, et al., "Picsel: Measuring user-perceived performance to control dynamic frequency scaling," in *Proceedings of the 13th International Conference in Architectural Support for Programming Languages and Operating Systems (ASPLOS)*, March 2008.

[9] B. Lin and P. Dinda, "Vsched: Mixing batch and interactive virtual machines using periodic real-time scheduling," in *Proceedings of ACM/IEEE SC (Supercomputing)*, November 2005.

[10] J. Lange, P. Dinda, and S. Rossoff, "Experiences with client-based speculative remote display," in *Proceedings of the USENIX Annual Technical Conference (USENIX)*, 2008.

[11] A. Shye, et al., "Power to the people: Leveraging human physiological traits to control microprocessor frequency," in *Proceedings of the 41st Annual IEEE/ACM International Symposium on Microarchitecture (MICRO)*, November 2008.

[12] C. Yun, et al., "O' game, can you feel my frustration?: Improving user's gaming experience via stresscam," in *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems*. New York, NY, USA: ACM, 2009, pp. 2195–2204.

[13] S. Tarzia, et al., "Sonar-based measurement of user presence and attention," in *Proceedings of the 11th International Conference on Ubiquitous Computing (UbiComp)*, September 2009, also see NWU-EECS-09-09.

[14] J. Bennett and H. Zhang, "Worst-case fair weighted fair queueing," in *Proceedings of IEEE INFOCOM 1996*, March 1996.

[15] D. Stiliadis and A. Varma, "Latency-rate servers: A general model for analysis of traffic scheduling algorithms," *IEEE/ACM Transactions on Networking (ToN)*, vol. 6, no. 5, pp. 611–624, 1998.

[16] Consumer Electronics Association, "CEA finds american adults spend $1,200 annually on consumer electronics products," April 2007.

[17] J. B. Horrigan, *Home Broadband Adoption 2008*. Pew Internet and American Life Project, 2008.

[18] C.-C. Tu, et al., "Oneclick: A framework for capturing users network experiences," in *Proceedings of ACM SIGCOMM 2008 (poster)*, August 2008.

[19] K.-T. Chen, C. C. Tu, and W.-C. Xiao, "Oneclick: A framework for measuring network quality of experience," in *Proceedings of IEEE INFOCOM 2009*, 2009.

[20] J. S. Miller, J. R. Lange, and P. A. Dinda, "EmNet: Satisfying the individual user through empathic home networks," Department of Electrical Engineering and Computer Science, Northwestern University, Tech. Rep. NWU-EECS-09-05, April 2009.

[21] H. Blom, "Modified iperf implementation," (web page), http://staff.science.uva.nl/ jblom/gigaport/tools/test%5Ftools.html.

[22] WikiMedia Foundation, Inc., "Wikipedia, the free encyclopedia," (web page), http://en.wikipedia.org/.

[23] Google, Inc., "Google image labeler," (web page), http://images.google.com/imagelabeler/.

[24] VideoLan team, "Vlc media player," (web page), http://www.videolan.org/vlc/.