Real-time probing of available bandwidth in home networks

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Real-Time Probing of Available Bandwidth in Home Networks

Archi Delphinanto, Eindhoven University of Technology and TNO

Ton Koonen, Eindhoven University of Technology

Frank den Hartog, TNO

ABSTRACT

Prioritization of flows in a home network based on traffic classification is still no guarantee that enough bandwidth will be available between a content server and a client. Besides, such QoS technologies need to be supported by every device in the end-to-end path to be effective, which is relatively expensive for the owners of home networks. In any small-scale IP network, best effort or QoS-enabled, at home or anywhere else, it is therefore preferable to diagnose the network in real time before admitting a new flow. In this article we analyze existing probing techniques, and demonstrate a new method to probe the available bandwidth between a server and a client in a heterogeneous IP-based home network. The tool works with existing end-user devices, is non-intrusive, has a short measurement time, does not require preknowledge of the link layer network topology, and is accurate enough to make decisions about the admission of high-throughput high-quality streams such as for IPTV services.

QoS IN HOME NETWORKS

Home networks are becoming ever more heterogeneous. This means that a single network consists of many different physical- and link-layer technologies and topologies, interconnecting many different devices with each other and the Internet, enabling many different services (e.g., [1, references therein]). The Internet Protocol (IP) suite is the main enabler for the required interoperability. Correspondingly, a growing amount of consumer electronics devices contain an IP stack. Unfortunately, IP has only limited support for quality of service (QoS), necessary to support many different services concurrently in a single shared home network. It mainly concerns reducing or avoiding jitter that occurs when multiple traffic streams contend for bandwidth. Furthermore, noticeable packet loss will happen when User Datagram Protocol (UDP) based services such as high-quality telephony and HDTV need to be supported. This is especially an issue for broadband service providers (e.g., [2, 3]).

Many additional QoS solutions for IP net-works are available, but most of them operate on the principle of traffic classification, where each data packet is placed into a limited number of traffic classes, and each router on the network is configured to differentiate the traffic based on its class. These solutions have not gained large popularity in the home networking marketplace, because they need to be supported by every device in the end-to-end (e2e) path to be effective. This makes them relatively expensive for consumers with many non-depreciated devices: to enjoy QoS they have to buy new devices. Besides, current solutions are different for different layer 2 technologies. Intermediate translators would then be needed to guarantee e2e QoS in a heterogeneous path. Though implementations for this exist (e.g., [4]), they are deemed to be too expensive for mass-scale application today. Finally, prioritization of flows is still no guarantee that enough bandwidth (here synonymously used for application-level data throughput) will be available between a server and a client.

Before admitting a new flow to the home network, or rather any small-scale IP network (best effort or QoS-enabled), we therefore propose to diagnose the network in real time. The information can be used for admission control, have the content service pragmatically adapt its properties to the actual condition of the network, or report intermittent issues to the user and/or the service provider's remote service management server. A crucial part of a diagnosing tool is the real-time assessment of e2e available bandwidth between the relevant client and server in the network. Although many e2e speed test applications exist for the Internet [5], none of them fulfills all of the requirements for use in today's home networks. Among these requirements are the following. The tool must:

1 Be easy to implement. It should work with as many as possible existing devices in the home without firmware upgrades. Preferably the tool should only require a simple software module added on the server side of the e2e path of a flow. For most use cases we can therefore assume the diagnosing application to be a service running on the home gateway, serving various clients in the home network, which only need to have a regular IP stack.

2 Be non-intrusive. It should not disrupt other traffic in the home noticeably.

3 Have a short measurement time. It should have a low convergence time from an end-user perspective, and it should be fast enough to react to major changes in the home network traffic pattern. We assume this to be on the order of a few seconds; fluctuations within this time frame may be dealt with by application-layer buffering, for instance.

4 Not require pre-knowledge of the link-layer network topology. Home networks may contain link-layer technologies that are not standardized or widely known.

5 Be accurate enough to make informed decisions about the admission of delay- and jitter-critical applications. In the case of IPTV and IP telephony, that means an accuracy of ~1 Mb/s and ~50 kb/s, respectively.

In this article we propose a new tool, the Available Bandwidth Estimator (Allbest), which fulfills all the above requirements to the extent that it is a software upgrade of only the probe server, it injects less than ~ 1 Mb/s of probe traffic, produces results with an accuracy better than ~ 1 Mb/s within less than 10 s, and can be applied on any Ethernet-WiFi combined topology. This is the first diagnosing application that successfully applies probe round-trip time (RTT) measurement to wireless LANs and does not assume any home network topology a priori. The tool is based on a new probing method we developed, which is described in the following section. In the second half of the article we detail our testbed, followed by our test results. We finish with a conclusion and a few words on our current and future experiments.

PROBING HOME NETWORKS

BOTTLENECK CAPACITY AND AVAILABLE BANDWIDTH

We follow [6] for defining capacity of a hop as the bit rate, measured at the IP layer, at which the hop can transfer maximum transmission unit (MTU)-sized IP packets. Therefore, the capacity of an e2e path is the maximum IP-layer rate the path can transfer from source to sink. In our work, we assume the MTU size to be 1500 bytes of Ethernet v2 (RFC 1191). As a path may consist of several links, the minimum link capacity in the path determines the path capacity. This link is called the narrow link. In contrast, the tight link is the link in the path with the maximum capacity utilization. This is the link with the least available bandwidth due to crossing traffic (i.e., other traffic in the path considered for admission of a new stream). In many cases, the tight link is in the narrow link, and the link is then referred to as the bottleneck.

For measuring bottleneck bandwidths in Internet paths, two types of tools can be distinguished: packet-pair dispersion tools (also called probe gap model, or PGM) and self-loading techniques. The latter probe the network with trains of packets [7] at an increasing rate, and thus rely on flooding the network. They therefore do not fulfill requirement 2. PGM techniques were first explored in [8], and send only a few packets at the rate C of the bottleneck capacity or somewhat slower. This allows crossing traffic to get in between the probe packets and disperse them (i.e., increase the difference in arrival time).

An issue with PGM techniques is that C needs to be known a priori. From requirement 4 follows that the tool must be able to estimate the available bandwidth A without such pre-knowledge of the path. This means that C needs to be determined first, and the estimation of A becomes a two-step process. In [9] we proposed and validated a new method for determining C in heterogeneous home networks based on RTT measurements of probe packets, and fulfilling all requirements listed in the previous section. However, we also learned that probing a wireless LAN with rate C may yield the correct average dispersion rate at the receiver when measuring in one direction, but will not if it needs to be derived from RTTs. This is easiest understood by looking at the details of our capacity estimation method first.

CAPACITY ESTIMATION

Our capacity estimation method is based on the packet-pair dispersion technique, which is usually implemented by sending two packets back-to-back on the network, thus minimizing the chance that crossing traffic will disperse the packets. It is then the bottleneck that will delay the second packet with respect to the first. C can subsequently be calculated simply from the minimum dispersion D and the packet size L as $C = L/D$. One then should minimize the chance that crossing traffic increases or decreases the bottleneck dispersion of the packets further down the path. To do so, we perform a series of n packet-pair probes, assume the crossing traffic stochastic, and then calculate D from the minimum RTT of the first packet (RTT_1) of a probe pair and the minimum RTT of the second packet (RTT_2) of a probe pair. C is then given by

$$C = L / (\min_{i=1...n} [RTT_2(i)] - \min_{i=1...n} [RTT_1(i)]). \quad (1)$$

RTTs can be measured without adaptation of the client side by using MTU-sized Internet Control Message Protocol (ICMP) Ping probe packets, or by sending MTU-sized UDP packets to a non-activated port. The client then automatically generates reply packets: ICMP Echo packets or ICMP Error packets (i.e., code 3 or "Destination port unreachable"), respectively. ICMP Error packets are much smaller than ICMP Echo packets and therefore experience hardly any delay on the way back to the probing sender/receiver, assuming that the return one-way capacity between the client and the server, $C_{reverse}$, is not much smaller than the sought-after one-way capacity $C_{forward}$ between server and client. The final result is then a good measure for $C_{forward}$. For symmetric media we may also use ICMP Ping probing packets, and assume that the delay and dispersion is the same for both directions of travel. Equation 1 then yields $C/2$ rather than C.

Equation 1 allows us to avoid unwanted contention of probe packets in the wireless medium. Existing packet-pair dispersion techniques will not

> Our capacity estimation method is based on the packet-pair dispersion technique, which is usually implemented by sending two packets back-to-back on the network, thus minimizing the chance that crossing traffic will disperse the packets. It is then the bottleneck that will delay the second packet with respect to the first.

Figure 1. *Conventional back-to-back packet-pair probing in heterogeneous home networks. PK1 and PK2 are the probe packets sent back-to-back on the path. REP1 and REP2 are the respective reply packets. a) PK2 has to wait until REP1 is off the wireless medium; b) REP2 has to wait until REP1 is off the wireless medium.*

work in wireless media round-trip, because the reply packet of the first probe packet contends with the second probe packet on the air interface (Fig. 1). Irrespective of which packet wins, the reply packet of the second probe packet will eventually arrive at the probing sender/receiver too late. PK2 is acting as crossing traffic to the reply packets; the method is basically self-disturbing. As a result, C will be underestimated.

To avoid this contention, we need to prevent the first reply packet from being put on the network. We achieved this (Fig. 2) by sending a single packet with size $2 \times$ MTU instead of two packets back-to-back. On the network, this packet will automatically be fragmented (and behave like two individual packets back-to-back), and only after defragmentation will a single reply be sent back by the client. This will provide us the correct RTT_2, that is, the RTT_2 only delayed by bottleneck dispersion, not by additional contention. Because we are not directly measuring D, but separate RTTs (Eq. 1), we can find the correct RTT_1 by sending a different series of single probe packets, well separated from each other and with size MTU.

Contention of probe and reply packets is the main reason why conventional packet-pair probing is not suitable for determining A. We can neither probe back-to-back nor with rate C without creating extra delay caused by contention. Unfortunately, we cannot solve the latter by using fragmentation, because that can only mimic back-to-back probing.

AVAILABLE BANDWIDTH ESTIMATION

A deep analysis of the various delays that constitute the RTTs observed during capacity estimation allowed us to make a good estimation of A also. In Fig. 3 a typical histogram is shown of RTT_1 that we measured in an IEEE 802.11b net-

work with 1.5 Mb/s crossing traffic. Besides a clear minimum value, the RTT undergoes two random effects: the random back-off mechanism of IEEE 802.11 (mostly at short additional delays) and the delay caused by queuing due to crossing traffic.

We assume that, for UDP probing, most of the random delay is experienced in the forward direction. This is justified by the fact that the reply packet is very small, and we assume that the queuing mechanism of the system is fair. The reply packet is therefore hindered relativ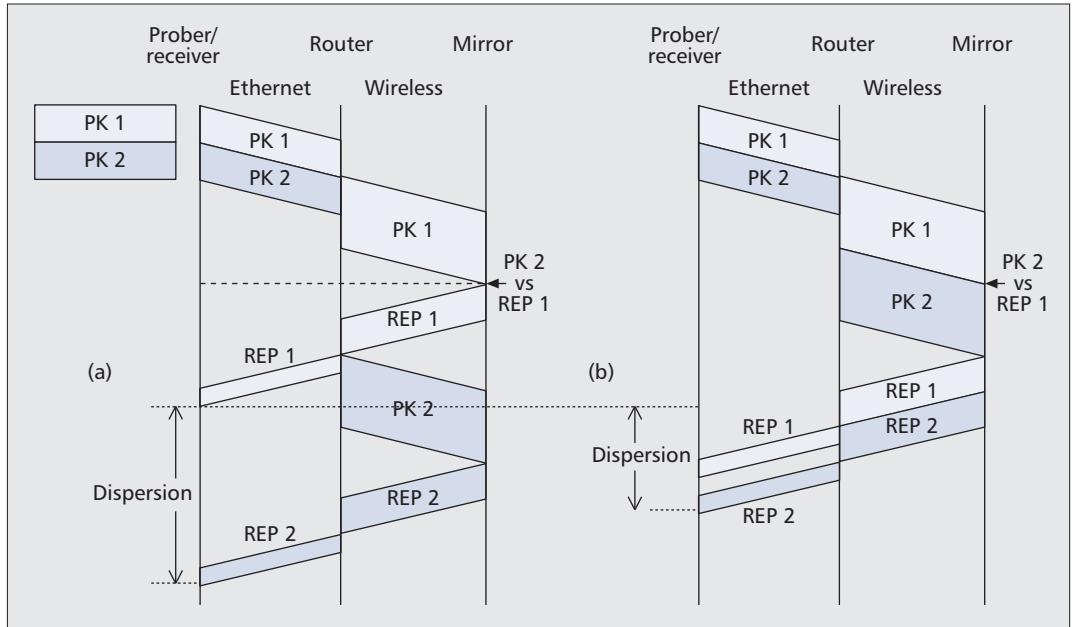ely little by the crossing traffic. We further assume that any systemic delay in the network (for instance processing delay) is either negligible or canceled when subtracting RTT_1 from RTT_2 [10], and that the delay caused by random effects is mainly happening in the bottleneck. A is then given by

$$A = L / \left(\frac{L}{C} + \bar{d}_r \right) \tag{2}$$

with L/C the delay in the bottleneck without crossing traffic following from Eq. 1, and the average delay caused by random effects in the bottleneck. The latter can be derived from RTT_1 as

$$\bar{d}_r = \underset{i=1...n}{avg} [RTT_1(i)] - \underset{i=1...n}{min} [RTT_1(i)]). \tag{3}$$

ALLBEST TESTBED

The setup of our testbed is schematically drawn in Fig. 4. The Allbest server runs on the "prober/receiver" computer, and probes the "mirror" via any heterogeneous topology of interest. The results presented in this article were obtained by configuring a WLAN IEEE 802.11b or 802.11g with a Linksys WRT54GL v.
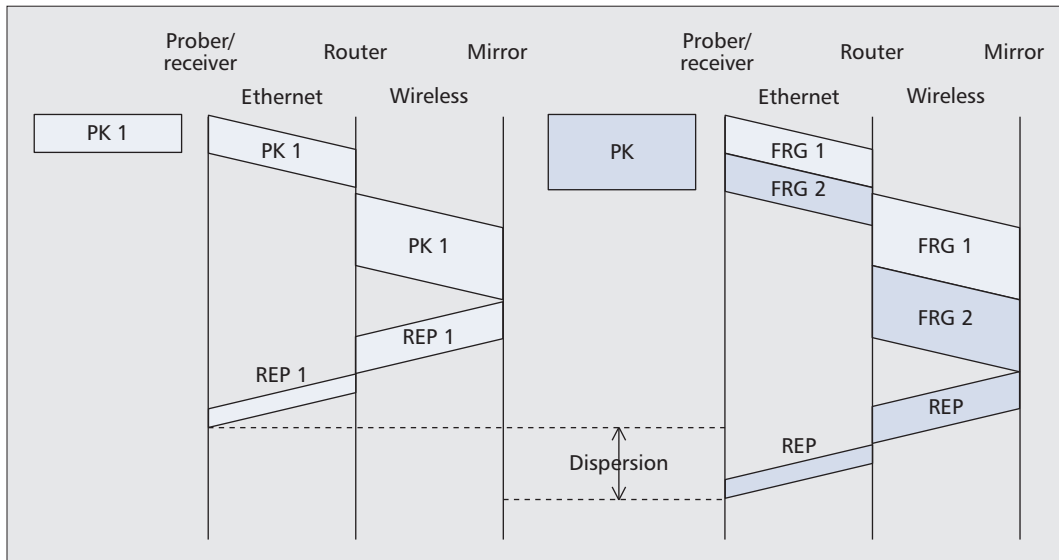
Figure 2. *Allbest's method of active probing in heterogeneous home networks. Probe packets PK1 and PK are sent far apart from each other. PK has size 2 × MTU and is automatically fragmented on the network. The fragments FRG are dispersed. A reply REP is sent only after defragmentation; thus, no contention has occurred.*

1.1 access point as a bottleneck link. We switched off the automatic rate adaptation and clear to send (CTS) protection mode, and run both networks on their maximum physical rates of 11 Mb/s and 54 Mb/s, respectively. The measurements were carried out in a Faraday cage to avoid uncontrolled interference.

We benchmarked Allbest against the well-known testing tool Iperf and Wbest [11]. Wbest is the only other real-time probing tool we know that is applicable to wireless networks. It requires the wireless hop to be in the last link, because it needs to be sure that the probing packets arrive at the bottleneck with rate *C*. For the estimation of *C* it uses standard PGM and packet-pair dispersion. Both Wbest and Iperf need to be installed on both the prober/receiver (which for Wbest and Iperf just acts as a prober) and the mirror (which for Wbest and Iperf acts as a receiver). The prober/receiver and mirror are laptop computers with a 2.0 GHz processor. Allbest and Iperf run on Windows XP Service Pack 3, and Wbest runs on Linux Ubuntu 10.04. To maximize the performance of the software, other processes running in the computers' background memory were switched off whenever possible.

Allbest basically consists of a home-built configurable UDP packet generator and a home-built configurable ICMP Ping packet generator, combined with Wireshark to measure high-precision RTTs. Any RTT > 2 × min[RTT(i)] is discarded, and we have verified that most of those long RTTs are caused by uncontrollable processing delay in the laptops due to other tasks of the operating system. A measurement takes 90 probe pairs and is repeated 6 times.

With Iperf we measured at which UDP injection rate which packet loss occurs with 1472-byte payload per packet. The result is fitted linearly, and the point where the fitted line crosses the transmission rate axes is interpreted as being the available bandwidth. Each Iperf measurement is set for 10 s with 1 s interval. The UDP packet
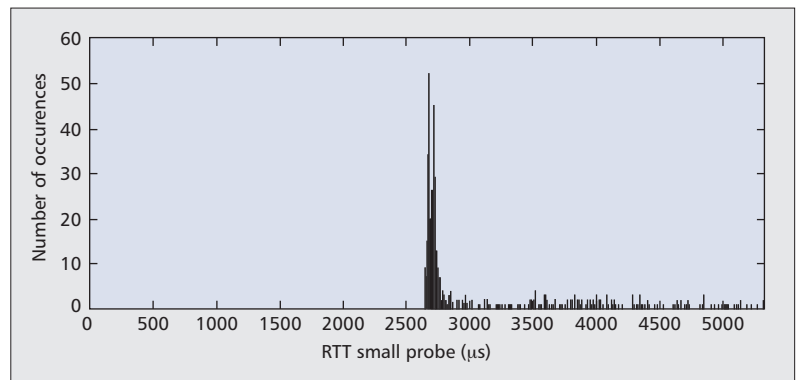

Figure 3. *Histogram of 1000 RTTs of MTU-sized probe packets in an IEEE 802.11b network with 1.5 Mb/s crossing traffic. Bin size = 20 μs.*

loss is averaged over 8–10 similar measurements, leading to 1 percent standard deviation. Also, Wbest was configured to use 1472-byte UDP payload. Each measurement of 90 packet pairs was repeated 30 times. Like Allbest, Wbest filters and discards unreliable results.

Random UDP crossing traffic is generated with the Distributed Internet Traffic Generator (D-ITG), available from the Universita degli Studi di Napoli ''Federico II.'' The UDP packets have uniformly distributed packet sizes (40–1472 bytes) and are sent at Poisson-distributed exponential time intervals. We distinguished crossing traffic and contending traffic, and follow [11] for their definitions.

AVAILABLE BANDWIDTH IN ETHERNET-WIFI NETWORKS

We have obtained the available bandwidth for three different topologies with Iperf, Wbest as well as Allbest (in its UDP probing variety):
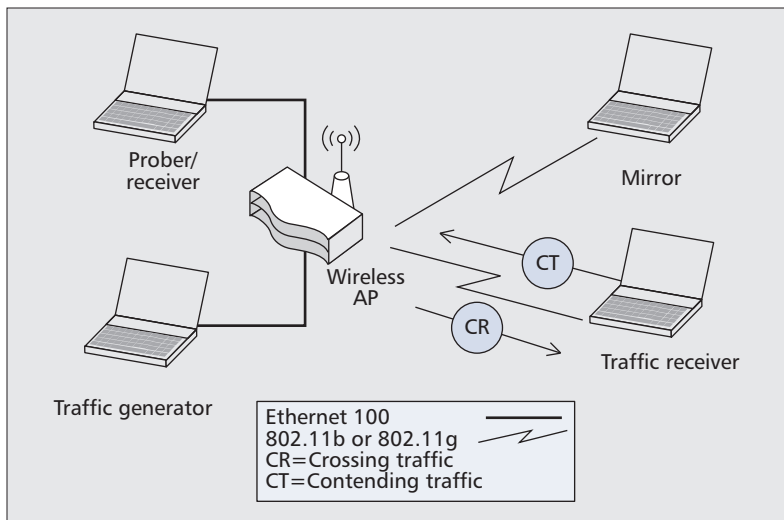1 Prober/receiver→100BASE-TX→IEEE 802.11b→mirror

Figure 4. *Schematic view of our heterogeneous wired/wireless LAN probing testbed. Allbest runs on the prober/receiver. The traffic generator generates crossing or contending traffic.*

2 Prober/receiver→100BASE-TX→IEEE 802.11g→mirror
3 Prober/receiver→IEEE 802.11g→100BASE-TX→mirror

For every topology we generated three different amounts of crossing traffic (X), at about 0, 25, and 50 percent of the capacity, and 25 percent of contending traffic.

For topology 1, the results are summarized in Fig. 5a. For this topology, with IEEE 802.11b, all tools yield similar results at first sight and within the error margins. Allbest estimates the capacity C of 802.11b on 7.6 ± 0.2 Mb/s, which is equal to the theoretical value [9]. For all tools, the available bandwidth A is lower than C for $X = 0$. This is caused by the random backoff mechanism of WLAN. If the random backoff algorithm were to be active for all packets, we would expect $A = 6.4$ Mb/s [9]. Since all tools estimate available bandwidths somewhat higher than that, we suspect there are still many packets that do not undergo random backoff.

On the whole, Allbest seems to find larger available bandwidths than Iperf, and Wbest finds even larger ones. Even though Iperf is a well-known benchmarking tool, it is probably underestimating A in our experiments. Because the crossing traffic is stochastic, some packet loss will already be recorded at relatively low Iperf injection rates. For both Allbest and Wbest, the values for A at $X > 0$ are also closer than Iperf to the expected value of A (obtained by simply subtracting the crossing traffic rate X from the available bandwidths A at $X = 0$ Mb/s). The error margins of the Allbest results are remarkably lower than the ones for Wbest, although we tried to have the results based on the same number of probes. We do not have an explanation for this yet.

For Allbest, the value for A at $X = 3$ Mb/s was calculated by discarding any RTT > 3 × min[RTT(i)], rather than using the default cutoff time of 2 × min[RTT(i)], as stated in the previous section. We found that with the latter, too many packets had been discarded that were clearly delayed by crossing traffic, and A was

grossly overestimated (5.1 ± 0.2 Mb/s). Many PGM techniques use a default cutoff time of 2 × min [RTT(i)]. This follows from their assumption of fair queueing congestion management in the router. This means that bottleneck delays can never be larger than $2L/C$, even if the utilization by crossing traffic is larger than 50 percent (which then just results in larger packet loss). Crossing traffic of 3 Mb/s with randomly distributed time intervals will utilize the bottleneck more than 50 percent for at least part of the time. Fortunately, the actual congestion management mechanism of the router (most probably store and forward) still allowed us to capture relevant packets with larger RTTs and compute a realistic value for A.

For topology 2, the results are summarized in Fig. 5b. For this topology, with IEEE 802.11g, Allbest shows clear supremacy. Allbest estimates the capacity C of 802.11g on 38 ± 2 Mb/s, which is equal to the theoretical value [9]. The value of A at $X = 0$ Mb/s is then expected to be 26 Mb/s if the random backoff algorithm is active for all packets [9]. Iperf and Allbest estimate somewhat higher again, but Wbest significantly underestimates A, also for larger X. The inventors of Wbest warn of underestimation when the probe packets arrive at the bottleneck at a rate larger than C [11]. Surprisingly, Wbest's capacity estimation for topology 2 is quite good, 38 Mb/s. The fact that Wbest arrived at plausible answers for A with topology 1 can be explained by it grossly overestimating the C of topology 1 (8.8 Mb/s). The results for Allbest are very close to the ones for Iperf, and have the lowest error margins of all. But as in Fig. 5a, it is not clear whether Iperf yields the correct values. More than Iperf, Allbest yields values for A at $X > 0$ close to what one obtains by subtracting X from $A(X = 0)$.

For topology 3, the results are much the same as for topology 2 (Fig. 5c). This shows that Allbest can function in either order of physical- and link-layer technologies. Unfortunately, we could not get any UDP results from Iperf, because it cannot inject faster than about 10 Mb/s when directly connected to an 802.11g network.

CONCLUSIONS AND FUTURE WORK

We achieved a breakthrough in available bandwidth probing of heterogeneous home networks by understanding and then solving the contention issues that PGMs traditionally had with wireless links in the e2e path. This allowed us to design a new probing method based on round-trip time measurements, with low intrusiveness and short convergence time, and without the need to know the home network topology a priori. Our tool, Allbest, is accurate enough to make informed decisions about the admission of IPTV streams and the like, and gives the service provider no more information than strictly needed. We have built a prototype and a testbed, and our performance measurements indicate that Allbest works well and outperforms Iperf and Wbest for various topologies based on 100BASE-TX, IEEE 802.11b, and 802.11g, for up to 50 percent crossing traffic.

Our work is opening up a whole field of research related to diagnostics of heterogeneous
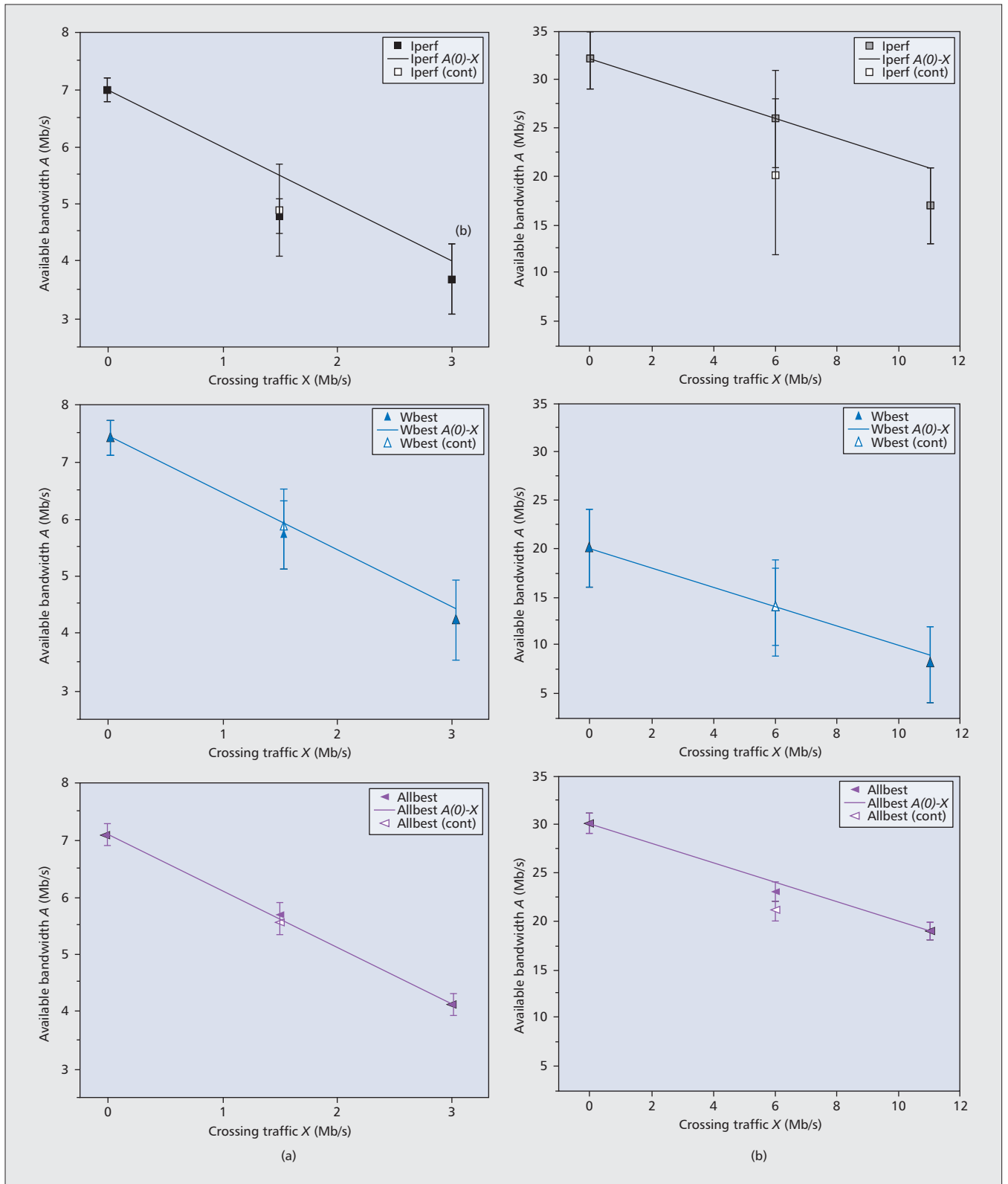
Figure 5. *Part 1. Available bandwidth measured with different tools and different amounts of crossing and contending (cont) traffic. a) topology 1; b) topology 2.*

home networks. Many different configurations and network parameters need to be investigated. New network technologies supporting IP are currently entering the home, such as HomePlug, MoCa, IEEE 1901, IEEE 802.11n, and G.hn. Some of them (e.g., HomePlug) are exhibiting very different physical- and link-layer properties (e.g., fast rate adaptation) from the networks studied in this article. Our method needs to be improved to include these novel techniques. Also different queuing mechanisms than fair queuing should be studied. To increase the accu-

racy of our method up to a level at which it can be used for voice-over-IP services, network tomography techniques may be applied.

Another matter is how the obtained values should lead to intelligent decisions on the service level, which probably needs some form of cross-layer optimization. One of the questions following from this is how frequent a measurement should be repeated. It will certainly depend on the dynamics of the traffic in the home network. It can be safely assumed that crossing traffic in homes cannot be modeled with the stochastic properties of Internet traffic. We are currently performing a series of experiments in Dutch homes in order to understand the in-home traffic dynamics.

Finally, the applicability of our method to other consumer networks should be studied. In-car networks, personal area networks, hotel networks, and others exhibit similar properties and management issues as home networks. At the IEEE CCNC 2011 conference, operators also showed interest in using Allbest for mobile access networks.

REFERENCES

[1] F. T. H. den Hartog *et al.*, "First Experiences with Personal Networks as An Enabling Platform for Service Providers," *Proc. 4th Annual Int'l. Conf. Mobile and Ubiquitous Systems: Networking & Services*, Philadelphia, PA, Aug. 2007, pp. 1–8.
[2] Home Gateway Initiative, "HGI Guideline Document: QoS white paper," HGI-GD013-R2, 2009, http://www.homegateway.org.
[3] C. Develder *et al.*, "Delivering Scalable Video with QoS to the Home," *Telecommunication Systems*, Springer, Online FirstTM, 9 June 2010.
[4] D. Jiang and M. Li, "Quality of Service in the Home Network," *Proc. 2nd Int'l. Conf. Future Generation Commun. and Networking*, Hainan Island, PR China, Dec. 2008, pp. 473–76.
[5] E. Goldoni and M. Schivi, "End-to-End Available Bandwidth Estimation Tools, An Experimental Comparison," *Traffic Monitoring and Analysis*, *LNCS*, Springer, vol. 6003, 2010, pp. 171–82.
[6] R. Prasad *et al.*, "Bandwidth Estimation: Metrics, Measurement Techniques and Tools," *IEEE Network*, vol. 17, no. 6, Nov./Dec. 2003, pp. 27–35.
[7] B. Melander, M. Bjorkman, and P. Gunninberg, "A New End-to-End Probing and Analysis Method for Estimating Bandwidth Bottlenecks," *Proc. IEEE GLOBECOM 2000*, San Francisco, CA, Nov. 2000, vol. 1, pp. 415–20.
[8] V. Jacobson, "Congestion Avoidance and Control," *Proc. ACM Symp. Commun. Architectures and Protocols (SIGCOMM'88)*, Stanford CA, Aug. 1988, pp. 314–29.
[9] A. Delphinanto *et al.*, "Path Capacity Estimation in Heterogeneous, Best-effort, Small-scale IP Networks," *Proc. 35th IEEE Conf. Local Computer Networks*, Denver, CO, Oct. 2010.
[10] W. Kampichler and K. M. Goeschka, "On measuring Quality of Service Limitations in Local Area Networks," *Proc. IEEE ICC '03*, Anchorage, AK, May 2003, vol. 1, pp. 291–95.
[11] M. Li, M. Claypool, and R. Kinicki, "WBest: A Bandwidth Estimation Tool for IEEE 802.11 Wireless Networks," *Proc. 33rd IEEE Conf. Local Computer Networks*, Montreal, Canada, Oct. 2008, pp. 374–81.

BIOS

ARCHI DELPHINANTO [M] received his B.Sc. in engineering physics from Bandung Institute of Technology, Indonesia, in 1998, and his PD.Eng. in ICT from Eindhoven University of Technology, the Netherlands, in 2003. He currently works as a scientist at TNO, the Netherlands, and is also pursuing his Ph.D. degree in the Electrical Engineering Department at Eindhoven University of Technology. His research interests focus on service discovery and access in heterogeneous networks, including interoperability, quality of service, and service monitoring.

TON KOONEN [F] is a full professor at Eindhoven University of Technology (TU/e) since 2001, and chairman of the Electro-Optical Communication Systems group. Before joining TU/e, he worked for more than 20 years in applied research in industry at Bell Laboratories in Lucent Technologies, and as a part-time professor at Twente University. His current research interests are in optical fiber access and in-building networks, including radio-over-fiber techniques. He is a Bell Labs Fellow.

FRANK DEN HARTOG [M] is a senior scientist, Future Internet at TNO, and deputy chair of the Technical Working Group of the worldwide Home Gateway Initiative (HGI). Between 1998 and 2003 he was an applied scientist at KPN Research, where he pioneered the home networking research area. He now acquires and manages large collaborative research projects on every aspect of heterogeneous consumer networking. He has authored over 100 conference articles, journal papers, patents,ß and contributions to standardization.
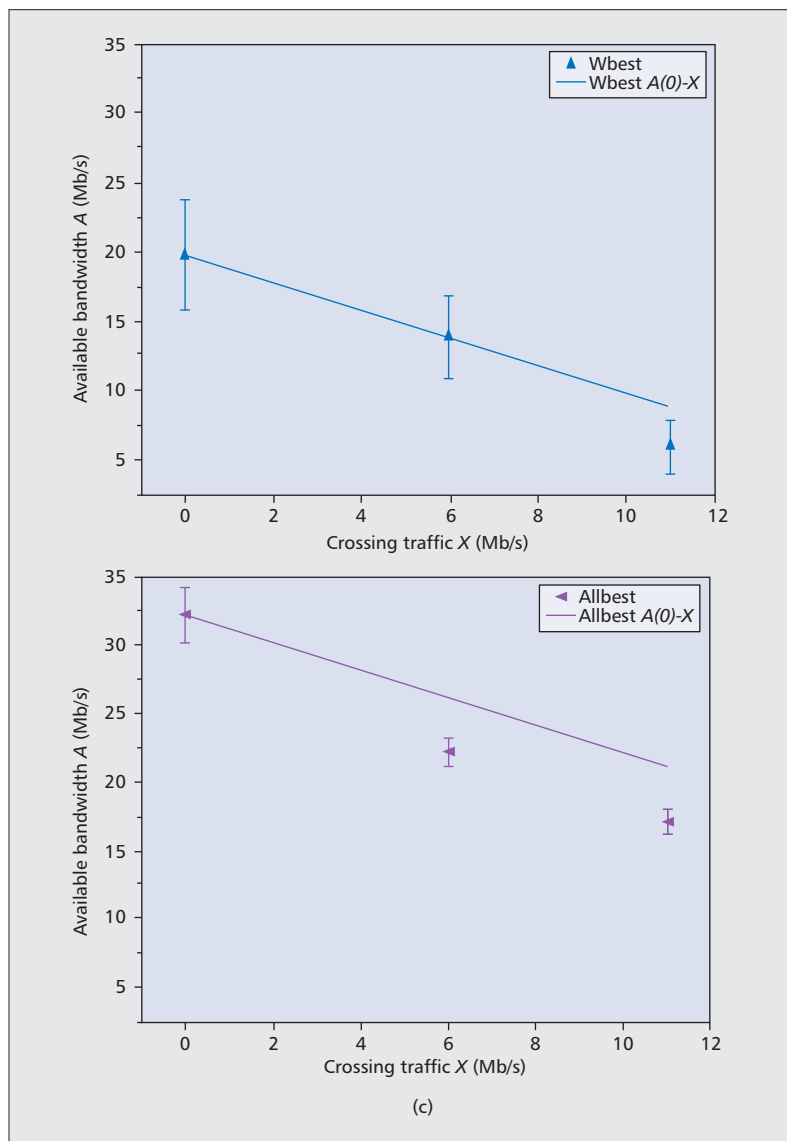
Figure 5. *Part 2. c) Topology 3.*