Comprehensive Mobile Bandwidth Traces from Vehicular Networks

Ayub Bokani[1], Mahbub Hassan[2], Salil S. Kanhere[3], Jun Yao[4], Garson Zhong[5]
[1,2,3] Email: {Bokani, Mahbub.Hassan, Salil.Kanhere}@unsw.edu.au - Data61|CSIRO
[4,5] Email: {jyao, gzho871}@cse.unsw.edu.au

ABSTRACT

Bandwidth fluctuation in mobile networks severely effects the quality of service (QoS) of bandwidth-sensitive applications such as video streaming. Using bandwidth statistics it is possible to predict the network behaviour and take proactive actions to counter network fluctuations, which in turn can improve the QoS. In this paper, we present comprehensive bandwidth datasets from extensive measurement campaigns conducted in Sydney on both 3G and 4G networks under vehicular driving conditions. A particularly distinguishing feature of our dataset is that we have collected data from repeated trips along a few routes. Thus our data can be useful to obtain statistically significant results on network performance in an urban setting. We outline the measurement methodology and present key insights obtained from the collected traces. We have made our dataset available to the wider research community.

CCS Concepts

•Networks → Location based services;

Keywords

Bandwidth dataset; Adaptive video streaming

1. INTRODUCTION

We have witnessed a significant growth in mobile data traffic all around the world. It is reported in Cisco's "Global Mobile Data Traffic Forecast" [9], that mobile data traffic will grow from 4.2 Exabyte (EB) in 2015 to 24.3 EB by 2019. This growth is the consequence of the significant rise in the number of smartphones and mobile subscribers. Moreover, there is also a rapid increase in bandwidth-intensive applications such as video streaming. While the peak data rates have increased with the introduction of 4G technologies, bandwidth fluctuations are still very common, particularly in high mobility scenarios such as when traveling in a vehicle. Ensuring good Quality of Service (QoS), particularly for bandwidth-sensitive applications such as video streaming thus becomes a challenge. Insights obtained from studying empirical bandwidth data can be crucial in improving QoS for such applications.

The state-of-the-art video streaming technologies [1, 2, 24] make use of adaptive streaming to deal with frequent bandwidth fluctuations. In this method, the media server stores different quality versions of the same video file to allow real-time quality adaptation of the video due to network bandwidth variation experienced by a client. It has also been shown that historical information about bandwidth variations could be useful for estimating the current bandwidth conditions and proactively adjusting the streaming quality to improve QoS. Yao et al. [23], Halvorsen et al. [18], Curcio et al [10], and Singh et al. [20] for instance, have investigated the use of bandwidth statistics to better estimate the available bandwidth at the next visited location and improve the video streaming performance. We also used historical bandwidth information in our previous studies [3–7, 25] to estimate the network throughput in next moment and improve the overall streaming quality. In these studies we used Markov Decision Process (MDP) as the underlying optimization framework and found that, comparing to a well-known non-MDP method, it increases the streaming quality significantly when the bandwidth model is known *a priori*.

There have been prior attempts in conducting empirical mobile bandwidth measurements to study network behaviour and characteristics. Li et al., [13] and Xiao et al., [21], for example, collected high-speed mobile network bandwidth samples from different networks and roads in China. However, they did not consider multiple trips in metropolitan areas and their measurements only includes inter-city roads. Therefore, their dataset is not very representative, since the network conditions can fluctuate significantly from one trip to the next along the same route (as observed in our experiments). Moreover, their data does not provide micro-scale statistics about bandwidth variations within a metropolitan area. Deshpande et al., [11] also measured the network throughput in vehicular environment over 3G and WiFi networks in 2010. They considered two driving scenarios of 500 and 9 mile routes on different New York City roads. In the longer drive scenario, they only considered 2 trips and in shorter one they repeated their measurements 10 times for the same route. Riiser et al. [19] conducted another valuable real-world bandwidth measurement campaign over 3G network and made this dataset publicly available. They collected throughput samples while performing adaptive HTTP streaming on mobile devices in multiple popular

MMSys '16 Klagenfurt am Worthersee, Austria

DOI: http://dx.doi.org/10.1145/2910017.2910618

routes around Oslo (Norway). In their measurements they considered different travelling types including private vehicle and public transport systems (i.e., bus, train, metro, tram and ferry). This dataset presents a limited number of traces (i.e., from 1 to 17 traces) for each particular route. Although these bandwidth traces give a high resolution picture about how the bandwidth is distributed in different locations, they can not be sufficient when more fine-grained network throughput statistics are required for specific purposes. In our previous study [5] for instance, we found that using more historical bandwidth traces provides higher prediction accuracy for the next available bandwidth. We showed how our MDP model can achieve better video streaming performance when more historical bandwidth samples are used to create a bandwidth model.

Therefore we chose to collect our own bandwidth dataset with a particular focus on collecting data from repeated trips along a limited set of routes. These repeated measurements will allow us to obtain statistically significant insights about micro-scale bandwidth variations. In this paper we discuss our two bandwidth measurement campaigns that are conducted in 2008 (3G) and 2015 (3G and 4G). Our bandwidth dataset includes 56,754 samples from 3G and 4G *real world* networks. Each sample consists of measurement time, fetching time of a particular file that is used in our measurements, the bandwidth value and device's geographical location on that moment. We made this large dataset freely available on: "https://github.com/aubokani/Bandwidth-Dataset.git".

The rest of this paper is organised as follows. We discuss our first bandwidth measurement campaign in Section 2. In Section 3 we discuss how our Android bandwidth measurement application is developed. Our second bandwidth measurement campaign is explained in Section 4 followed by some examples of using this dataset in Section 5. We finally conclude the paper in Section 6.

2. BANDWIDTH MEASUREMENT CAMPAIGN 1: 3G

We developed a simple client-server measurement system using off-the-shelf hardware. The server was housed in our lab at the University of New South Wales (UNSW). The client (as shown in Fig. 1) comprised of two Soekris Net4521 boards interconnected via 10 Mbps Ethernet and configured in master/slave mode. The boards were enclosed in a protective casing, connected to a power generator and housed in the boot of a car. Three PCMCIA cellular modems were housed in the system (two in the master board and one in the slave board). To account for network and technology diversity, we simultaneously conducted measurements over two HSDPA [17] networks (A and B) offering different service rates and a network (C) supporting a pre-WiMAX proprietary standard, iBurst [12]. To enhance the wireless signal reception, the cellular modems were connected to external antennas mounted on the car windshield. A Garmin GPS18 GPS sensor was connected to the client for recording the vehicle location.

We developed a lightweight packet-train program to measure the WWAN bandwidth, which achieves fast convergence and generates minimal network traffic. We refer readers to [22] for further details about the program and validations. We collected one bandwidth sample for approximately every 10 second. The samples are tagged with location co-

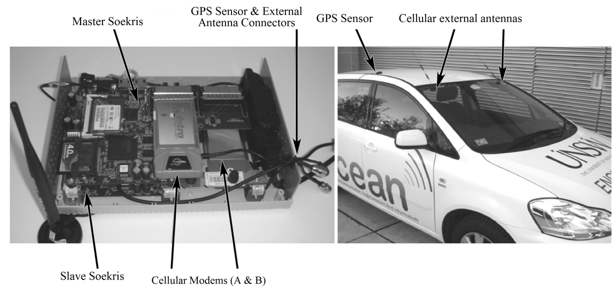


Figure 1: Client hardware

Table 1: Illustrative example of 3G network bandwidth traces

	time	latitude	longitude	bandwidth (Kbps)
1	1186549400	-33.919785	151.228913	1663.1440
2	1186549410	-33.919635	151.227787	1964.7330
3	1186549420	-33.91958	151.227322	2038.8659
4	1186549430	-33.91958	151.227322	2011.2631
5	1186549440	-33.91953	151.22692	1838.6578
6	1186549450	-33.91905	151.226322	1208.2767

ordinates and time, and stored in a repository. Table 1 illustrates 6 example of collected samples. As it's shown, the first column shows the samples number, the second column represents the time when samples are recorded, the third and forth columns present the latitude and longitude respectively and the last column includes the measured bandwidth.

We collected bandwidth samples by driving the car along a distinct route in the Sydney metropolitan area in 2008. Fig. 2 depicts the trajectory of the chosen route. The starting point of this 24 Km route is the University of New South Wales (UNSW) in the eastern suburbs, with the final destination being Macquarie University (MQ) located in the northwestern suburbs. The chosen route is a typical representation of daily commute. In total, 41,130 samples were collected in 71 trips. Fig. 3 illustrates the empirical and normal distribution of bandwidth samples as well as the bandwidth fluctuation from this 3G dataset. As it's shown, the empirical CDF is very close to the Normal CDF of bandwidth samples. This is a true indicator that using many bandwidth traces and assuming normal distribution, the network behaviour can be estimated.

3. ANDROID BANDWIDTH MEASUREMENT APPLICATION

With rapid evolution of network technologies, we witnessed a tremendous growth on quality and coverage of mobile data signals. Nowadays, 4G networks outperform WiFi and provide several times greater performance than 3G [8]. Such ubiquitous deployment of cellular data networks, has increased the demand for internet access among mobile users drastically. Recall that studying on network behaviour and characteristics can have important role in increasing the QoS of different applications. This has motivated us to conduct another bandwidth measurement campaign and build a 4G bandwidth dataset. On the other hand, due to the frequent changes and improvements on 3G technologies, there was a need to collect up-to-date 3G bandwidth samples as well. Therefore we chose to conduct our second bandwidth measurement campaign over both 3G and 4G cellular networks

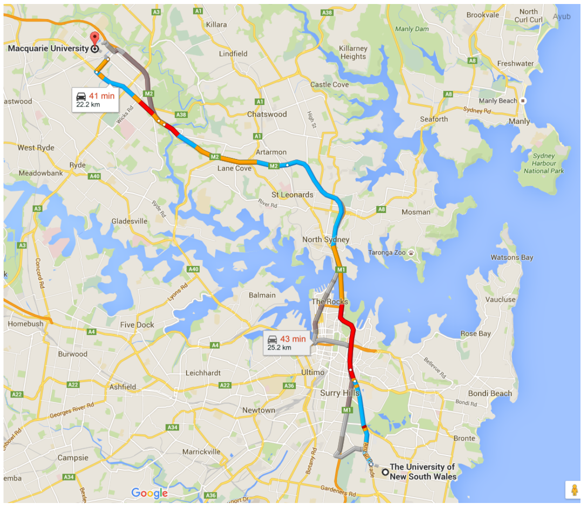


Figure 2: UNSW - MQ Univeristy, source: Google map

in 2015. In these measurements, we considered using a simple user-friendly mobile application instead of our previous bandwidth measurement system mainly for following reasons:

- Ease of use for future measurements.
- Enormous improvement in smartphones with numerous built-in features such as accurate GPS sensors.

Our desired mobile application should ideally be able to measure and store downstream bandwidth characteristics from any given network. Although there are several well-known bandwidth measurement applications such as Speed-Test.net [14] and OzSpeedTest [15] nowadays, we chose to develop our own application for two reasons: First, these services do not provide an option to the users to record their own measured bandwidth samples and build their own statistics. Second, a homegrown application allows us to collect data at the granularity that we desire. This is important for our measurements that supposed to be conducted in a fast moving environment and we need to collect samples repeatedly at a fine granularity.

After testing our initial design, the application's user interface (UI) which had several buttons and users had to set some features manually, was revised multiple times. The final design includes one button only to start and stop performing bandwidth measurements. The functionality of the application was modified to include automatic upload of each measurement to a database as they are completed, ensuring that no bandwidth measurements are lost due to human error. Figure 4 demonstrates two screen shots of our application before and after pressing the *Start Collecting* button.

3.1 Data Structure/Schema

Listing 1 presents the list of data we chose to collect from the mobile client. As it's shown, the sampling time, file size, download duration and time, geographical coordinates before and after file download, network operator's information and country name are all recorded along with each bandwidth sample. In our design, we considered *active* approach

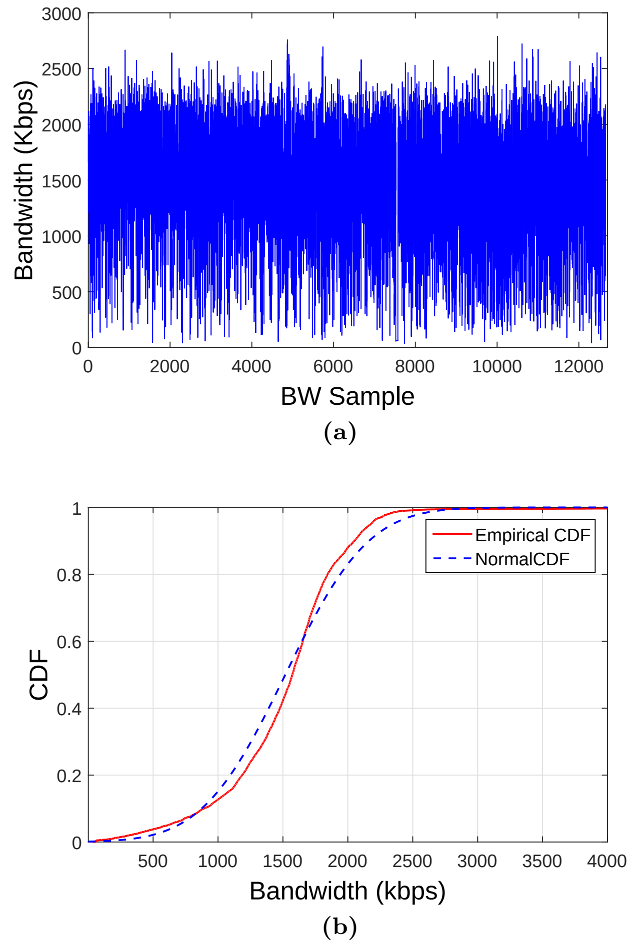


Figure 3: 3G-V1 Badnwidth dataset: (a) bandwidth fluctuation, (b) Empirical Vs Normal CDFs of samples with $\mu = 1518.35$ Kbps and $\sigma = 503.10$ Kbps

which obtains measurements by actively downloading a 1 MB file from UNSW-CSE web server using the HTTP protocol and attempts to saturate the connection. Our bandwidth measurement approach closely follows the methods used by the Ookla SpeedTest.net application [14].

Listing 1: Data Structure in our Android bandwidth measurement tool.

```
<?xml version=" 1 . 0 " encoding="UTF-8"?>
<BWData>
    <DateTime>26/05/2014 12:03:03 </DateTime>
    <DownloadSize>8388608</DownloadSize>
    <DownloadDuration>3.26</DownloadDuration>
    <DownloadRate>2512.1128</DownloadRate>
    <Latitude>-33.94764671</ Lat i tude>
    <Longi tude>151.12730787</Longi tude>
    <EndLatitude>-33.94764671</EndLatitude>
    <EndLongitude>151.12730787</EndLongitude>
    <OperatorCode>50502</OperatorCode>
    <OperatorName>Virgin </OperatorName>
    <CountryISO>au</CountryISO>
 </BWData>
```

3.2 Database Server

In building our bandwidth measurement system, the underlying network architecture is an important design con-

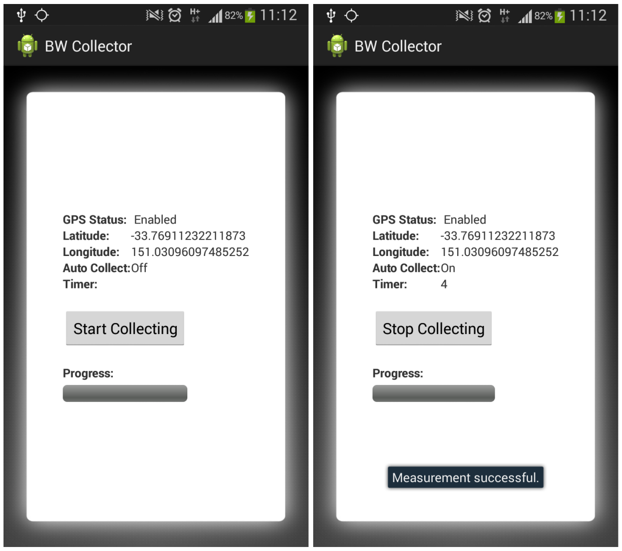


Figure 4: Android bandwidth measurement application: before and after pressing the button.

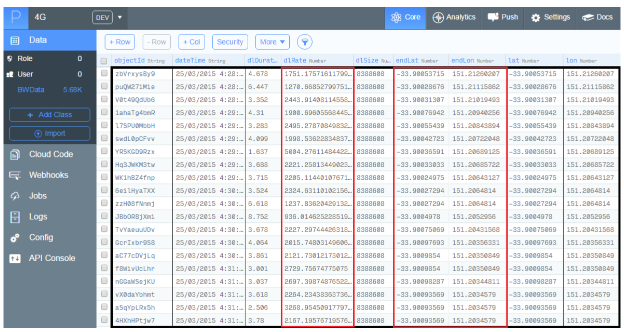


Figure 5: Parse: some recorded data samples with JavaScript based bandwidth measurement application

sideration. The course taken will ultimately affect the performance, storage, backup, monitoring and scalability of the end solution. One option is to host the required database using an SQL based database on our server at UNSW/ CSE. Potential issues involved with such a traditional approach include having to set up the server, monitoring for issues and the server maintenance. It is foreseeable that there will be also performance issues due to CPU share as many other people may host their projects on the same machine. A better option would be using one of the existing cloud based hosting systems. We chose to use Parse [16] as a suitable platform to build our dataset. Parse provides *Android* SDKs and *JavaScript* APIs for developers which considerably reduce the required time to develop a web backend and improve applications' scalability. The mobile client's backend library can be readily connected to the Parse's backend using a predefined private/public key and store or retrieve the data.

4. BANDWIDTH MEASUREMENT CAMPAIGN 2: 3G & 4G

Using our developed Android bandwidth measurement application, we conducted our second bandwidth measurement

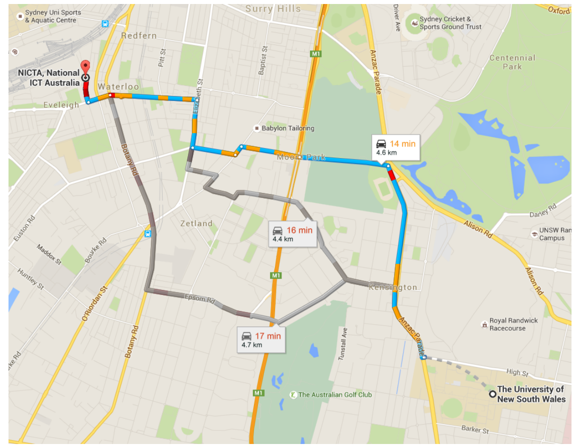


Figure 6: UNSW - NICTA, source: Google map

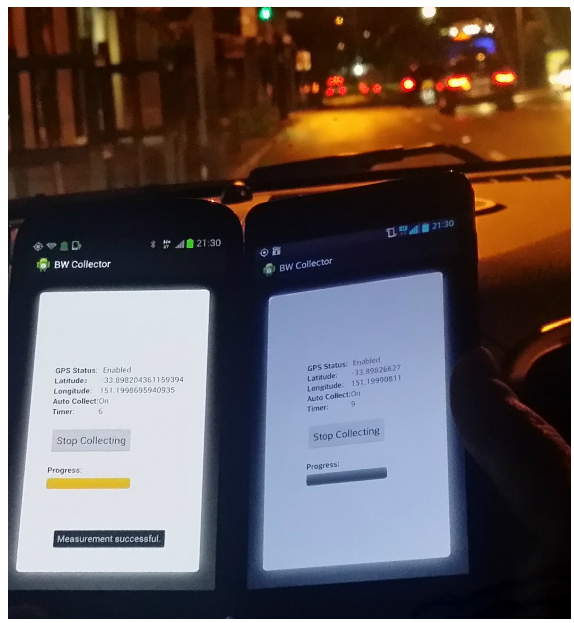
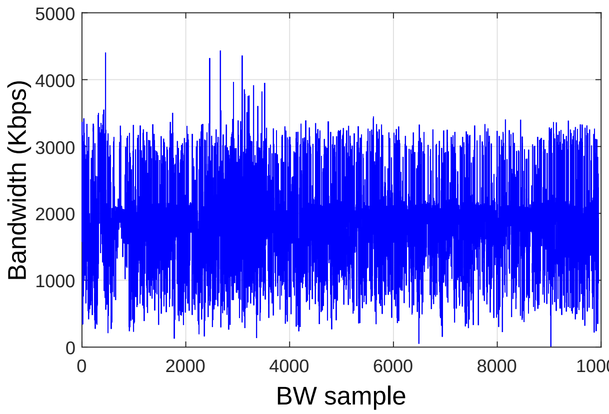
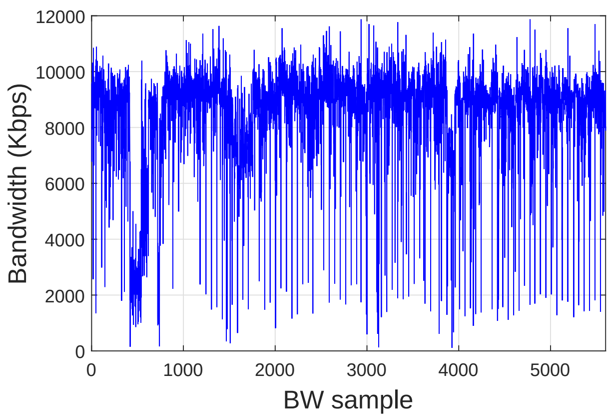


Figure 7: Bandwith collection: UNSW - NICTA, night time

campaign in 2015. This time, we considered a 4.7Km route from UNSW to NICTA in Sydney, Australia (Fig. 6). Two different Android phones were used to collect 3G and 4G bandwidth samples simultaneously. A total of 15,624 samples were collected within 72 trips in different day and night times to consider On and Off peak hours of traffic (see Fig. 7). Some recorded samples on our Parse server are shown in Fig. 5. As it's shown, each measurement is recorded in a separate row with different sampling frequencies of 10 and 15 seconds for 4G and 3G networks respectively. The higher sampling frequency for 4G network is considered to collect more bandwidth samples on this faster network. Fig. 8 demonstrates how bandwidth fluctuates in 3G and 4G networks. We also plot the empirical and normal CDFs of bandwidth samples from both networks in Fig. 9. As it's shown, similar to our previous measurements, the normal CDFs of bandwidth samples could indicate the real bandwidth distribution.

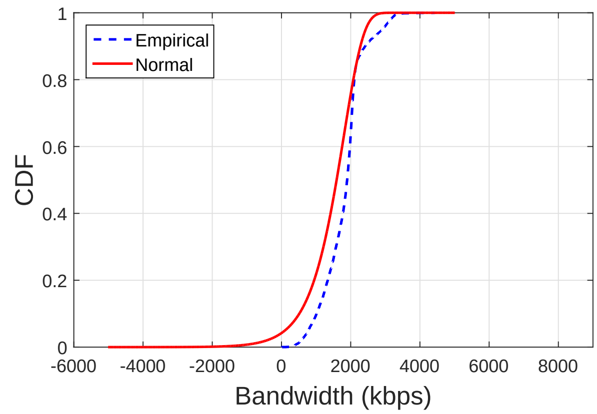


(a)

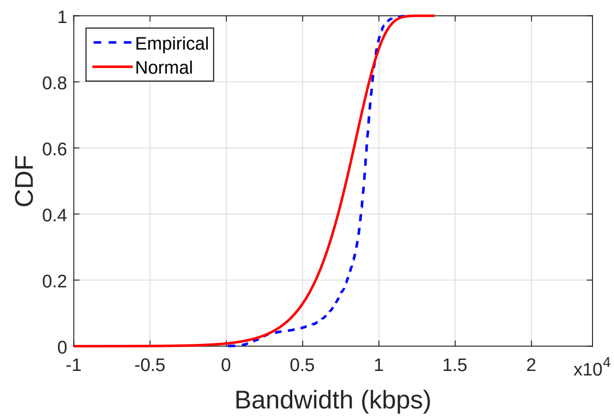


(b)

Figure 8: Bandwidth fluctuations from different real-worls networks: (a) 3G, (b) 4G



(a)



(b)

Figure 9: Empirical Vs Normal CDFs of bandwidth samples collected from different real world networks: (a) 3G with $\mu = 1802.83$ Kbps and $\sigma = 572.77$ Kbps, (b) 4G with $\mu = 8522.11$ Kbps and $\sigma = 1778.86$ Kbps.

5. EXAMPLES OF USING THIS DATASET

In [4] and [3], we used the 3G bandwidth statistics from our first bandwidth measurement campaign to estimate the available bandwidth in different locations. Using simulations we showed how such successful estimation can enhance the quality of video streaming in vehicular environment. Without this dataset, it could not be possible to conduct experiments with our streaming model and measuring its performance for similar network conditions. Such experiments are not feasible over real-world networks due to the bandwidth fluctuations over time and not repeatable network behaviour. In another study [5], where we used multiple learning algorithms, again it was essential to use our bandwidth dataset to repeat the same network condition and evaluate the effect of multiple tuning parameters in our models.

As another example, Yao et al., [23] used bandwidth statistics to predict the available bandwidth in different road segments to provide higher QoS of video streaming for commuters. They found that there is a strong correlation between WWAN throughputs and geographical locations. Their results show that in most of locations, the average of previous bandwidth values is a better indicator of the current bandwidth at that location compared to the observed bandwidth value in previous location. The majority of video streaming methods rely on online bandwidth observation to adapt the streaming quality to the available bandwidth. However, using historical bandwidth information makes it possible to predict the network behaviour with a considerably higher accuracy.

6. CONCLUSIONS

In this paper we discussed our two bandwidth measurement campaigns on 3G and 4G networks in a vehicular settings. A particularly salient feature of our campaign is that we conducted several repeated measurements on selected urban routes. Thus, our data offers interesting insights about the bandwidth fluctuations that can occur in a realistic urban setting. The bandwidth traces will be particularly of interest to researchers wishing to conduct experiments to evaluate the performance of their network protocols under real-world settings. The numerous traces from selected routes will allow researchers to draw statistically significant results. We have made our dataset freely available to all researchers.

7. REFERENCES

[1] Adobe. HTTP Dynamic Streaming on the Adobe Flash Platform. [Online accessed 04-March-2013], URL: http://www.adobe.com/au/products/hds-dynamic-streaming.html.

[2] Apple. HTTP Live Streaming Overview. [Online accessed 01-October-2014], URL: http://goo.gl/CElrwV.

[3] A. Bokani. Location-Based Adaptation for DASH in Vehicular Environment. In *Proceedings of the 2014 CoNEXT on Student Workshop ACM '2014*, Sydney, Australia. 2-5 December 2014.

[4] A. Bokani, M. Hassan, and S. Kanhere. HTTP-based Adaptive Streaming for Mobile Clients using Markov Decision Process. In *Packet Video Workshop (PV)*, San Jose, USA. 12 December 2013.

[5] A. Bokani, M. Hassan, S. Kanhere, and X. Zhu. Optimizing HTTP-Based Adaptive Streaming in Vehicular Environment using Markov Decision Process. *Multimedia, IEEE Transactions on*, 17(12):2297–2309, 2015.

[6] A. Bokani, S. A. Hoseyni, M. Hassan, and S. S. Kanhere. Empirical Evaluation of MDP-based DASH Player. In *International Telecommunication Networks and Applications Conference (ITNAC)*, Sydney, Australia. 18-20 November 2015.

[7] A. Bokani, S. A. Hoseyni, M. Hassan, and S. S. Kanhere. Implementation and Evaluation of Adaptive Video Streaming based on Markov Decision Process. In *IEEE International Conference on Communication (ICC)*, Kuala Lumpur, Malaysia. 23-27 May 2016.

[8] Y.-C. Chen, E. M. Nahum, R. J. Gibbens, D. Towsley, and Y. sup Lim. Characterizing 4G and 3G Networks: Supporting Mobility with Multi-Path TCP. *University of Massachusetts Amherst, Tech. Rep*, 2012.

[9] Cisco. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update 2014âĂŞ2019 White Paper. [Online accessed 04-June-2015], URL: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html.

[10] I. D. Curcio, V. K. M. Vadakital, and M. M. Hannuksela. Geo-Predictive Real-Time Media Delivery in Mobile Environment. In *Proceedings of the 3rd workshop on Mobile video delivery (MoViD)*, Firenze, Italy. 25 October 2010.

[11] P. Deshpande, X. Hou, and S. R. Das. Performance Comparison of 3G and Metro-Scale WiFi for Vehicular Network Access. In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, Melbourne, Australia. 1-3 November 2010.

[12] A. Greenspan, M. Klerer, J. Tomcik, R. Canchi, and J. Wilson. Ieee 802.20: Mobile broadband wireless access for the twenty-first century. *Communications Magazine, IEEE*, 46(7):56–63, 2008.

[13] L. Li, K. Xu, D. Wang, C. Peng, Q. Xiao, and R. Mijumbi. A Measurement Study on TCP Behaviors in HSPA+ Networks on High-Speed Rails. In *IEEE Conference on Computer Communications (INFOCOM)*, Kowloon, Hong Kong. 26 April-1 May 2015.

[14] Ookla. The Global Standard in Interrnet Metric. [Online accessed 01-March-2015], URL: http://www.speedtest.net/.

[15] Optus. Oz Broadband Speed Test. [Online accessed 01-March-2015], URL: http://www.ozspeedtest.com/.

[16] Parse. Perfect Cloud to Power Your App on any Platform. [Online accessed 01-March-2015], URL: https://www.parse.com.

[17] M. Y. Rhee. High speed downlink packet access (hsdpa). *Mobile Communication Systems and Security*, pages 167–196.

[18] H. Riiser, T. Endestad, P. Vigmostad, C. Griwodz, and P. Halvorsen. Video Streaming Using a Location-Based Bandwidth-Lookup Service for Bitrate Planning. *ACM Transactions on Multimedia Computing, Communications and Applications*, 8(3):24, 2012.

[19] H. Riiser, P. Vigmostad, C. Griwodz, and P. Halvorsen. Commute Path Bandwidth Traces from 3G Networks: Analysis and Applications. In *Proceedings of the 4th ACM Multimedia Systems Conference (MMSys '13)*, Oslo, Norway. 1 March 2013.

[20] V. Singh, J. Ott, and I. D. Curcio. Predictive Buffering for Streaming Video in 3G Networks. In *IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, San Francisco, USA. 25-28 June 2012.

[21] Q. Xiao, K. Xu, D. Wang, L. Li, and Y. Zhong. TCP Performance over Mobile Networks in High-Speed Mobility Scenarios. In *22nd International IEEE Conference on Network Protocols (ICNP)*, Raleigh, USA. 21-24 Oct 2014.

[22] J. Yao, S. S. Kanhere, and M. Hassan. An Empirical Study of Bandwidth Predictability in Mobile Computing. In *Proceedings of the third ACM international workshop on Wireless network testbeds, experimental evaluation and characterization (MobiCom-WiNTECH)*, San Francisco, USA. 14-19 September 2008.

[23] J. Yao, S. S. Kanhere, and M. Hassan. Improving QoS in High-Speed Mobility Using Bandwidth Maps. *IEEE Transactions on Mobile Computing*, 11(4):603–617, 2012.

[24] A. Zambelli. IIS smooth streaming technical overview. *Microsoft Corporation*, 3, 2009.

[25] G. Zhong and A. Bokani. A Geo-Adaptive JavaScript DASH Player. In *Proceedings of the VideoNEXT '2014 Workshop on Design, Quality and Deployment of Adaptive Video Streaming ACM '2014*, Sydney, Australia. 2-5 December 2014.