

Mobile Network Traffic: a User Behaviour Model

Ilias Tsompanidis

Computer Science Dept.
University College Cork, Ireland
Email: it1@cs.ucc.ie

Ahmed H. Zahran

Electronics and Electrical Communications Dept.
Cairo University, Egypt
Email: azahran@eece.cu.edu.eg

Cormac J. Sreenan

Computer Science Dept.
University College Cork, Ireland
Email: cjs@cs.ucc.ie

Abstract—What would the same users do in a different network? The performance of the network has a significant effect on the traffic profiles of users, which cannot be easily identified from network traces. This work combines a number of studies to compile a new responsive traffic model for mobile networks that realistically mimics user behaviour. Users continuously evaluate the performance of the network, and initiate or terminate their sessions accordingly. The presented model utilises Markov chains to capture this behaviour, while the performance of a session is depicted with the binary distinction of good and bad quality. We analyse and showcase the implications of this model, and discuss its benefits on network planning and research applications.

I. INTRODUCTION

The landscape in mobile data traffic is changing rapidly, as reported by Cisco in their yearly Index [1]. Past studies and traffic models for mobile devices have focused on network traces, usually from laptops connected to campus networks. Today, mobile devices include highly portable smartphones and tablets, with better battery and connectivity capabilities. Networks have evolved from isolated hotspots to highly integrated heterogeneous wireless networks. Traffic patterns have changed, with peer-to-peer services no longer being so popular and emerging legal HTTP-based media streaming services gaining much popularity [2]. Music streaming (Pandora, Spotify), movie streaming (Netflix, Hulu, iPlayer), and casual video streaming (Youtube, Vimeo) are just examples of services that were either nonexistent, or just emerging in mobile a few years ago.

Traces provide only part of the traffic story, constrained to what the clients were able to consume while limited by the network capacity, their position and movement at the time of measurement. Unfortunately, it is very difficult to record the intentions of the users, and accurately model the extra traffic that would be transmitted if conditions were different. Information on session drop rates after unsuccessful transmissions and long buffering times can help us understand the potential of the clients. A few studies, mainly addressing mobile media streaming, *e.g.* [3], shed light on the behaviour and tolerance of mobile clients. Some of their important findings are the relation between startup delay/rebuffering and abandonment rate, as well as the effect streaming problems have on starting another session.

In this work we argue that models blindly mimicking traces are not realistic for traffic generation, especially for protocols that optimise usage of network resources. A traffic generator

should typically react to network conditions and performance, as real users do. However, trace-driven models are based on static views of user requests. We present a novel user behaviour-based traffic model.

The remainder of this work is organised as follows. Section II presents the behaviour-driven traffic model. We proceed with discussing the requirements of some popular types of traffic and their supporting models in Section III. In Section IV we showcase the performance of our model using the supporting traffic types. We identify the place and relevance of this work within other traffic models in Section V and discuss its implications in Section VI. Finally, we conclude in Section VII.

II. MODEL DESCRIPTION

We use the term *Session* to describe all traffic and flows that are needed to fulfill a single task, such as stream a movie, browse the web, sync e-mails, or download a file. This abstract view allows us to define Sessions with Average Bit Rate requirements, while more detailed views can be developed if needed. We introduce the term *Activity* as a group of consequent Sessions of the same type. A user begins an Activity by engaging with a mobile device, with the intent to carry out a Session. Each Session may cause the desire to continue the Activity with a new Session.

Figure 1 shows a high-level view of the proposed model. A user is trying to perform an Activity, such as browse the web or stream videos. For example, an Activity would be a user watching videos on Youtube, with each streamed video being a Session. The user may select a new one to continue his Activity. He might even stop watching a video before it is finished and try a different one (abandon the Session), or even stop watching altogether (abandon the Activity).

The probabilities to start a new Session and to continue the active Session (P_N and P_C , respectively) can change during the Session. For example, if a new Session fails to start transferring data for more than 15 seconds, P_C could be set to zero and the user would abandon the stagnant Session. We formally define these changes in user behaviour with two thresholds, the Quality threshold (T_Q) and the Abandon threshold (T_A). When a Session is performing below T_Q , the user is experiencing nuisances that may cause him to abandon the Session, *e.g.* slow loading time or rebuffer events in video streaming. A Session performing below T_A leads the user to immediately abandon the Session.

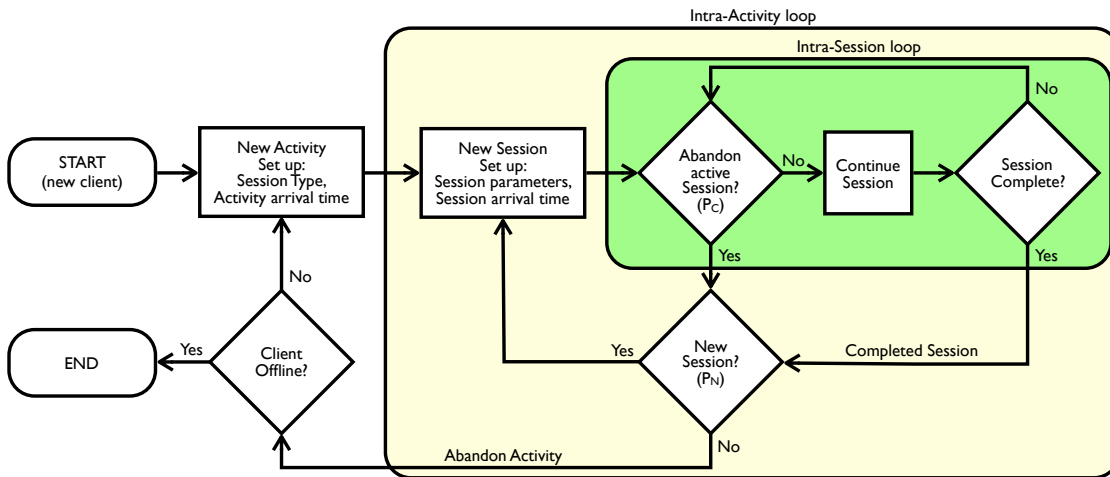


Fig. 1. Traffic model flowchart.

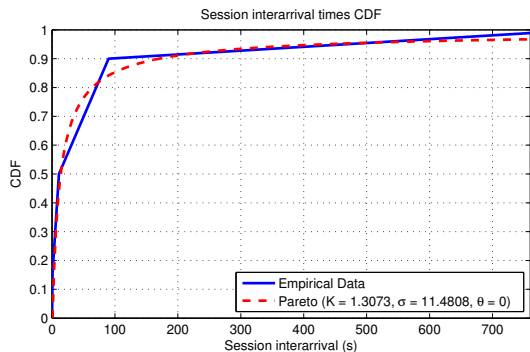


Fig. 2. Empirical [4] and Pareto distributions of inter-Session time.

A. Activity and Session arrivals

This section describes the arrival of data Sessions of a cellular client, allowing them to form Activities. The inter-Activity and inter-Session times follow distinct distributions. New Sessions are generally initiated within a matter of seconds, while a user is actively using a device. Afterwards, the device would be typically left aside for some minutes, before it is used again for a new Activity.

We proceed with suggestions for these distributions, as identified from currently available data. Session arrivals have been thoroughly covered in numerous studies. For example, in [4], the authors studied a number of different Session types and report on the interarrival times. 50% of Sessions occur in less than 11 seconds, and 40% in between 11 seconds and 1.5 minutes. Based on this information, we fit the Pareto distribution using Maximum Likelihood Estimation, as shown in Figure 2.

An Activity is initiated when a user begins engaging a mobile device. Only Activities that produce network traffic are of interest for this work. Internet analytics data [5] claim that on average, people check their phones 150 times a day, for reasons varying from checking the time to streaming movies.

The presented usage breakdown [5] suggests that about half of these interactions generate network traffic. We consider a Poisson Burst Process (PBP) to model the arrival of Activities, with an arrival rate of $\lambda = 0.05$ per minute (≈ 70 per day). Thus, the time between two consecutive Activities can be modelled with an exponential distribution, with $\mu = 1/\lambda \approx 1200$ seconds. The Sessions belonging to the same Activity constitute the bursts of the PBP. Transient and diurnal behaviour, such as reduced device usage during the night, can be expressed by means of a time-varying PBP.

The Activity arrival rate in this work is related to the Probability of Return (PoR) after a *visit* in [3]. More specifically, the authors studied the PoR of failed and normal visits, with a visit consisting of views with a low interarrival time. The PoR for the immediate future is significantly affected by the success of the last Session. In fact, the PoR after a successful Session is about double the PoR after a failed Session. This corroborates our insight of considering Sessions in the same Activity as a group of related tasks with short return time, and that the Activity interarrival depends on the success of the previous Sessions. We extend the Activity arrival rate of the PBP to $\lambda_S = 0.05$ and $\lambda_F = 0.025$ per minute, following an Activity with a successful and a failed last Session, respectively.

A Session is considered as failed when the quality of experience is bad, as the user understands that it is under-performing. A failed Session may be dropped by the network, or the user may decide to abandon it, or even to complete it. However, a good quality Session is considered as successful even when the user stops it mid-Session by his own decision. Therefore, not all abandoned Sessions are failed, and not all successful Sessions are fully completed.

B. Intra-Activity dynamics

The intra-Activity loop controls the creation of new Sessions within the same Activity. Since there are two states of finished Sessions, P_N is better described with two probabilities, P_{NS} and P_{NF} , the probabilities to start a new Session after a successful and a failed one, respectively.

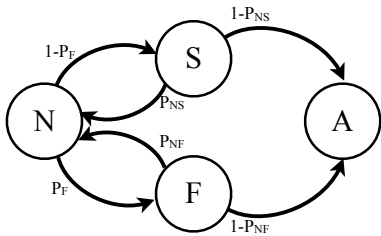


Fig. 3. Intra-Activity Markov model. State N stands for an active Session that can end successfully (S) or fail (F). From both states, there can either start a new Session (N) or abandon the Activity (A).

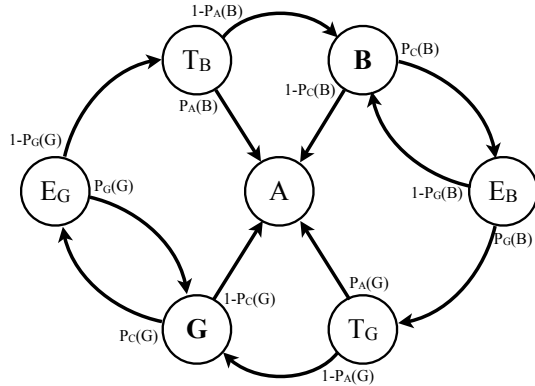


Fig. 4. Intra-Session Markov model. States B and G stand for Sessions in Bad and Good quality and any of them can be the starting state. States E stand for the evaluation of the Session quality by the user. States T_B and T_G are used as transitional states to Bad and Good quality Sessions, respectively. The user abandoning the Session is depicted by state A.

A four-state Markov model is employed to depict this behaviour, as shown in Figure 3. P_F is the probability for a Session to fail, which is affected by the network performance. Finally, if no new Session is created, the Activity is abandoned.

C. Intra-Session dynamics

Figure 1 describes the Intra-Session loop as a process where a user continuously evaluates the performance of the Session and decides to continue or abandon. We identify three possible cases for a user to drop a Session: a) there is a sudden severe quality drop, and the user decides to immediately abandon the Session, b) the user is fed up and abandons the Session after enduring bad quality for a while, and c) the user loses interest and abandons the Session regardless the quality. The two first cases are caused by Sessions with bad quality, whereas the third one is valid for both good and bad quality. Modeling this behaviour is not trivial. Figure 4 shows the discrete Markov model that describes this process.

The three behaviours and the Session performance are modeled with the different transition probabilities. P_C is the probability for a user to continue a good (G) or a bad (B) quality Session. P_G is the probability for a Session to be considered as of good quality during evaluation. P_A is the probability of a user immediately abandoning a Session upon a quality change. We expect $P_A(G) = 0$, as a transition to good quality is generally considered as an improvement and

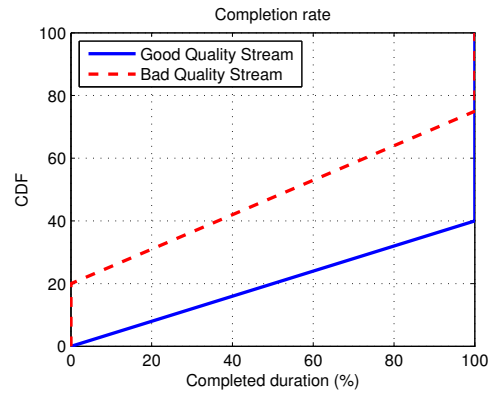


Fig. 5. CDF of modeled Time-to-Abandon (TTA).

does not cause irritation to users.

D. Session dropping

Users often do not finish their Sessions. For example, approximately 50% of video streams are dropped before being completed [6]. More interestingly, a simple metric is devised to identify the good and bad quality, the stream's *normalized rate*. 60% of good quality streams are completed, in contrast to bad quality streams where only 25% are fully watched and 20% are immediately dropped. We base the completion rate CDF in Figure 5 on these findings.

The users continuously evaluate the connection quality and decide whether to continue their Session in the process we described as the Intra-Session loop in Section II-C. The effects of this continuous evaluation are shown in Figure 5, providing a straightforward model of expected playback duration, the *Time-to-Abandon* (TTA). We expect other applications to have a similar TTA model, as users can stop a Session frustrated due to poor quality or lose interest even with good quality. With this convention, we fit the model probabilities to Figure 5, from the data available from [6]. Thus, $P_A(B) = 0.2$, while $P_C(B)$ and $P_C(G)$ depend on the evaluation frequency and duration of the Session. A Session either completes its duration while in states G or B, or the user abandons it in state A.

Since networks and services usually cannot initiate a Session instantaneously, users tolerate a brief inactivity time during the start-up of a Session. A start-up delay is considered as normal if its duration is within the *grace period*. However, during the *waiting period*, users may opt to abandon the Session. Drawing on [3], we use a grace period of 2 seconds, while the time that a user will wait for a new Session to reach at least T_A , before abandoning it, follows a uniform distribution of [2, 19.2] seconds.

III. SESSION TYPES

Various Session types can be introduced to describe virtually any kind of traffic. Since the quality of a Session is assessed and fed back to the behaviour model, the traffic models have to support some quality parameters. For the scope of this paper, the quality parameters of the Session are covered by average throughput requirements.

The first characteristic of a Session is its duration (d) in seconds or its length (l) in bytes. A Session also has a Maximum (D_{max}) average demand, and two throughput thresholds, T_Q as a binary limit of good and bad quality and T_A for lowest acceptable quality. T_Q resembles the minimum average throughput for good user experience, T_A the point where a user will abandon the Session and D_{max} the maximum sustainable throughput of this Session. Finally, the probabilities to start a new Session (P_{NS} and P_{NF}) are also required.

In the following subsections we identify the required Session parameters for a selection of popular traffic types: media streaming over HTTP and web traffic. In particular, Video on Demand streaming will either have a satisfactory throughput level that enables seamless playback for a given video encoding rate, or keep causing rebuffering events due to low throughput. Similar tolerance thresholds for web traffic are described by ITU-T, by the means of preferred and acceptable transfer time for browsing and file transfer [7], [8].

A. Web traffic

Web browsing typically features a number of short flows followed by significant network inactivity intervals during which the user views and processes the information. A traffic Session with completely different behaviour than web browsing would be a file download, such as an email with a sizeable attachment, or a file transfer.

1) *Interactive web traffic*: Web browsing Sessions usually include the transfer of data and a client dwell time, while the user physically reads through the content. It is also common for users to continue the Sessions, visiting multiple web pages. This behaviour is consistent with the self-similar nature of web browsing traffic, as it has been described as an ON-OFF model in the literature, *e.g.* [9]. The average throughput (T) of such a sequential browsing Session would be:

$$T_{avg} = \frac{bits}{t_{transfer} + t_{dwell}} = \frac{bits}{\frac{bits}{T_{transfer}} + t_{dwell}} \quad (1)$$

where t_x is the time it takes for action x to complete.

Dwell time is expected to vary depending on the amount of content and the attention span of the user. Indeed, researchers have showed that dwell time follows a Weibull distribution [10], with the scale parameter λ highly dependent on content amount and size. In over 80% of their data, dwell times are limited between 2 and 70 seconds.

Many researchers have reported on the size of websites, best modelled by heavy-tail distributions. In 2011 [11], the average size for mobile-optimised webpages was 160KB and 770KB for non-optimised. Since mobile web browsing includes both optimised and non-optimised web pages, it is not unreasonable to consider web page sizes in the range of 60KB and 1500KB.

Finally, acceptable load times are also of interest. Web pages are expected to load in 15 seconds [8]. Load times longer than 10 seconds make the user believe the web page is unresponsive, leading to the user abandoning the Session.

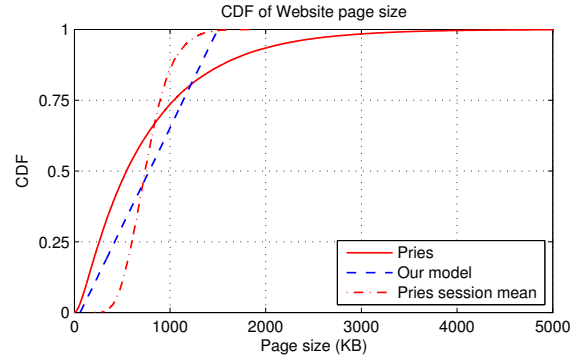


Fig. 6. Web page size under our model and Pries [9]. Pries Session mean considers the average web page size for Sessions visiting 10 web pages.

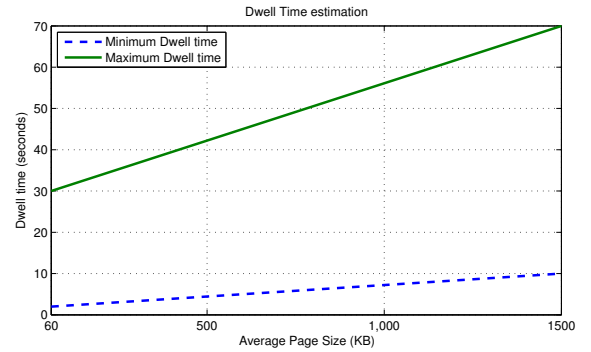


Fig. 7. Dwell time for interactive web traffic.

We combine the information above to describe a simple model for interactive web traffic. The average webpage size is defined by a uniform distribution between 60 and 1500 KB. This simple approach provides comparable average page sizes for web browsing Sessions to current models, as shown in Figure 6. For a given average webpage size, the following parameters are defined: dwell time, D_{max} , T_A and T_Q . Dwell time follows a uniform distribution between the limits shown in Figure 7. The thresholds for load time are shown in Figure 8, and by applying Eq 1 D_{max} , T_Q , and T_A are computed.

Web browsing Session size is traditionally defined in bytes, as in how many bytes long the visited web pages and embedded objects are. However, for a Session of consecutive visits to web pages with an associated dwell time, a more appropriate Session size metric is time, as a measure of user engagement. The duration of such Sessions in WAP browsing was measured in [12], and was found to follow a Weibull distribution.

It should be noted that while more sophisticated models exist, to the best of our knowledge there are no models which combine dwell and acceptable load times that depend on webpage size. In this section, we presented a simple approach that includes these factors, important for setting QoE targets, while conforming to the average web site size and dwell time reported in the bibliography. However, a more detailed web browsing model is not within the scope of this work.

2) *File download*: In mobile devices, it is common for the user to actively wait for a download to finish. Long downloads

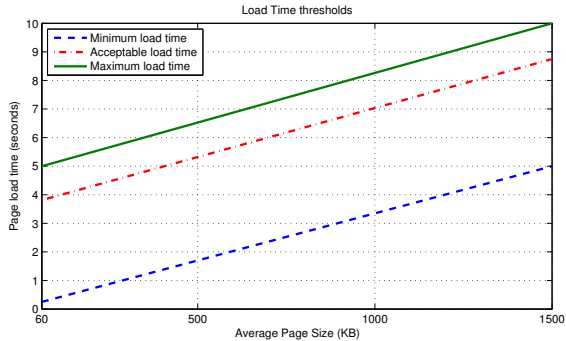


Fig. 8. Load time thresholds for interactive web traffic.

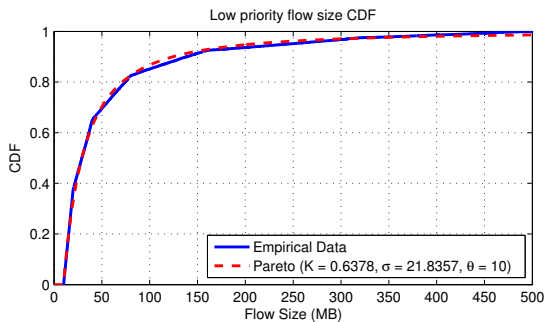


Fig. 9. Empirical [14] and Pareto distributions of file size.

typically transfer content such as music, or smartphone apps. T_A and T_Q should describe the tolerances of users to slow and very slow file transfers. As WSP invest in better infrastructure and more capacity, users grow accustomed to better download speeds. Thus, T_A and T_Q for file downloads depend on user expectations, rather than content-specific requirements.

Various parties perform hands-on user-side measurements of cellular network speeds (e.g. [13]). These real-life measurements can offer significant insight for the user-perceived performance of mobile networks. From the data offered by these studies, we assume the lowest average download rate in 3G networks sets $T_A \approx 0.66$ mbit/s (5 MB per minute), the average download rate across carriers for 3G sets $T_Q \approx 1.33$ mbit/s (10 MB per minute), and the average download rate across carriers for 4G sets $D_{max} \approx 10$ mbit/s.

A research with data from 2007 from a number of campuses identified the flow size distribution of low priority web traffic [14]. Even though they considered mainly wired traffic, we believe that the observations made for wired data from 2007 can cautiously apply to mobile traffic today. Therefore, we use the Pareto distribution fitted to the empirical data, as shown in Figure 9.

B. Media streaming

Media streaming is rapidly becoming the prevalent traffic in mobile networks. The de facto method is HTTP streaming, while two main strategies have been identified, Progressive Download (PD) and Dynamic adaptive streaming over HTTP (DASH) [15]. PD uses minimal signalling and is transferred

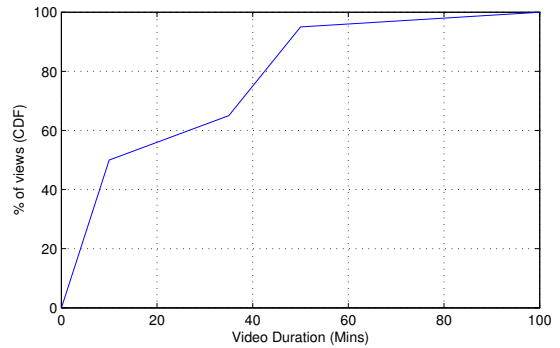


Fig. 10. CDF of modeled video request duration.

in chunks similarly to a file download, while DASH offers adaptive video quality.

Media streaming over HTTP/TCP has a significant throughput overhead. The analytic model presented by Wang *et al.* [16] provides comprehensive insight on the overheads incurred by TCP. One of the most important findings is that the throughput required for a continuous playback with no rebuffering is about double the encoding rate of the video.

Krishnan *et al.* [3] studied the video stream viewing behaviour over a large number of Sessions, for wired and mobile clients. They mainly focus on the abandonment and revisiting rates, providing useful insights about the tolerance of users. Additionally, the paper details the split of video requests into two logical groups, the videos with a duration of under 30 minutes, and the ones longer than that. We identify two different logical groups, the short ‘casual’ ones with a duration shorter than 10 minutes, and the longer ‘movie’ group (> 10 minutes). About 50% of the requests target casual videos with a duration less than 5 minutes. As for movie videos, about 15% of the requests regard videos 20-22 minutes long, about 30% target 35-45 minutes and a small 5% of requests are for videos longer than 45 minutes. We base our duration distribution for requested videos on the same trends, shown in Figure 10, and provide a simple but realistic approximation of the empirical distribution.

1) *Casual streaming ($d \leq 10$ mins)*: As the videos are short, we assume that PD is used. The video resolution is chosen randomly from 240p - 720p, defining its encoding rate (R_{enc}). We use as a reference the average bitrates for various video resolutions as served by Youtube, shown in Table I.

Rebuffering has a significant effect on media streaming QoE. Even a few rebuffering events lead to shorter playback times, with a higher probability of dropping the Session [17]. For seamless playback, we consider $T_Q = 2 * R_{enc}$, as previously mentioned. A conservative estimation of the lowest acceptable quality would be $T_A = R_{enc}$.

Since PD is essentially a file download, any $D_{max} \geq T_Q$ suffices. Some PD strategies initially download a small portion fast and then just stream at R_{enc} , while others just try to download the file as fast as possible. We consider both behaviours by selecting D_{max} randomly from $[2 * R_{enc}, 4 * R_{enc}]$.

TABLE I
YOUTUBE AVERAGE BITRATE PER RESOLUTION

Resolution	kbit/s
240p	310
360p	560
480p	1010
720p	2100
1080p	3750

2) *Movie streaming ($d > 10$ mins)*: Movie streaming usually employs adaptive video techniques, so R_{enc} is considered variable. We assume 360p to be the lowest acceptable quality and 720p to be the maximum streaming quality. We set T_A , T_Q and D_{max} accordingly from Table I, to $R_{enc}(360p)$, $2 * R_{enc}(360p)$ and $2 * R_{enc}(720p)$ respectively.

C. Fitting Sessions to the behaviour model

Lacking more detailed data, we formulate some assumptions for the P_N of the various Session types. For movie streaming and file downloads a user would typically initiate an Activity to watch a specific movie or download a specific big file. We typically expect an average of 1.5 consecutive Sessions in the Activity, when successful. Casual streaming and web browsing are types of Activities that returning Sessions are more easily initiated. We assume an average of 2.5 consecutive Sessions for successful Activities of these types.

For Activities with only successful Sessions ($P_F = 0$), the model in Figure 3 is easily reduced to a simple 2-state Markov chain. Thus, for movie streaming and file downloads P_{NS} is 33.3%, and 60% for casual streaming and web browsing. In case of failed Sessions of any type, we consider $P_{NF} = 20\%$, meaning that Activities with $P_F = 1$ have an average of 1.25 consecutive Sessions.

The rest of the parameters used are particular to each Session type. A summary of all proposed model parameters is shown in Table II. We have detailed the process of identifying these parameters for throughput-based quality assessment and provided suggested values based on current trends. Nevertheless, we encourage adjustments or even new definitions of Session types for a more appropriate application of our behaviour model.

IV. ANALYSIS

The analysis of each model follows, with the assumption that the quality of each Session is set from the beginning and does not change in its duration. We calculate the expected traffic from each Session type, and show the difference in produced traffic for various quality levels, as a means to demonstrate the user behaviour under different conditions.

A. Interactive Web

From its Weibull distribution the Session duration has a mean value of 265 seconds. From the TTA model in Figure 5, Good Quality Sessions have an average duration of $(0.6 * 100\% + 0.4 * 50\%) * 265 = 0.8 * 265 = 212$ seconds. The duration for Bad Quality Sessions is $(0.2 * 0\% + 0.55 * 50\% + 0.25 * 100\%) * 265 = 0.525 * 265 = 139$ seconds.

The mean of the uniform distribution of page size is 780 KB. For this page size, the dwell time follows a uniform distribution within $[6, 50]$ seconds, with a mean of 28 seconds. Both the average page size and dwell time are comparable with the relevant averages from the HTTP model in [9]. For Good Quality Sessions the load time would be within $[2.625, 6.2812]$, with an average of 4.45 seconds. Similarly, for Bad Quality Sessions load times would be within $[6.2812, 7.5]$, averaging to 6.89 seconds.

From Eq. 1, the average throughput is 192.3 and 178.8 kbit/s for good and bad quality Sessions respectively. Thus, a Session of good quality is expected to produce 4.97 MB of traffic, while a bad quality one 3.03 MB.

Finally, an Activity with only bad quality Sessions will transfer 3.79 MB, with one with only good quality Sessions will reach 12.44 MB.

B. File Download

The mean file size from the Pareto distribution is about 60 MB. Good quality Sessions have an average completion rate of 80%, thus the expected transferred size is 48 MB. Similarly, for Bad quality streams it is 31.5 MB. Thus, the Activity will transfer on average 39.375 and 72 MB for only bad and only good quality streams, respectively.

C. Casual Streaming

From the uniform distribution of R_{enc} and stream duration, the average streaming bitrate is 1990 kbit/s and duration 300 sec, thus an average transfer size of 72.876 MB per Session.

After applying the known completion rates, good and bad quality streams transfer on average 58.3 and 38.26 MB, respectively. Bad quality Activities average 47.825 MB, while good quality Activities reach an average of 145.75 MB.

D. Movie streaming

The adaptive nature of movie streaming can mask minor changes in network performance. By definition bad quality streams will have a throughput of $[560, 1120]$ kbit/s, while good quality streams $[1120, 4200]$ kbit/s. As the actual bitrate distribution is unknown, we assume the average of the two ranges to be representative of the two quality levels, at 840 and 2660 kbit/s respectively.

The average video length per Session is 2385 seconds, and the estimated completed duration is $0.8 * 2385 = 1908$ seconds for good quality streams, and $0.525 * 2385 = 1252$ seconds for bad quality streams. Therefore, for good quality streams, the average total data traffic per Session is 619.54 MB, adding up to an average of 929.31 MB per Activity. Similarly, Activities with only bad quality streams produce an estimated traffic of 160.5 MB, with each Session contributing 128.4 MB on average.

E. Overall

We have identified the average Activity size of all traffic types for $P_F = 0$ and $P_F = 1$. The average traffic of single-type clients for various values of P_F is presented in Figure 11.

TABLE II
MODEL PARAMETERS

		Interactive Web	File Download	Casual streaming	Movie Streaming
Activity Arrival	Model	Poisson Process (min)			
	Assumptions	Following successful Session: $\lambda_S = 0.05$, following failed Session: $\lambda_F = 0.025$			
Session Interarrival	Model	Pareto (sec)			
	Assumptions	From empirical data (Fig. 2)			
Session Size/Duration	Model	Weibull (sec)	Pareto (MB)	Piecewise Uniform (min)	
	Assumptions	Shape: 0.4, Scale: 80	Fig. 9	Empirical data (Fig. 10)	
Max Throughput D_{max}	Metric	Max avg. throughput	Max download rate	Max download rate	Max Adaptive Streaming rate
	Assumptions	From avg. website size and dwell time (Fig. 8)	10 mbit/s current 4G WSP	$[2 * R_{enc}, 4 * R_{enc}]$	$2 * R_{enc}(720p)$
Quality Threshold T_Q	Metric	Acceptable avg. load time	Acceptable transfer rate	Rebuffer-free playback	Acceptable Streaming quality
	Assumptions	Linear to avg. page size Fig. 8	1.33 mbit/s 10 MB/minute	$2 * R_{enc}$	$2 * R_{enc}(360p)$
Abandon Threshold T_A	Metric	Maximum avg. load time	Minimum transfer rate	Low-quality playback	Low-quality Streaming
	Assumptions	Linear to avg. page size Fig. 8	0.66 mbit/s 5 MB/minute	R_{enc}	$R_{enc}(360p)$
Start-up delay tolerance SDT	Model	Constant grace period with Uniform waiting period			
	Assumptions	[2,19.2] sec			
Time-to-abandon TTA	Model	Piecewise Uniform			
	Assumptions	$T > T_A$: TTA on remaining Session duration/size (Fig. 5) $T < T_A$: Abandon immediately			
New Session in Activity P_N	Model	Markov chain (Fig. 3)			
	Assumptions	$P_{NS} = 60\%$	$P_{NS} = 33.3\%$	$P_{NS} = 60\%$	$P_{NS} = 33.3\%$
		$P_{NF} = 20\%$			

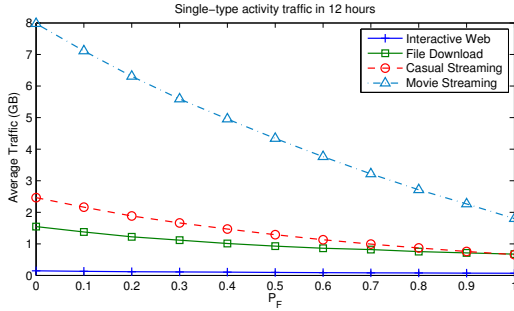


Fig. 11. GB of traffic over 12 hours of single-type Activities in relation to different Session failure rates (P_F)

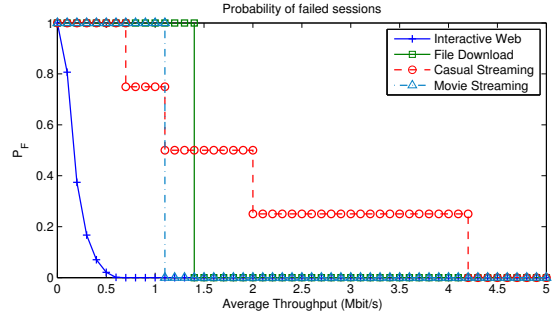


Fig. 12. Probability of failed Sessions (P_F) as a function of throughput.

The difference in total traffic volume between movie streaming and the other types of traffic is significant, even though movie streaming with its longer duration produces less Sessions.

It becomes obvious that the increase of the capacity of an under-performing network results in a significant increase of traffic and a change in the traffic type mix. Moreover, for a given throughput rate each traffic type will have a different P_F , as displayed in Figure 12. For example, Web browsing is the least demanding traffic type, with 600 Kbits/s ensuring good quality Sessions under all circumstances. For the rest of the traffic types T_Q typically resides within the range of 1 to 2 Mbits/s. The sharp changes in P_F are caused by the binary nature of T_Q , also described in [6].

Depending on the network, the relation between P_F and average throughput can differ. Unless some kind of prioritisation is performed, a general rule would be that either web browsing has $P_F \simeq 0$, or all other traffic types have $P_F \simeq 1$. The implication of this statement makes sense in the real world. For example, when web browsing is not performing satisfactorily,

media streaming is certainly expected to under-perform.

V. RELATED WORK

Realistic data traffic is required for many aspects of network simulations, with models and traffic generators constantly being revised by the research community. The two most popular approaches are trace-driven and source-level models. Trace-based models have a significant advantage, as they utilise real traffic that actual users produced. Repositories of trace sets such as CRAWDDAD [18], have allowed many research teams to provide models [19] and tools [20] that can be used for synthetic traffic generation in simulation or real testbed environments.

However, simple trace-based models can have a significant drawback, as they describe networks and users under specific conditions and restrictions. Source-level models attempt to mimic the mechanisms behind specific protocols and applications, and alleviate the issues of blindly using trace-based models [21]. Two popular topics that are approached in this manner are Web traffic and File downloads. File downloads

are usually a testing ground for TCP models, capturing the packet-level dynamics of the protocol, *e.g.* [22]. Web traffic is commonly described with an ON-OFF model (*e.g.* [9], with the ON periods modelling the HTTP GET requests, the responses, and the requests for the embedded objects, while the OFF periods are used to model the processing and client dwell time.

Similar to application source-level models, we present a source-level model of user behaviour, tuned to persons operating mobile devices. Whereas the former describe the flow dynamics of specific applications, our model treats the user as the source of network traffic and describes how users respond to changes in Session quality. Our work extends and combines the realistic creation of bursts of Sessions [23] with the understanding that users behave differently when provided with different amounts of bandwidth [24] in a cross-layer closed-loop model.

VI. DISCUSSION

The effect of performance feedback on the traffic that is produced raises concern over the validity of trace-driven studies and a range of simulations, such as network expansions, capacity planning and resource management. The main claim of this work is that network performance has a significant impact on traffic generation. Trace-driven models that are not responsive to performance cannot sufficiently capture the user-behavioural aspects of traffic generation.

This work builds upon the findings of a number of studies, and is mainly based on recent, mobile-focused measurements. However, the rapid evolution of mobile networks and offered services has not allowed for adequate research of user QoE and behaviour. Thus, in this work, a number of assumptions are made. In order to ascertain validity to our findings, we have been conservative with our parameter and traffic requirement selection. Nevertheless, the basic trends that form the basis of the overall traffic model remain qualitatively relevant, even if some of the contributing models are changed.

VII. CONCLUSION

The user can be considered as the source of the majority of network traffic today, especially for mobile devices. At the same time the user is the entity that evaluates the quality of experience of said traffic. Existing traffic models do not fully capture the mutual interaction between network performance and user generated traffic. This paper presents a novel traffic model that captures such interaction and demonstrates its importance by examining two popular types of traffic: Web traffic and media streaming.

We plan to acquire detailed network traces that capture the user behaviour and conduct extensive user studies to validate and extend the model. Such data will allow better estimation of the various model parameters, such as the Activity and Session interarrival times. A possible extension is a more sophisticated Activity arrival model, as we expect to encounter users that adapt to the network performance by reducing the number of Activities with high throughput requirements, while continuing to use their devices for low throughput Activities.

ACKNOWLEDGMENT

This work is supported by Science Foundation Ireland (SFI) under Research Grant 10RFP/CMS2952. The authors would also like to acknowledge the support of the National Telecommunication Regulation Authority (NTRA) of Egypt.

REFERENCES

- [1] Cisco White Paper, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2012-2017," Feb. 2013.
- [2] S. Gebert, R. Pries, D. Schlosser, and K. Heck, "Internet access traffic measurement and analysis," in *Traffic Monitoring and Analysis*, ser. Lecture Notes in Computer Science, 2012, vol. 7189, pp. 29–42.
- [3] S. S. Krishnan and R. K. Sitaraman, "Video stream quality impacts viewer behavior: inferring causality using quasi-experimental designs," in *ACM IMC*, Nov. 2012.
- [4] R. Kumar and A. Tomkins, "A characterization of online browsing behavior," in *WWW*, Apr. 2010.
- [5] New Relic. (2013, Mar.) Mobile APPEal: Exploring the Mobile Landscape. [Online]. Available: <http://blog.newrelic.com/2013/03/13/mobile-appeal-why-the-future-is-mobile/>
- [6] L. Plissonneau and E. Biersack, "A longitudinal view of HTTP video streaming performance," in *ACM MMSys*, Feb. 2012.
- [7] ITU-T, "Recommendation G.1010. End-user multimedia QoS categories," Nov. 2011.
- [8] —, "Recommendation G.1030. Estimating end-to-end performance in IP networks for data applications," Nov. 2005.
- [9] R. Pries, Z. Magyari, and P. Tran-Gia, "An HTTP web traffic model based on the top one million visited web pages," in *NGI*, June 2012.
- [10] C. Liu, R. W. White, and S. Dumais, "Understanding web browsing behaviors through Weibull analysis of dwell time," in *ACM SIGIR*, July 2010.
- [11] Z. Wang, F. X. Lin, L. Zhong, and M. Chishtie, "Why are web browsers slow on smartphones?" in *HotMobile*, Mar. 2011.
- [12] G. Zhao, Q. Shan, S. Xiao, and C. Xu, "Modeling web browsing on mobile internet," *Communications Letters, IEEE*, vol. 15, no. 10, pp. 1081–1083, 2011.
- [13] Mark Sullivan. (2012, May) 3G/4G Performance Map: Data Speeds for AT&T, Sprint, T-Mobile, and Verizon. [Online]. Available: <http://goo.gl/etXa31>
- [14] M. Marcon, M. Dischinger, K. Gummadi, and A. Vahdat, "The local and global effects of traffic shaping in the internet," in *COMSNETS*, Jan. 2011.
- [15] J. Erman, A. Gerber, K. K. Ramakrishnan, S. Sen, and O. Spatscheck, "Over the top video: the gorilla in cellular networks," in *IMC*, Mar. 2011.
- [16] B. Wang, J. Kurose, P. Shenoy, and D. Towsley, "Multimedia streaming via TCP: An analytic performance study," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 4, no. 2, pp. 16:1–16:22, May 2008.
- [17] F. Dobrian, V. Sekar, A. Awan, I. Stoica, D. Joseph, A. Ganjam, J. Zhan, and H. Zhang, "Understanding the impact of video quality on user engagement," in *ACM SIGCOMM*, Aug. 2011.
- [18] D. Kotz and T. Henderson, "CRAWDAD: A Community Resource for Archiving Wireless Data at Dartmouth," *Pervasive Computing, IEEE*, vol. 4, no. 4, pp. 12–14, 2005.
- [19] M. Ploumidis, M. G. Papadopoulou, and T. Karagiannis, "Multi-level application-based traffic characterization in a large-scale wireless network," in *WoWMoM*, 2007.
- [20] A. Botta, A. Dainotti, and A. Pescapé, "A tool for the generation of realistic network workload for emerging networking scenarios," *Computer Networks*, vol. 56, no. 15, pp. 3531–3547, 2012.
- [21] A. Dainotti, A. Pescapé, P. S. Rossi, F. Palmieri, and G. Ventre, "Internet traffic modeling by means of hidden markov models," *Computer Networks*, vol. 52, no. 14, pp. 2645–2662, 2008.
- [22] M. C. Weigle, P. Adurthi, F. Hernández-Campos, K. Jeffay, and F. D. Smith, "Tmix: A Tool for Generating Realistic TCP Application Workloads in ns-2," *SIGCOMM Comput. Commun. Rev.*, vol. 36, no. 3, pp. 65–76, 2006.
- [23] D. Ammar, T. Begin, and I. Guerin-Lassous, "A new tool for generating realistic internet traffic in ns-3," in *SIMUTools*, Mar. 2011.
- [24] S. Sarvotham, R. Riedi, and R. Baraniuk, "Network and user driven alpha-beta onoff source model for network traffic," *Computer Networks*, vol. 48, no. 3, pp. 335 – 350, 2005.