# Energy-Efficient On-Demand Streaming in Mobile Cellular Networks

*Yong Cui*, Xiao Ma*, Jiangchuan Liu† and Yayun Bao**
**Department of Computer Science and Technology, Tsinghua University, Beijing, P.R. China*
*†School of Computer Science, Simon Fraser University, Burnaby, BC, Canada*
**cuiyong@tsinghua.edu.cn, max11@mails.tsinghua.edu.cn, jimmybao0730@gmail.com*
*†jcliu@cs.sfu.ca*

## I. Introduction

With the rapid development of wireless access technologies and mobile terminals (e.g. tablet, smartphone), end users are able to enjoy abundant applications with heavy workload on mobile platforms from anywhere at anytime. However, due to limited terminal size and portable requirement, mobile devices are difficult to provide sufficient computation and battery capacity, hence making the Quality of Experience (QoE) of mobile platform applications hard to guarantee [1].

Among the different sorts of heavy-workload applications, media streaming is a representative one [2]. Since media content needs to be transmitted, decoded and played on mobile devices, the energy of both wireless network interface, processor and screen is cost during the entire process. Recent literature has shown that transmission costs up to 50% of total energy on average [3, 4]. In the literature, media content encoding/decoding [5, 6] or traffic shaping [7, 8] are usually employed to reduce transmission energy cost. The former one needs to modify the media encoding/decoding pattern to reduce the amount of transmission or change the traffic pattern in a more energy-efficient way, while the latter one is transparent to both source and destination and is more flexible to multiple video/audio encoding/decoding formats.

In the last one decade, 4G LTE standardized by 3GPP [9] has witnessed a rapid development. With the advantages as high speed, wide coverage range, etc., it has attracted an increasing number of researchers to devote themselves to it. However, recent literature has shown that both 3G and 4G LTE are not adapt for media streaming transmission from energy consumption perspective [10, 11]. The power control mechanisms of 3G (RRC) and 4G (DRX/DTX) both have their inherent limitations that make the wastage of energy even more severe. For instance, the inappropriate setting of the inactivity timer could increase tail energy cost by nearly 60% percentage in 3G network [12]. Moreover, real-environment experiments have shown that the signal strength of cellular radio can also significantly influence the unit transmission energy cost per bit. According to Bartender [13], when the signal is weak, the energy consumed per bit is as much as 6x higher than when it is strong. Therefore, the traffic scheduling objective is to make transmission happens during superior network condition and to avoid too many tail states.

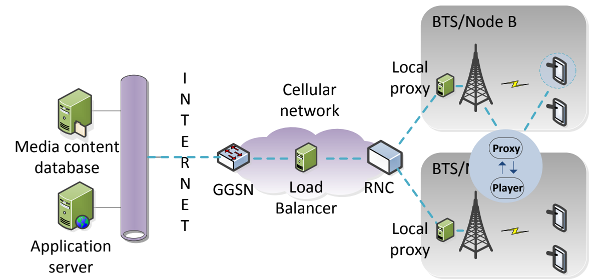



Figure 1. System Architecture

In this paper, we consider energy efficient design of media streaming applications in mobile cellular networks. Particularly, since the energy saving in our solution is achieved by traffic scheduling, to ensure sufficient schedulable time interval, on demand video streaming is considered as the objective application since the server possesses the entire media content at the beginning of the scheduling process. We employ proxy-based traffic shaping in our solution to avoid transmission during high energy-consuming radio state and to reduce unnecessary occurrence of tail energy state. To ensure high quality of experience (QoE) of streaming user, we propose a deadline-based strategy. However, since the performance is based on accurate signal strength prediction and it may cause too many tail states, we further employ Lyapunov Optimization as a tool to solve multi-client traffic scheduling problem to achieve the objective of minimizing total energy cost and guaranteeing users' QoE simultaneously. Simulation result has shown that our scheme can achieve a feasible tradeoff between the QoE of media streaming applications and the energy saving of mobile clients.

## II. System Architecture

In this section, we intend to illustrate the basic idea of our scheduling mechanism: traffic shaping by proxy between media streaming server and client to make transmission happens during superior network condition; and control the number of bursts to avoid

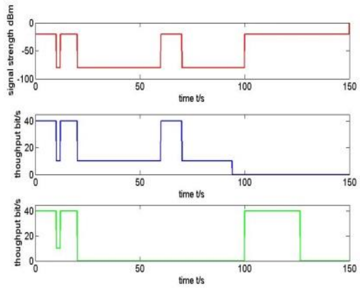



Figure 2. An example of traffic shaping

too much energy wasted by tail state. Figure 1 illustrates our basic system architecture. It mainly consists of a Load Balancer (LB), several Local Proxies (LPs) and client-side proxy (CP) locating on each mobile client.

Local proxy locates behind each base station (BTS)/Node B and is responsible for shaping the streaming traffic directed to each client within the coverage area of the BTS according to the signal strength of each client. Therefore, client-side proxy should transmit its signal strength variation to the local proxy hence the proxy can make scheduling decision in a more energy-efficient way. Since caching streaming data on local proxy may cause congestion from network operator perspective, a load balancer is involved in mobile backbone network to dynamically schedule traffic among different BTSs.

While deferring traffic to avoid bad-signal periods is a simple idea, it has many potential problems. First, media streaming is delay-sensitive, how to ensure user's quality of experience during the entire scheduling process is a problem. Second, after traffic shaping, the original continuous traffic may be broken into several traffic bursts, which may cause more time of cellular radio spent on tail state. Third, consider multiple clients locating within the coverage area of one base station and require media content simultaneously, the signal strength based scheduling may lead to traffic distribution imbalance on base station hence the channel resource is not sufficiently utilized. In this paper, we intend to design a longterm online optimization to achieve high QoE and low energy consumption and make the system stable.

Figure 2 provides a simple example of our traffic shaping mechanism. The sub-figure on the top is the signal strength variation from 0s to 180s. Without traffic shaping, the entire transmission may perform like the figure in the middle. Since signal strength also has an impact on throughput, the variation trend of throughput is consistent with that of signal strength. There is only one tail state since the transmission is not interrupted. If traffic shaping is enabled, the transmission from 20s to 100s is deferred to avoid bad-signal period, two tail states are incurred. If the energy cost by one additional tail state is lower than the energy wasted by transmitting during bad-signal period and the QoE can be guaranteed throughout the entire process, traffic shaping obtains positive profit. Furthermore, if the user is extremely power hungry, the energy saving can be achieved even more under sacrificing a portion of QoE.

**III. Solution Analysis**

During the entire transmission process on local proxy, the cached media content cannot be infinitely delayed since otherwise the user-side playback may be stuttered. If the video/audio playback rate is obtained by local proxy, the deadline of each data packet can be predicted by the proxy. We define the schedulable time duration of a data packet to be the time from entering local proxy to the deadline. Therefore, a simple strategy would be to choose the time slot within the schedulable time duration in which the signal strength is better. However, this strategy may need to predict the signal strength in the near future, the effect highly relies on the accuracy of the prediction. Moreover, it does not consider the overhead brought in by additional tail states (since one continuous flow is broken into multiple bursts). In the literature, there are several related work following this prediction-based strategy [13, 14]. Since the strategy is not practical and has many disadvantages, we try to propose an online traffic scheduling mechanism that only relies on current state information.

Assume there are $N$ $(1, 2, ..., N)$ mobile clients located within the coverage area of one BTS. To model the transmission process of client $i$, we denote the amount of data on the local proxy to be transmitted to client $i$ at the moment t to be $Q_i(t)$, the amount of data belonging to client $i$ that arrives proxy at the moment t to be $A_i(t)$. Since the throughput and signal strength vary with time, we denote the throughput at the moment t to be $w_i(t)$, the transmission energy cost per bit to be $P_i(t)$. Therefore, the decision should be made to determine whether transmitting data to client $i$ at time slot $t$ or not. Define a decision parameter $\alpha_i(t)$ to be,

$$\alpha_i = \begin{cases} 1 & transmit\ Q_i(t) \\ 0 & idle \end{cases} \qquad (1)$$

Therefore, we can define the amount of data that leaves local proxy and prepares to transmit to client $i$ to be $b_i(t) = \alpha_i(t) w_i(t)\tau$, Where $\tau$ is the duration of one time slot. Hence, the energy cost of client $i$ at time slot $t$ can be denoted by $E_i(t) = \alpha_i(t)( P_i(t)\tau + P_{tail} t_{interval})$, where $P_{tail}$ is the average tail power and $t_{interval} \in (0, T_{tail})$ is the duration of the tail state. Note that we both consider the transmission energy and tail energy cost.

As a result, the average energy cost of the N clients over a long time duration can be depicted as follows,

$$\overline{E} = \limsup_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^{N} E_i(t) \quad (2)$$

To depict the requirement of user's QoE, the average queue length *Q(t)* of the *N* clients throughout a long time duration should be stable,

$$\overline{Q} = \limsup_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^{N} Q_i(t) < \infty \quad (3)$$

Having the equations above, our energy-QoE tradeoff problem can be formulated by:

$$\min\ E \quad s.t.\ Q < \infty \quad (4)$$

To achieve system stability in the long run, we employ Lyapunov optimization in control theory to address the problem formulated by Eq.4. We first define a Lyapunov function $L(Q(t)) = \frac{1}{2}\sum_{i=1}^{N} Q_i^2(t)$ to depict the congestion level of the queue. The larger the *L* function is, the more the clients will be suffered from congestion. To keep the system stable, we involve Lyapunov drift $\delta(Q(t))$ to depict the difference of the *L* functions between two adjacent time slots,

$$\delta(Q(t)) = L(Q(t+1)) - L(Q(t)) \quad (5)$$

Therefore, we can minimize the following equation to tradeoff between delay and energy,

$$\delta(Q(t)) + VE(t) \quad (6)$$

Where *V* is a non-negative weight that is chosen as desired to affect an energy-delay tradeoff. The more the V is, the more the energy saving effect will be and vice versa.

By bounding the upper-bound of Eq.6, it can be derived that we can minimize the following simplified equation to minimize the upper-bound of Eq.6,

$$VE(t) - \sum_{i=1}^{N} Q_i(t) b_i(t) \quad (7)$$

As a result, the traffic scheduling algorithm can be easily designed based on the analysis above. At every time slot *t*, choose an appropriate decision vector $\alpha(t) = \{\alpha_1(t), \alpha_2(t), ..., \alpha_N(t)\}$ to minimize Eq.7. At the next time slot *t+1*, update $Q_i(t+1) = Q_i(t) + A_i(t) - b_i(t)$ based on the decision made before. The QoE and energy saving can be simultaneously well-guaranteed since Lyapunov optimization has provided such good characteristics.

## IV. Simulation

In this section, we evaluate our scheme on C++ platform. The throughput at the beginning of each time slot follows a uniform distribution on [0, 200*bps*]. The number of users is $N = 20$. The average tail power is set to 0.7*W*. The duration of each time slot $\tau$ is 100*ms* and data packet size is 100*bits*.

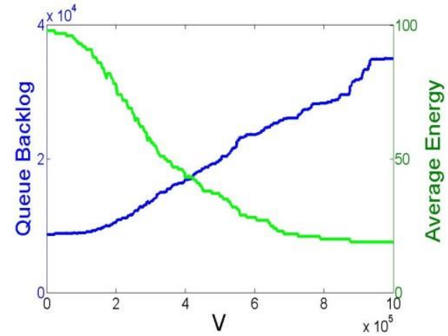



Figure 3. The influence of tradeoff parameter V

Figure 3 illustrates the variation of the time-averaged queue backlog and the average energy consumption with the change of tradeoff parameter *V*. We can see from the figure that the average energy consumption falls quickly at the beginning and then tends to decrease slowly while the time-averaged queue backlog grows linearly with *V*. It verifies our theoretical analysis that the parameter *V* can significantly influence the tradeoff between energy saving and QoE of media streaming application.

## V. Conclusion

Mobile streaming service has become increasingly popular on mobile handset platforms in the last one decade. However, a large amount of energy is wasted during media content transmission due to the inherent limitations of power control mechanisms of cellular radio. This paper introduces an energy-aware traffic shaping mechanism that saves media streaming energy consumption by re-scheduling traffic to avoid transmission during bad network condition and to decrease tail energy wastage. To achieve load-balancing from network operator perspective, we develop a network-wide long-term optimization mechanism in an online fashion to avoid rush hour on BTS/node B in cellular networks. Simulation results have approved that our approach is able to achieve a feasible tradeoff between the QoE of media streaming applications and the energy saving of mobile clients.

## References

[1] N. Thiagarajan, G. Aggarwal, A. Nicoara, D. Boneh, and J. P. Singh, "Who killed my battery?: analyzing mobile browser energy consumption," in Proceedings of the 21st international conference on World Wide Web. ACM, 2012, pp. 41–50.

[2] R. Trestian, A.-N. Moldovan, O. Ormond, and G.-M. Muntean, "Energy consumption analysis of video

streaming to android mobile devices," in Network Operations and Management Symposium (NOMS), 2012 IEEE. IEEE, 2012, pp. 444–452.

[3] M. Hoque, M. Siekkinen, and J. Nurminen, "Energy efficient multimedia streaming to mobile devices: A survey," 2012.

[4] H. Zhu and G. Cao, "On supporting power-efficient streaming applications in wireless environments," Mobile Computing, IEEE Transactions on, vol. 4, no. 4, pp. 391–403, 2005.

[5] Y. Eisenberg, C. E. Luna, T. N. Pappas, R. Berry, and A. K. Katsaggelos, "Joint source coding and transmission power management for energy efficient wireless video communications," Circuits and Systems for Video Technology, IEEE Transactions on, vol. 12, no. 6, pp. 411–424, 2002.

[6] S. Mehrotra, W.-g. Chen, and K. Kotteri, "Low bitrate audio coding using generalized adaptive gain shape vector quantization across channels," in Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on. IEEE, 2009, pp. 9–12.

[7] S. Sharangi, R. Krishnamurti, and M. Hefeeda, "Energy-efficient multicasting of scalable video streams over wimax networks," Multimedia, IEEE Transactions on, vol. 13, no. 1, pp. 102–115, 2011.

[8] C.-H. Hsu and M. M. Hefeeda, "Broadcasting video streams encoded with arbitrary bit rates in energy-constrained mobile tv networks," Networking, IEEE/ACM Transactions on, vol. 18, no. 3, pp. 681–694, 2010.

[9] "3gpp lte," http://www.3gpp.org/LTE.

[10] F. Qian, Z. Wang, A. Gerber, Z. M. Mao, S. Sen, and O. Spatscheck, "Top: Tail optimization protocol for cellular radio resource allocation," in Network Protocols (ICNP), 2010 18th IEEE International Conference on. IEEE, 2010, pp. 285–294.

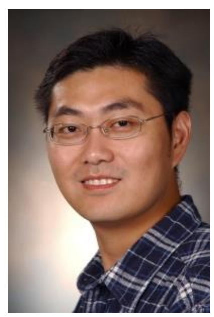
[11] J. Huang, F. Qian, A. Gerber, Z. M. Mao, S. Sen, and O. Spatscheck, "A close examination of performance and power characteristics of 4g lte networks," in Proceedings of the 10th international conference on Mobile systems, applications, and services. ACM, 2012, pp. 225–238.

[12] N. Balasubramanian, A. Balasubramanian, and A. Venkataramani, "Energy consumption in mobile phones: a measurement study and implications for network applications," in Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference. ACM, 2009, pp. 280–293.

[13] Schulman, V. Navda, R. Ramjee, N. Spring, P. Deshpande, C. Grunewald, K. Jain, and V. N. Padmanabhan, "Bartendr: a practical approach to energy-aware cellular data scheduling," in Proceedings of the sixteenth annual international conference on Mobile computing and networking. ACM, 2010, pp. 85–96.

[14] S. Chandra and A. Vahdat, "Application-specific network management for energy-aware streaming of popular multimedia formats." in USENIX Annual Technical Conference, General Track, 2002, pp. 329–342.

**Yong Cui,** Ph.D., Professor in Tsinghua University, Council Member in China Communication Standards Association, Co-Chair of IETF IPv6 Transition WG Softwire. Having published more than 100 papers in refereed journals and conferences, he is also the winner of Best Paper Award of ACM ICUIMC 2011 and WASA 2010. Holding more than 40 patents, he is one of the authors in RFC 5747 and RFC 5565 for his proposal on IPv6 transition technologies. His major research interests include mobile wireless Internet and computer network architecture.



**Jiangchuan Liu** is an Associate Professor in the School of Computing Science at Simon Fraser University, British Columbia, Canada. His research interests are in networking, in particular, multimedia communications, peer-to-peer networking and cloud computing. He is a co-recipient of IEEE IWQoS'08 and IEEE/ACM IWQoS'2012 Best Student Paper Awards, IEEE Communications Society Best Paper Award on Multimedia Communications 2009, IEEE Globecom'2011 Best Paper Award, and ACM Multimedia'2012 Best Paper Award. He is a recipient of the NSERC 2009 Discovery Accelerator Supplements (DAS) Award, and a recipient of 2012 Research Excellence Award from SFU Faculty of Applied Science.