Internet usage and performance analysis of a rural wireless network in Macha, Zambia

D.L. Johnson, Elizabeth M. Belding,
Kevin Almeroth
University of California, Santa Barbara
davidj,ebelding,kalmeroth@cs.ucsb.edu

Gertjan van Stam
Linknet, Macha, Zambia
gertjan.vanstam@machaworks.org

ABSTRACT

There have been a number of rural wireless networks providing Internet access over the last decade but little is known about how the Internet is being used, how these networks perform and whether they follow similar trends when compared with Internet usage patterns in developed regions. We analyse a set of network traces from the Linknet wireless network in Zambia, which provides Internet access to approximately 300 residents of a rural village using a satellite link and a combination of point-to-point links, hotspots and wireless mesh networks. Our analysis reveals largely web-based traffic as opposed to the peer-to-peer traffic dominance that one finds in urban areas. Social networking sites receive the most hits, and large file downloads from operating system repositories contribute the most to the bandwidth consumption. A number of network pathologies in the gateway as well as the wireless mesh network are also analysed and a set of recommendations conclude the work.

Categories and Subject Descriptors

C.2.2 [**Computer-Communications Networks**]: Network Protocols-*Applications*; C.4 [**Performance of Systems**]: Measurement Techniques

General Terms

Measurement, Performance

Keywords

Rural wireless, Performance analysis, Mesh network, WWW, P2P, Proxy

1. INTRODUCTION

There are many rural wireless networks that have been operational since 2005. Some key examples are Airjaldi in India [12], the Peebles Valley Wireless network in South Africa [7] and the Linknet wireless network in Zambia [1]. These networks are unique in that they need to overcome challenges such as long distances between wireless nodes, low-bandwidth gateways to the Internet, lack of reliable power and high cost of Internet connectivity.

Rural wireless networks usually share a low-bandwidth, costly link to the Internet amongst a large user base. This means that any inefficiencies in the network can render a slow shared Internet link almost unusable. Analysing and understanding the traffic distribution, web usage patterns and source of bottlenecks can facilitate network designs that are optimized to give rural users a better Internet experience and bring down usage costs.

There has been a significant amount of work that has tracked Internet usage behaviour in urban areas over the past decade, but there is a large gap in analysis of Internet usage in the small set of rural wireless networks that are now in existence. For example, most recent Internet usage studies show that over half the Internet traffic is peer-to-peer (P2P) traffic. However, P2P traffic over a satellite link is costly and inefficient, and hence is likely to be less prevalent in a rural network.

In this work we analyse a set of Internet gateway trace logs, proxy access logs and mesh network logs from the Linknet wireless network in Zambia over a period of 2 weeks in February 2010. The wireless network provides Internet access to approximately 300 residents of a rural village, as well as numerous international visitors. The results show that the network is heavily biased towards social networking-based web traffic and much of the potentially cacheable traffic, like Youtube videos, are not cached. A number of networking anomalies are also uncovered such as a large portion of packets with low TTL values that cause "TTL expired" responses and route flapping in the mesh network.

Based on observations from the traffic traces, we make suggestions to deal with the unique challenges of rural networks using slow shared Internet links. Some of the suggestions involve using data ferries, intelligent proxy caching, time-shifting large file downloads and packet filtering. We also suggest improvements for the mesh networks to prevent excessive route flapping. This paper is the first contribution that we are aware of that analyses Internet usage on a large rural wireless network with incorporated wireless mesh networks.

NSDR '10, June 15, 2010, San Francisco, California
Copyright 2010 ACM ISBN: 978-1-4503-0193-0 ...$10.00.

2. RELATED WORK

The most up-to-date multi-year study of the distribution of Internet traffic in the developed world was carried out by Ipoque in 2008/2009 [11]. It showed that P2P traffic had decreased significantly and Web traffic had increased due to the extensive use of content servers. In Germany, for example, P2P traffic decreased from 69% to 52% and Web traffic rose from 14% to 26%.

The work that is the closest to ours is one that examines Web traffic usage in Internet Cafes and community centres in Cambodia and Ghana [5]. This paper only studied HTTP traffic and there was no wireless network aggregating traffic to the Internet connection. The traffic classification revealed that portal sites like Yahoo and MSN, advertisement sites and multimedia content accounted for the bulk of the traffic.

Wireless network performance in rural networks has been analysed in [2] and [12] but the Peebles Valley mesh network in South Africa [7] had the most similar network configuration. Analysis showed that in an unloaded network, the mesh would become the bottleneck, rather than the satellite, after 7 hops.

3. BACKGROUND AND NETWORK ARCHITECTURE

We briefly summarize the Linknet network in Macha, Zambia and refer the reader to [9] for more detail. Macha is a typical poor rural area in Africa with scattered homesteads, very little infrastructure and people living a subsistence lifestyle. However, in the middle of this rural area is a mission hospital, medical research institute and community centre that has provided connectivity to approximately 300 community workers and visitors since 2004 using satellite-based Internet. During the period of this measurement 7% of the users were visitors and the rest were local Zambian nationals. Over the past few years the Internet connection has been spread to a large portion of staff in the area, as well as the community centre, which has an Internet Cafe. The VSAT connection has a committed download speed of 128 kbps bursting to 1 Mbps and a committed upload speed of 64 kbps bursting to 256 kbps with no monthly maximum. The total monthly cost of the C-band VSAT connection is $1200 (US dollars).

Figure 1 shows a scaled view of the core of the Linknet network with the position of each wireless router represented by a symbol. Two radio masts, one shown in the diagram at the IT tower, spread the connectivity over a 6 square kilometre area using a combination of 802.11 point-to-point links, hotspots and mesh networks. The gateway server was installed with *ClarkConnect*[1], which is a pre-configured firewall and Internet gateway. It has no user-management functionality, but some rudimentary user-management was done by checking host-names and their associated IP addresses in cases where bandwidth was used excessively.

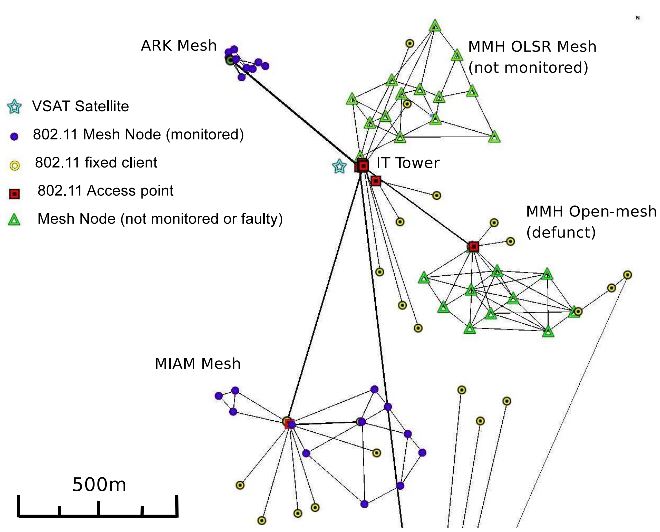

Figure 1: Scaled view of the Macha network.

4. GOALS AND MEASUREMENT PROCESS

Our goals are three-fold: The first goal is to understand the usage patterns of users in a rural context to evaluate differences from an urban setting and gain insight into the needs of users in a rural setting. The next goal is to understand the performance of the network from the application level down to the physical level to find out what unique challenges are prevalent in rural networks. Our final goal is to make use of the learning from the first two goals and suggest ways in which the performance can be improved.

To meet these goals, two measurements points were required. The first measurement point was located at the gateway and captured all Internet traffic on the interface to the satellite and to the wireless network. The packets were captured in pcap format and a capture length (snaplen) of 96 bytes was used to minimize the size of the log file. This snaplen size was chosen to capture enough information from all the headers in the network packet. In order to analyse the HTTP traffic, the squid proxy access logs were also archived for analysis. All IP addresses were anonymized in order to protect the privacy of the users.

The second measurement point was located at the mesh gateway nodes to monitor the quality of the links in the mesh network. The monitoring daemon sent the ETX[2] values of every link in the mesh back to the gateway server every minute. The number of hops taken by mesh network packets was calculated by examining the TTL values of packets entering the gateway from the mesh.

14 days of traffic were captured from midnight, Sunday 31 January to midnight Sunday 14 February in 2010. Approximately 50 GB of packets were captured, consisting of about 6 million packet flows. Captured traffic was compressed and sent to the USA for analysis during off-peak hours. User-management software installed in April 2010 established that 10% of the traffic was from International visitors. A similar traffic distribution between the local population and visitors was likely for February but this could not be used to establish the type of traffic visitors were generating in the original data set.

5. TRAFFIC USAGE ANALYSIS

In this section we present an analysis of the usage patterns from squid proxy logs over 14 days organised into 1 hour bins. The plot in Figure 2a shows the total number of requests per hour as well as the total megabytes sent and received per hour. It follows a typical diurnal pattern but the off-peak period is very short due to users staying up late or waking up early to make full use of the extra available bandwidth during these hours. The number of requests generally tracks the amount of data downloaded except, for example, early Friday and Saturday morning where a single user was very active.

[1] http://www.clarkconnect.com/info/

[2] Expected Transmission Count (ETX) measures packet loss and is used as a routing metric.

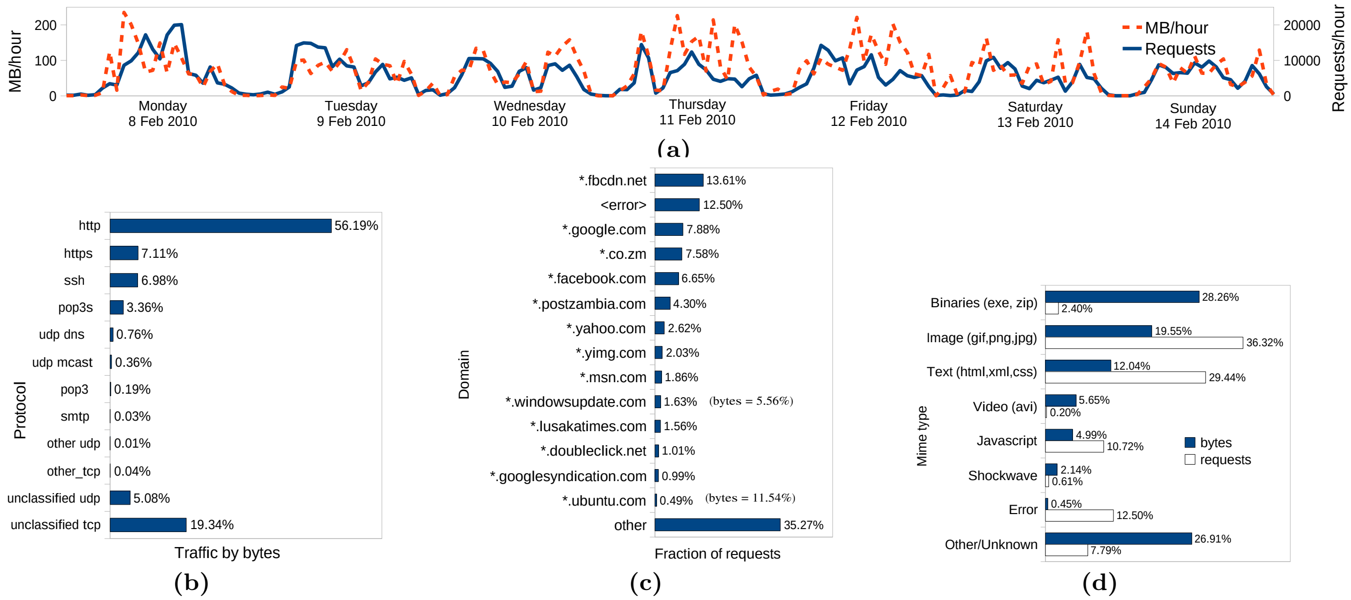

Figure 2: Usage analysis over 14 days: (a) Plot of the total web traffic for the satellite link over the second week; (b) Traffic classification by protocol; (c) Web traffic classification by domain visited; (d) Breakdown of traffic by MIME Type.

The proxy had a cache hit rate of 43% with an actual bandwidth saving of 19.59%. This low fraction of bandwidth saved is fairly common due to the dynamic nature of the Internet today; we offer some recommendations in section 7 to improve this. The cache size was set to 1 gigabyte and studies have shown there is very little gain from increasing the cache size beyond 1 gigabyte [5].

5.1 Content distribution

To help understand the spread of traffic types, we classify packets as either TCP or UDP and then further classify these by their known Internet Assigned Numbers Authority (IANA) port numbers. Combined outbound and inbound traffic from the gateway to the satellite link was analysed.

Figure 2b shows the results of our classification. TCP traffic accounts for 93.24% of the traffic, which is consistent with most modern Internet usage trends. Web traffic accounts for 68.45% when standard HTTP and secure HTTP are combined and is clearly the dominant protocol. This is in sharp contrast to developed countries in which 2008/2009 studies show that Web traffic accounts for between 16% and 34% of Internet traffic due to the high prevalence of P2P traffic [11]. The large portion of ssh traffic was due to our traffic downloads.

There was also a significant amount of email traffic using pop3 store-and-forward delivery models. This traffic primarily occurred early in the morning before people arrived at work and made optimal use of off-peak hours. 26.47% of traffic could not be classified with simple port based techniques. This most likely consisted of Skype traffic, which is extensively used in Macha, as well as applications not using known assigned IANA ports. In future work we will use more advanced packet inspection tools to understand the breakdown of this unclassified traffic.

5.2 URL analysis

As most of the traffic is Web based, studying the sites that are visited as well as the traffic types will provide useful insights into user behaviour. Figure 2c classifies web traffic into the top 15 site domains. The most startling pattern that emerges is the dominance of Facebook. The Facebook host site and CDN make up 20.26% of the total requests. This is close to 3 times greater than the next most visited site, which is Google. A recent news article reported that Facebook overtook Google in the US in terms of number of hits in March 2010 [4]. From this data it appears that the crossover point in Macha occurred many months ago. Portals such as Google, Yahoo (*.yahoo.com, *.yimg.com) and MSN are clearly the next dominant site category (14.39% combined).

Local web sites in Zambia form the third most dominant category. This is encouraging in the sense that local relevant content and language is available to Zambians. Zambian news sites are also in the top 15 web sites visited, showing that digital news is a viable alternative when newspapers and radio are not available locally. Downloading software packages and updates from operating system sites such as Windows and Ubuntu are also in the top 15. Although they show a relatively low hit rate, the Ubuntu site, for example, has the highest share of traffic by bytes. Web site access follows a common pattern of a relatively small number of popular URLs and a long tail of other URLs. This long tail is the cause of a large pool of URLs in the "other" category.

2% of the total web requests are to known advertising domains and there are most likely many others that could be extracted from the "other" category by using pattern matching lists. The key issue, which is also raised in [5], is that the target population in developing regions is highly unlikely to be a customer of the advertised services. This essentially constitutes wasted bandwidth, which is particularly detrimental over a slow satellite link.

5.3 Traffic type and sizes

Because certain visited sites warrant large file downloads, we now move our focus to correlating types of web traffic with their size distributions. Figure 2d shows the MIME types of the traffic by bytes. There is a clear penchant towards binary downloads with large file sizes. The URL analysis in the previous section revealed that the sites that consume the most bandwidth are Windows update sites and package distribution sites for Linux. This corroborates with the evidence of binary files contributing to the largest fraction of bandwidth consumed. Large file sizes make up the tail end of the distribution, but they contribute to a large fraction of the traffic. For example, 99.9% of the traffic contains objects that are less than a megabyte, but the remaining 0.1% of the objects contribute to 50% of the total bytes downloaded.

Images and text form the bulk of the traffic types due to the dominance of web traffic in the network. The large amount of Javascript can be attributed to an increasing use of web applications when using web based email, for example. Multimedia also plays a significant role in the Linknet network as can be seen from the Video and Shockwave traffic, which together make up 7.79% of the traffic. Inspection of the squid proxy logs showed that a significant amount of Youtube requests were made but all resulted in a cache miss. Youtube videos are served by a CDN and cannot be cached using a traditional proxy caching application like squid, which depends on a unique IP address and URI combination for its lookup.

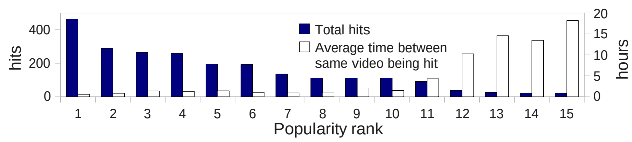

Figure 3: Youtube popularity.

We rank the 3162 Youtube hits over the two week measurement period according to total number of requests to the same video. The results of this analysis are shown in Figure 3 for the 15 most popular videos. There were 451 unique videos, but the top 15 ranked videos contained 75% of the total requests. For the top ranked video, there were instances where the video was requested by six different users within a 10 minute interval. This would saturate the satellite link and the downloads would fail or be abandoned by the users. Anecdotal evidence from users in the Zambian network confirms this.

Youtube could have potentially consumed 60% of the satellite bandwidth if all 3162 Youtube videos were downloaded. However, Figure 2d shows that only 2.14% of the traffic by bytes is shockwave traffic. This is a result of about 87% of streaming media being abandoned by users in the first 10 seconds [6] as well as potentially slow or non-responsive Youtube downloads due to congestion. If a more advanced caching server was used and every requested Youtube video was cached in full, 86% of the bandwidth used for Youtube requests could have been saved. This is calculated by noting that there were 451 unique Youtube videos and 3162 Youtube video hits. Some of the repeated hits were due to users visiting a web site containing an embedded Youtube video with the auto-play option set. Future work will quantify the extent of this problem as well as ways to limit its impact.

6. NETWORK PERFORMANCE AND TROUBLESHOOTING

This section seeks to understand how the limited satellite connection bandwidth is distributed amongst the potential 300 users in Macha and how the mesh network performs under load. To do so we use traffic traces divided into 10 minute bins from the gateway and make use of mesh network traces from the mesh gateway nodes.

Figure 4a shows the number of unique IP addresses that were present in each 10 minute interval. This gives some indication of the number of concurrent users that were active in the network and ranged between 40 on weekdays to 25 on weekends. The four dips on Wednesday and Thursday of between 10 minutes and 2 hours were due to power failures. Some of these dips do not drop to zero as a few wireless routers are on UPS backup and laptops can also absorb short power failures. The mesh networks were only reflected as a single IP address and they would likely yield up to 5 additional concurrent users in a 10 minute window.

Figure 4b shows the total throughput on the gateway interfaces to the wireless network and the maximum throughput to a single IP destination in the network. This shows very clearly that there were single users that consume a large fraction of the bandwidth (50% on average). A single user would most likely be downloading large files while the balance of users were involved in more interactive activities such as web browsing.

Only the MIAM mesh was analysed as the ARK mesh gateway node was typically powered off or not as active as the MIAM mesh. Figure 4c confirms that the MIAM mesh was not a large contributor to the overall bandwidth consumption. Further analysis shows that it consumed the maximum bandwidth in the Linknet network only 4.74% of the time. The peak amount of traffic from the gateway to the MIAM mesh was 30 MB in 10 minutes, whereas single machines could achieve up to 80 MB. This is most likely due to a combination of a poor quality wireless link from the mesh network to the gateway and route flapping, which is discussed in Section 6.2.

There is a clear diurnal pattern in the traffic traces where almost no traffic was present between midnight and 6:00 AM. This pattern is broken on Friday night and early Saturday morning, where there was a continuous stream of activity from a low number of users. Analysis of this period of high activity showed that 52.86% of this traffic was from unclassified TCP ports compared to the 19.34% average on unclassified TCP ports over the 14 day measurement period. This may be torrent traffic for weekend entertainment. Note that there is no TV or radio reception in this area and no DVD rental store nearby.

The flow distribution over the 14 day measurement window exhibits a long tailed distribution with the most demanding flow using 22% of the bandwidth on average but as much as 90% of the bandwidth during quieter periods. The average flow size is 8.4 kB and the average flow length lasts 70 seconds. The top 0.1% of the largest flows contain 50% of the data; this correlates with earlier analysis of traffic types and sizes.

6.1 Performance analysis

To determine the source of networking problems, ICMP messages and HTTP messages were analysed. There were 828 “TTL expired in transit” messages *per hour* and 3900 “destination host unreachable” messages *per hour*. Further analysis showed that “TTL expired” messages were caused by the same number of packets *per hour* leaving the mesh network with TTL set to 3, 4 and 5. This made up 7% of the total traffic by packet count leaving the mesh networks. These were UDP DNS requests and were most likely caused

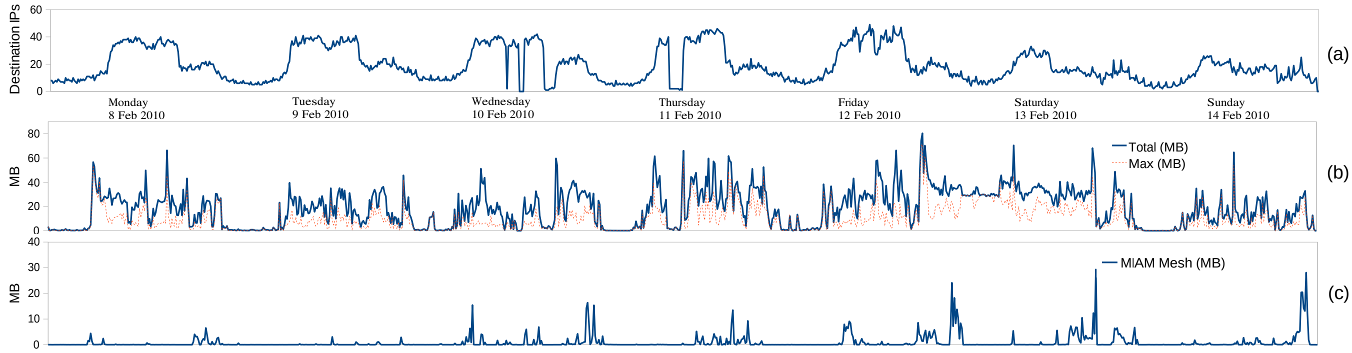

Figure 4: The distribution of all inbound traffic on the satellite link in the final week of the 14 day measurement period using 10 minute bins: (a) Total number of unique IP addresses detected; (b) Total traffic vs maximum traffic for a single IP address; (c) Total bandwidth used in the MIAM mesh.

by a malfunctioning DNS server or possibly a computer virus that wastes satellite bandwidth. Either the packets with TTLs below 32 should be discarded or the TTL should be reset before leaving the gateway.

Looking at the HTTP messages, there are 723 "503 Service not available" messages *per hour* making up 9.75% of all HTTP message responses to an HTTP GET. This is a very large fraction of HTTP requests not successfully completed. There are two reasons that this could occur. The first is that the DNS request could not be completed by the gateway server. The second is that there was no response to the HTTP GET from the upstream server. Both would be due to congestion in the network and confirm that interactive browsing is being hindered by a fully saturated satellite link often due to large downloads. Another possible cause of network congestion is a large amount of computer virus traffic. Many computers in Macha do not have updated virus scanners and it is highly likely that they are infected.

6.2 Mesh analysis

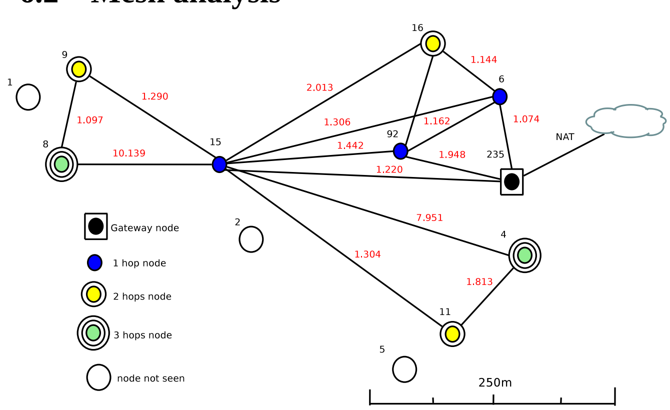

Figure 5: Scaled snapshot of the miam mesh during the 2 week measurement period with ETX values indicated per link. Note that the number hops is determined by the route with the lowest sum of ETX values.

The two mesh networks described in Section 3 make use of the Optimized Link State Routing (OLSR) protocol with the ETX link metric. An in depth study of the latency and throughput performance of the MIAM mesh network was carried out using active measurement techniques [1]. The 1 hop routers could achieve an average of 5 Mbps to the gateway router, but 2 and 3 hop neighbours could not achieve over 500 kbps. As a result, when the satellite link

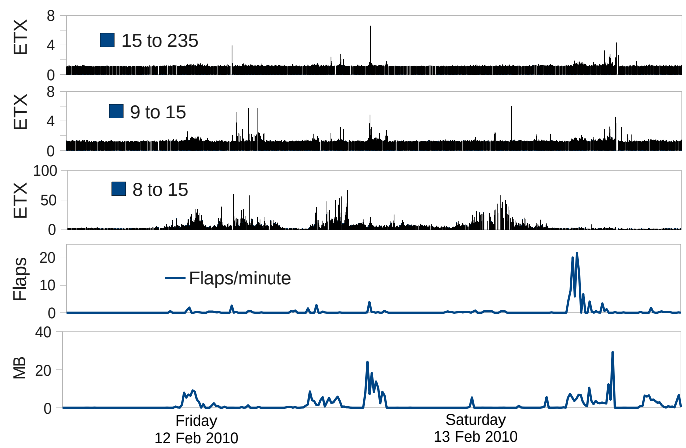

Figure 6: Graph showing how load on the mesh network affects the ETX link metric as well as route flapping.

is not saturated, the bottleneck could become the mesh for 2 and 3 hop neighbours, but this was never observed in the 14 day measurement period.

Figure 5 shows the connectivity between all nodes in the mesh network and a snapshot of the ETX values present in the mesh at an instant in time. The average node degree in the network is 3.33 and no node has less than two potential links. However, some of the links are very poor with ETX values above 5. As an example, if node 9 or 11 go down, node 4 and 8 would need to route through very poor links. To understand the stability of the mesh network, the ETX values were tracked over 2 weeks. A Friday and Saturday portion of this is shown in Figure 6 together with the load on the mesh network.

It is clear that as the load on the network increases, the ETX of the weaker links tends to increase rapidly. Another effect of the changing ETX values is route flapping; this is also shown in Figure 6. A small amount of route changing can usually benefit a network as the routing protocol tries to discover a more optimal route. However, route flapping that approaches two route changes per second, as seen here, is detrimental to network flows.

7. RECOMMENDATIONS

There are two key components of rural wireless networks that need optimization to connect to the Internet through low-bandwidth links. The first is the gateway server and the second is the wireless mesh network.

7.1 Gateway server

Due to the limited amount of expert networking skills in the area, a *ClarkConnect* pre-configured gateway server was used in Macha. However, this and other similar pre-configured gateways are not optimized for severely constrained Internet up-links and the following customizations would have a significant positive impact on the network:

- Changing the caching behaviour of squid to be able to identify when the same content is being served by a different URL when accessing content from a CDN like Youtube. This can be done by rewriting the URL request to only preserve the static video ID. Other techniques, such as value-based web caching, deal with dynamic URLs by generating indices based on document content [5].
- Making use of cache optimizations such as HTTP chunking, time shifting and data compression discussed in [5], which would improve the throughput and response times for users. Time shifting would disallow large file downloads during peak hours and delay the download to off-peak hours.
- Filtering requests to advertising domains such as *doubleclick.net*.
- Filtering requests that are guaranteed to fail, such as packets with small TTL values.
- Installing a network monitoring tool, together with traffic measurement daemons at strategic congestion points, to help a network administrator understand network usage and performance as well as diagnose network problems. For per-user monitoring, the network subnets will need to be re-designed to avoid using NAT at the mesh gateways.
- Making use of a data-ferry synchronization station using a dedicated small-footprint PC such as "PC in a plug" and external hard drive. This will be used to leverage frequent travellers visiting cities with low-cost high-capacity Internet links as data-ferries to fetch large popular repositories such as Ubuntu.

The Macha network is in the process of extending bandwidth and installing user authentication systems with preset and varied "service level agreements". This should help alleviate some congestion, especially once monetary value is associated with the number of bytes downloaded and the unit value increases during busy times. User management is a sensitive issue and it is highly likely that the set of rules will evolve as the balance between what is best for the community and individual users is accommodated.

7.2 Mesh network

In a similar fashion to the gateway server, the OLSR mesh network utilized in Macha was chosen based on being widely adopted and well documented. However it has been shown to be a sub-optimal solution and the following possible improvements could be made:

- Replace the OLSR routing protocol with a protocol that has proven better performance; protocols such as ExOR [3] or B.A.T.M.A.N. [8] perform better in the presence of poor links.
- If there is still a desire to continue to use OLSR, then replace the ETX link metric with Expected Transmission Time (ETT), which improves the overall throughput.
- Install mesh network monitoring daemons that make use of differential monitoring [10] on the wireless routers and send this information back to the gateway server.

8. CONCLUSION

It is clear from these results that the behaviour of the Internet in a rural wireless network connecting through a satellite is very different from that of an urban network in a developed region. From the usage analysis, it is apparent that the traffic is mostly web based; however, most of the traffic is not cached even though some simple proxy changes could make this traffic cacheable. A significant portion of the traffic consists of large binary files to operating system repositories. Most of this traffic could be eliminated by using people who travel to cities as data-ferries. Bandwidth is also wasted on advertisement servers, which have a very small likelihood of being relevant, as well as outgoing packets with small TTLs, which are likely to fail. The mesh network routes flap severely under load. Other routing strategies such as stigmergic or opportunistic routing will most likely be better suited to this environment.

More attention should be given to building pre-packaged networking solutions for rural wireless networks that are cognizant of the characteristics that have been highlighted in this paper. Some of the problems were addressed in the early days of the Internet when last-mile access was similar to what is currently experienced in rural networks, but many of the problems are new as the Internet has become more dynamic and media-rich. As the average web page size continues to grow, the digital divide will widen unless innovative networking techniques, which mitigate these increasing bandwidth demands, are employed.

9. ACKNOWLEDGEMENTS

The authors would like to thank Linknet in Macha, Zambia for their cooperation on this project.

10. REFERENCES

[1] J. Backens, G. Mweemba, and G. van Stam. A Rural Implementation of a 52 Node Mixed Wireless Mesh Network in Macha, Zambia. In *Africomm '09*, December 2009.

[2] P. Bhagwat, B. Raman, and D. Sanghi. Turning 802.11 inside-out. *ACM SIGCOMM Computer Communication Review*, 34(1):33, January 2004.

[3] S. Biswas and R. Morris. ExOR: opportunistic multi-hop routing for wireless networks. *Computer Communication Review*, 35(4):133, 2005.

[4] N. Clark. Facebook overtakes google, March 2010. `http://www.independent.co.uk/news/business/news/facebook-overtakes-google-1923102.html`.

[5] B. Du, M. Demmer, and E. Brewer. Analysis of WWW traffic in Cambodia and Ghana. In *World Wide Web conference WWW '06*, May 2006.

[6] L. Guo, S. Chen, Z. Xiao, and X. Zhang. Analysis of multimedia workloads with implications for Internet streaming. In *World Wide Web conference WWW '05*, May 2005.

[7] D. Johnson. Evaluation of a single radio rural mesh network in south africa. In *ICTD '07*, December 2007.

[8] D. Johnson, N. Ntlatlapa, and C. Aichele. A simple pragmatic approach to mesh routing using BATMAN. In *IFIP International Symposium on Wireless Communications and Information Technology in Developing Countries*, October 2008.

[9] K. Matthee, G. Mweemba, A. Pais, G. Van Stam, and M. Rijken. Bringing Internet connectivity to rural Zambia using a collaborative approach. In *ICTD '07*, December 2007.

[10] R. Raghavendra, P. Acharya, E. Belding, and K. Almeroth. MeshMon: a multi-tiered framework for wireless mesh networkmonitoring. In *MobiHoc S3 workshop*, May 2009.

[11] H. Schulze and K. Mochalski. ipoque Internet Study 2008/2009, 2009.

[12] S. Surana, R. Patra, S. Nedevschi, M. Ramos, L. Subramanian, Y. Ben-David, and E. Brewer. Beyond pilots: Keeping rural wireless networks alive. In *NSDI*, April 2008.