Mobile Video Chat : Issues and Challenges

Shraboni Jana, Amit Pande, An (Jack) Chan and Prasant Mohapatra
University of California, Davis, CA 95616
Email: {sjana, pande, anch and pmohapatra}@ucdavis.edu

Abstract—The next generation of cellular communication needs to endure a higher demand for enhanced real-time multimedia support to end users. Amongst interactive video applications, use of mobile video chat are on the rise in both enterprise and consumer worlds. Mobile video chat are resource-constrained and heterogenous with varying display sizes, processing powers, network conditions and battery level and poses real-time delivery constraints unlike traditional video streaming applications. With the demand for anywhere any-time connectivity for video chat, it becomes quintessential to develop techniques that adapt to the underlying network, optimize the device architectures and provide best possible video chat experience to end users.

In this paper, we identify main limitations and challenges in mobile video chat. We further discuss the possible solutions and propose research directions to make video chat a good experience.

1 INTRODUCTION

Video chat, also referred to as video telephony or video calls, is the new data hog in current cellular networks. A simple phone call pales in comparison to a face-to-face video chat. Recently, group interactive video chat applications have also emerged. In group video chat, multiple users can see and chat with each other on a single screen at the same time. GigaOM Pro [1] projects video chat consumers via mobile phones will make noticeable gains by 2015 (Figure 1). Undoubtedly video chat is now playing a bigger presence in our daily lives and changing the paradigm of interpersonal communications [2].
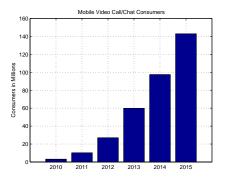

Fig. 1: Exponential increase in mobile video chat customers (Source : GigaOM Pro [1])

Many video chat applications for mobile phones have been rolled-out in the market. Different applications have different features and user experience. Fring, Tango, Skype, Vtok, FaceTime, ooVoo, Google talk are just few examples of those applications.

Video chat requires a high Quality of Service (QoS) to meet user needs and expectations in terms of packet loss rate, end-to-end delay, jitter etc. Many standards bodies and industry organizations have come up with various classifications of services and associated QoS parameters, for example, the video chat over 3G (Universal Mobile Telecommunications System, UMTS) falls under interactive-class with best effort service. Due to its interactive nature, video chat has very strict timing constraints and is intolerant of packet loss. The ITU G.114 standard recommends a maximum delay of 150 ms for video chat in order to ensure lip-synchronization. For a "good" to "excellent" performance, current uplink and downlink data rate for video chat is in the range from $\approx$ 130 kbps to 330 kbps which is subject to change with more powerful devices making into the hands of end-users.

The growth of video chat becomes a source of extra revenue for mobile operators. But as the trend will not stop after 2-3 years with approx. 140 million video chat consumers as in Figure 1, the future growth of video chat traffic is a concern for the network service providers. The operators are concerned about how video chat traffic may affect their networks. In addition, facilitating such growth requires advancement in application-level coding techniques, network-layer video delivery, quality assessment strategies, security etc. and an interplay of these issues.

In this paper, we present issues and challenges in design of successful mobile video chat applications. Figure 2 gives a system overview of main issues in mobile video chat. The subjective and objective user experience measurements, privacy and secrecy concerns, current network (communication channel) characteristics, end-device capabilities and inter-operability issues all play a crucial role in providing best user experience to end user. These issues will be registered by a controller which can decide right parameter choice for all components, while resolving any conflicts.

Our study is spread out in the next three sections. In Section 2, we give a summary of overall mobile video chat and relevant network architecture. In Section 3,
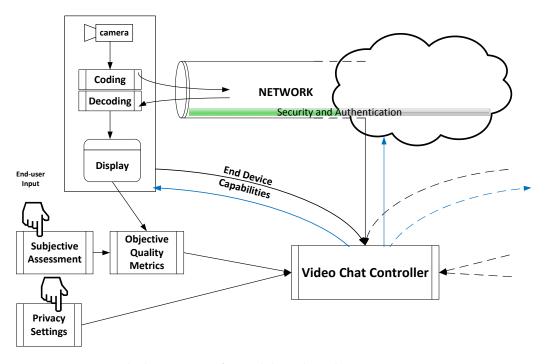
Fig. 2: Block Diagram of a Mobile Video Chat Engine

we discuss the limitations of end-to-end cellular networks and possible solutions that can improve video chat performance. Section 4 discusses the challenges in providing additional features, to make video chat easy to use application. We conclude our article in Section 5.

2 MOBILE VIDEOCHAT

In this section we discuss various components - mobile phone transmitter & receiver, video chat software, access segment, core network which are involved in the mobile video chat over cellular network and affect the performance of video chat. Access uplink-downlink segment is wireless radio link. Core network is wired.

We broadly divide the network into three parts - Mobile Devices, Application Provider and Mobile Carriers. We will discuss them in details in the next few subsections.

2.1 Mobile Devices

The video content generators and receivers are the mobile phone devices (end devices, Figure 2) which are gradually being replaced by smartphones. We use the term smartphone(s) or mobile phone(s) to address both the transmitter and receiver devices in this article. The improvement in processing power, speed, screen resolution, RAM and storage space boosts rich functionalities on mobile phones making high quality video chat possible in these devices.

Processing power and processing speed play a key role in the performance of any application in a device.Processors like NVIDIA TEGRA 4 have already

reached speeds $\approx$ 1.9 GHz with 72 GPU cores supporting ultra-HD video processing at low power consumption. Current smartphones are equipped with high resolution, low latency front-facing camera to capture end-user video while end-user is viewing the smartphone screen with pixel densities of 1080p. Display resolution has increased to 1080 x 1920 pixels, 5.0 inches ($\sim$ 441 ppi pixel density, in Samsung GS4) in the mobile phone devices, reducing the impact of resolution loss. There are now megabytes of RAM ($\approx$ 512 MB) available and also gigabytes ($\approx$ 32 GB) of space available in smartphones.

With the current advancements in mobile phones, high-quality video chat applications can be developed and deployed on mobile devices which was not possible before.

2.2 Application Provider

With the advancement in mobile phone hardware, obviously the onus is on video chat application providers to provide good quality of experience to end-users. Fring, Tango, Qik, Google Talk are some of the video chat applications contending to be frontrunner in mobile phones.

Video chat application needs to coordinate with the mobile phone hardware and network for efficient content creation satisfying all parties including end-users and network operators. Any video chat application needs to provide a range of functionalities - capturing, processing and transmission of video data. In Figure 2, though we show interaction of coding, decoding, camera for capturing with Video Chat Controller, we elaborate each of such functionalities below.
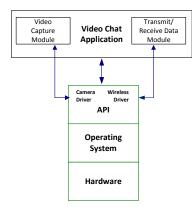
Fig. 3: A simplified view of smartphone hardware-software architecture and application providers' functions.

Video capture: The operating systems of the mobile phones provide Application Programming Interfaces (APIs) for video chat software to coordinate camera functionality to capture the content and transmit the content data through wireless channel (Figure 3). The software decides the frame-rate at which the camera will capture unique consecutive images and then save it to memory. Appropriate frame-rate is very important to maintain a balance between bandwidth, delay, computational power and user expectations. When users are communicating via sign language, for extreme cases, at least 21 frames-per-second (fps) is recommended in order to support finger spelling. In contrast, various studies suggest that audio and video are perceived as synchronized at minimum 5 fps [3] and therefore the frame rate requirements for video chat is around 5 fps. But with increasing user expectations, the frame-rate of 5 fps will not suffice.

With high-end smartphones making into the hands of end-users and increased network bandwidth due to 4G, future video chat is expected to provide higher bitrate to the end users.

Video compression: To reduce the bandwidth requirements, video application transmit compressed video. This encoding is commonly referred as source coding or simply video compression and is performed by the codec in an application. The higher the compression ratio, the more computational power is required but lesser is the bandwidth requirement. Current video coding standards, H.264 and VP8 both claim to provide the efficient compression and low bit-rates. In the study conducted by authors in [4], H.264 shows lower bandwidth usage and better video quality compared to VP8. High Efficiency Video Coding (sometimes referred to as H.265) is on its way to standardization and gives upto 40-50% improvement in bitrates.

Though current state-of-art H.264 codec has become very matured with efficient compression and low-latency, the codec needs to be further customized for video chat. The parameters of codec - frame rate, group of picture, size, quantization parameter and resolution need to be suited to the end-system capabilities. For example - Two mobile devices with resolutions x X y (camera) and x_1 X y_1 (display) in a group video chat may decide for a common shared resolution x_2 X y_2 based on camera resolution of mobile devices rather than using H.264 spatial scalability feature which may involve high latency.

Layered video coding approaches such as H.264 SVC (Scalable Video Coding) lead to increased computational cost and increased overall bitrate which is detrimental to performance of one-to-one video communications. However, it may play significant role in group video chat where multiple users have varying network conditions.

Video transmission: The application uses wireless driver APIs to transmit the data.

2.3 Mobile Carriers

In today's highly competitive environment, users have the option of choosing from a plethora of carriers. Therefore, it is not enough to simply make the services available to users. The mobile carriers must deliver those services in such a way that users fully enjoy a rich experience at a reasonable price. In this section we discuss the key technologies used by the mobile carriers for video chat. The mobile carriers are responsible for end-to-end seamless connectivity of video chat application. Video chat applications demand strict QoS. As mentioned before it requires packet delay $\leq$ 150ms and bandwidth requirement $>$ 100 kps. The frame error rate needs to be $\leq$ 1% [5]. Therefore, continuous feedback is required between the network and Video Chat Controller (Figure 2) for balancing between networking conditions and end-users video chat expectations. We now focus our discussion on network architecture relevant to video chat.

3G uses both circuit-switch mobile core network and packet network while 4G is a packet network. 3G-324M protocol is used to support real-time multimedia services over wireless circuit-switched 3G networks. The circuit-switch network provides with 64 kbps circuit switched path but published measurements [6] with live 3G networks, gives uplink speed to be only 50 kbps. As current upstream-downstream minimum bandwidth required by the video chat application is 128 kbps, circuit-switched mobile core network is unsuitable for video chat.

Figure 4 gives high-level network architecture in an end-to-end video chat. The dominant standard for transmitting video telephony in packet switched networks is RTP/UDP/IP. Internet Protocol, IP is a connectionless network communications protocol. It may provide greater bandwidth but the bandwidth is not guaranteed. Currently two versions of IP - version 4 (IPv4) and version 6 (IPv6) are in use. IPv4 is best effort service and IPv6 is designed to ensure QoS. To enhance real-time multimedia mobile services for IP, agnostic of access network, IP Multimedia Subsystem (IMS) provides an
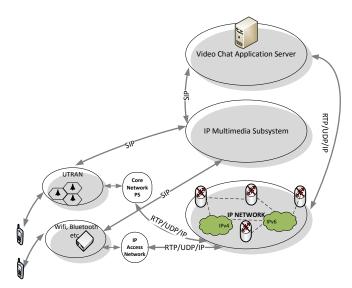
Fig. 4: High level network architecture in an end-to-end Video Chat

Fig. 5: Evaluation of video chat in 4G/wifi network.

architectural framework for a flexible multimedia management. IMS ensures QoS, negotiating mobile device requirements using Session Initiation Protocol (SIP), during session set-up or modification. Once end-to-end QoS is established, the end-user terminals use RTP protocol to packetize video chat data and send it using transport layer protocol, UDP over IP.

To transport 3G/4G services through IP networks and for providing end-to-end QoS provisioning for an IP packet, IETF defines two models - IntServ and DiffServ. The IntServ model uses Resource Reservation Protocol (RSVP) to signal and reserve the desired QoS for each flow in the IP network. Under IntServ, video chat has a very strict guaranteed service providing firm bounds on end-to-end delay and assuring bandwidth for the traffic. Though it is possible theoretically, to provide such QoS for each flow in the network, practically it is very hard as every device along the path of a packet, needs to be fully aware of RSVP and be capable of delivering the required QoS. The DiffServ model is relatively simple and coarse as they group the network flows based on different classes, also called Class of Service (CoS) and applies distinct QoS parameters for each class. The Type of Service Octet in IPv4 stores the 6-bit Differentiated Services Code Point in IP header to identify the CoS whereas in IPv6 Traffic Class Octet is used. Under Diff-Serv model, the delay-sensitive video chat traffic falls under conversational class of service.

Upcoming 4G communication network, Long Term Evolution-Advanced (LTE-A) promise a data rate of 500 Mbps uplink and 1 Gbps downlink at peak. With such increased bit-rates, in networks, the quality of delivered video is also expected to increase. Full High Definition videos require a bandwidth of nearly 2 Mbps, which may become a requirement for video chat in coming years.
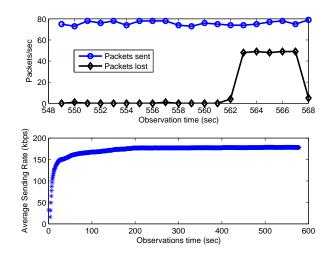
3 LIMITATIONS

We evaluate the network QoS (Figure 5) in an experiment carried out with video chat application "Vtok" (based on Google talk APIs) between two end-users, one connected to 3G and another connected to wifi network. The end-devices used are smartphones Samsung Galaxy II. We use Tcpdump to capture packets sent and received at two stations performing video chat. In a random packet trace, we observe that while the video chat is within the limits (Average Sending Rate, Figure 5), the packet loss is sometimes around 50%, which is way above the required network QoS thus reaffirming our belief that a robust video chat is far from realization.

Current wireless mobile video applications performance is limited due to the constrained network resources and fading and interference caused by the wireless medium. In this section, we discuss limitations of any real-time video application on mobile phones. We identify techniques which might come to aid in addressing these limitations for mobile video chat.

3.1 Bandwidth

Video chat requires real-time communication. If the application over utilizes the link, it causes unfairness to other traffic and if it under-utilizes the link, it may cause low quality of video chat. In addition to this, the uplink-downlink connection in 3G is asymmetric. When congestion happens, the video chat on the uplink will suffer first, thus determining the quality for the whole video chat. Hence, lower uplink bandwidth needs to be widely available for live video chat.

Multicasting can be one of the option to reduce bandwidth requirement when more than one participating users are present in the video chat. The multicasting allows for the serving of a single stream that is replicated throughout the network, thereby, reduces the bandwidth required for both uplink and downlink. Multimedia Broadcast/Multicast Service (MBMS) is introduced by

UMTS to provide high-speed multimedia multicasting and broadcasting services. The e-MBMS, which is the evolved version of the legacy MBMS system is considered an important architecture of LTE-A in this regard. Though multicasting has not been a reality in 3G so far, but has a potential to minimize bandwidth requirement in group video chat.

The bandwidth consumed by the video chat can be further reduced with a mechanism to notify end-users about incoming video chat call. The reason being, video chat applications keep sending packets even during the idle period. The voice call, in such a case, has an edge over video chat in not requiring end-users to be always logged into the data network of the application. With automatic login during an incoming call, the end-user can receive the video call without being logged into the video chat application. As per the Skype users and technology news, when idle, Skype consumes (1.3 GB/month [1]) considerable bandwidth compared to the current capped data plans available for cellular end-users.

In other approaches, to address asymmetric uplink-downlink connection, video chat applications can use available WiFi hotspots. Though WiFi brings a realistic and comfortable solution for bandwidth-constrained 3G to provide such high-end applications, maintaining QoS during such vertical handover is an issue.

3.2 *Face to Eye Delay*

Mobile video chat poses unique challenge due to its requirement for small face (transmitter) to eye (receiver) delay. ITU G.114 standard recommends a maximum delay of 150 ms for video chat in order to ensure lip-synchronization. Video chat having strict delay constraints, the retransmission mechanism of TCP may not be used. TCP retransmission increases delay in the system making it difficult to adhere to the delay constraints of video chat QoS. With UDP on the other hand, the network is unable to recover from packet losses. Delay and packet loss in the network causes jitters, frame-freeze, and video stalls. We will discuss the impact of packet-loss in later sub-sections.

Other causes of delays are horizontal and vertical handovers. Video chat can freeze or stall the video during handover when the delay is not within the limits. As the mobile moves, its associated access point changes as per the current location. The delay in such horizontal handover is inevitable for 3G. As mentioned before, currently, mobile carriers to solve the bandwidth constraints, prefer video chat applications to use available WiFi hotspots. Maintaining delay constraints during such vertical handover is also a key limitation [7].

IEEE 802.21 aims for seamless handover across heterogenous networks reducing disconnection times and

packet loss during handover of real-time communications. LTE-CoMP, has been proposed by 3GPP standards [8] community to take care of seamless connectivity for edge users and handovers in 4G networks. In LTE-CoMP the edge users are also supported by multiple neighboring base stations, therefore giving more diversity to the data received. When the edge users cross the cell during chat, with LTE-CoMP, they will be still supported by adjacent base stations and hence will not result in video freeze due to handover delay.

3.3 *Interference, Fading & Congestion*

Interference and fading in the wireless medium and congestion in the network causes packet losses. Packet loss may cause visible distortions in video chat. The two main visible distortions are blocking and blurring. We find with a random trace, (Figure 5) that packet loss is a major issue in video chat. Though bandwidth and delay are within the required QoS limits, limitations imposed by them may cause packet-loss.

Packet loss is inevitable in wireless medium, because of its very nature. The simultaneous transmissions from or to different mobile phones in the same cell may result in interference. As the mobile moves to the edge of the cell, its sustainable data rate may change due to interference from adjacent cells. Moreover, the impact of packet loss is further increased as there is error propagation among successive frames of the codec. For real-time applications, it will be more appropriate for the application to dynamically change the parameters of the codec to minimize impact of packet loss.

For delay tolerant applications, the network operators can employ better loss-resistant network protocols. But as the packets are streamed over UDP for video chat, the onus mainly falls over video chat application providers to implement packet loss recovery techniques to preserve network stability. To combat packet loss, it is very important for the application to identify the reason for data loss. The data losses can be due to the video coding losses, random losses or congestion losses in the network. Random losses are caused by wireless access link - signal fading, interference and channel quality losses. End-to-end packet loss in cellular network and internet can be caused by congestion loss occurred due to buffer overflow in the network. Decreasing transmission rate will not help in case of non-congestion related losses. To identify the root cause of the data-loss, Spike [9], an end-to-end loss differentiation algorithm can be used.

The data recovery techniques (Figure 6(a)) in literature can be divided into - feedback based and non-feedback based. Non-feedback based techniques are implemented at the encoder or decoder. The decoder-based approaches such as interpolation, filtering, spatial and temporal smoothing are not very efficient. Encoder-based technique though efficient, adds redundant bits to the video chat data and therefore increases bandwidth

---

1. http://tech.blorge.com/Structure:%20/2009/02/24/skype-steals-bandwidth--even-when-you-are-not-using-it/, last accessed: March 27, 2013

Dataloss Recovery Techniques

Feedback based   Non-feedback Based

Full Retransmission   Partial Retransmission   Encoder-based   Decoder-based

(a)

1. Partial Retransmission

+ BW availability
+ Random Losses

+ BW availability
+ Congestion Losses

2. Encoder-Based

3. Decoder-Based

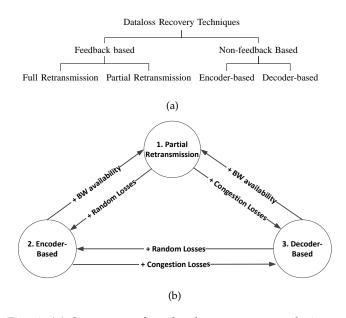+ Random Losses
+ Congestion Losses

(b)

Fig. 6: (a) Summary of packet-loss recovery techniques and (b) 3-Stage packet-loss recovery for video chat (+ indicates increase)

requirement. Forward Error Correction (FEC) and Joint Source and Channel coding are some of the encoder-based error control approaches. Such techniques require significant changes to video codec and has high computational complexity. Of all the encoder-based approaches, though FEC is a popular solution, studies using FEC in video chat application (Skype) show increase in bandwidth overhead by 25% to 50% [10]. Hence, in the congested environment FEC will likely increase the loss rate.

In feedback-based error-control techniques, information sent by the decoder is used by the encoder to adjust the coding parameters or retransmit lost packets to achieve better data recovery. In full retransmission entire data is sent again unlike partial retransmission. Both the approaches increases delay due to the increased round trip time. But partial retransmission takes up less bandwidth compared to full-retransmission. Some of the examples of partial retransmission are - Reference Picture Selection, Intra Update, media rate control protocols such as TFRC, DCCP etc.

Limitation imposed by the packet-loss for video chat is a challenge that needs to be addressed for giving end-users better quality for bandwidth-constrained network and delay intolerant video chat application. Generally a good video chat application, when incurred with heavy data-loss, should drop the video stream in order to preserve the audio stream as much as possible. In Figure 6(b) we propose 3-stage packet-loss recovery technique for mobile video chat adhering to its delay constraints. The plus sign (+) indicates increase in the network parameter. With bandwidth availability in the network, partial retransmission gives the best result.

Depending on the loss-type, the system can adapt to encoder or decoder based approaches in presence of losses. Encoder based approaches are more appropriate for random losses but have negative impact in presence of network congestion. The stages are numbered in the order of preference. To find the thresholds for switching between these 3 stages and satisfying end-user video chat experience is an important research area for video chat.

4   CHALLENGES

Video Chat Controller needs to address additional challenges apart from network and application limitations, discussed in the previous sections. It should measure upto end-user expectation both in terms of quality and security and end-device capabilities (Figure 2). In this section, we discuss these challenges in detail.

4.1   *Measure & Maintain user QoE*

Traditionally, Quality of Service (QoS) metrics such as packet loss, delay, and jitter have been used to measure the application performance on an end device. However, with increased diversity in application requirements of multimedia services and different content types, users Quality of Experience (QoE) instead of networks QoS is becoming an increasingly popular term in mobile video segment. QoE metrics tend to measure user perception of delivered multimedia services instead of counting on network service parameters. Many tools have evolved for evaluating video quality delivered to end user using subjective or objective methods.

Subjective quality assessment refers to algorithms which measure the "user-perceived" quality of received video. Typically users give a numerical score (1-5) on the perceived video quality. Methodologies such as oneclick or crowdsourcing can be employed to carry out video chat subjective assessment in a more economical way. Oneclick [11] uses a single dedicated key-click to convey the dissatisfaction of end users towards the video application in an on-line manner. Crowdsourcing applications like Amazon Turk[2] takes feedback from the mobile end-users in exchange of monetary benefit. These end-users are crowd sourced instead of selecting a pool of observer in a controlled environment.

On the other hand, objective methods refer to mathematical models that approximate results of subjective quality assessment (Figure 2), but are based on criteria and metrics that can be measured objectively and automatically using computational techniques. The full-reference objective metrics such as PSNR and SSIM are not suitable for real-time interactive content because they require copy of original video. However, metrics such as blocking and blurring which are no-reference metrics [12] can be used to quantify the network impairment in the video for video chat.

2. https://www.mturk.com/mturk/welcome, last accessed: March 27, 2013

Ideally, a mobile video chat must have robust mapping of subjective scores to objective measurements. While the chat engine can automatically assess the video quality using objective metrics, subjective framework (such as one-click or its variant) can be used to fine-tune its performance. This feedback can be used by service providers to manage network and other service parameters of end-to-end system.

4.2 *Battery consumption*

Battery consumption has become a major issue for running high-end applications in the smartphones. The mobile phones, being small and supporting such wide range of applications, increases the pressure on battery life for such devices. Further, battery technology is not improving at the same pace as the processing power and speed of the mobile devices. With nearly 1500 mAh battery power in current smartphones, if a typical video application consumes roughly 300mW of power per hour when enabled, so neglecting all other power consumption on the phone, the battery is expected to last 4 - 5 hours. Effort has been made by Tango in this regard. Tango uses Google push notifications to wake up the mobile phone on receiving a call. Though this technique will avoid draining of mobile phone battery, but users do not get the provision to sign out if they do not intend to receive calls.

The best way to solve this issue is to design battery status aware and end-user expectation aware video chat applications. One such effort for image compression is presented in Poly-DWT architecture [13] which can morph its hardware requirements and image reconstruction quality at run-time leading to considerable savings in power.

4.3 *Interoperability*

The future video chat requires to be ubiquitous like voice call. The content receiver can be a different mobile phone device with different requirements under different operators and running different video chat applications. Two end-users can strike a voice call irrespective of the system and network differences. PSTN (Packet Switched Telephone Network) has been the enabling technology in this regard. Headway has been achieved in standardizing multi-vendor interoperability and operator interconnect using IMS based services. SIP, an open standard is used for control plane signalling in IMS. Devices from any vendor supporting SIP video chat will be able to interoperate with the SIP-based devices from any other vendor. As mentioned earlier, user plane traffic in IMS is RTP/UDP/IP based.

Currently, end-users using different video chat applications cannot communicate with each other. This is because video chat applications use different video codecs and technologies for system negotiation which have not been standardized so far. To give similar flexibility to end-users as voice call, video chat application providers need to work towards standardization.

4.4 *Security & Privacy*

Security and privacy concerns are inherent in any real-time interactive communication.

Encryption techniques are well-studied and can be employed to provide required level of security for video chat users. Joint compression and encryption schemes have been recently developed, which lead to significant savings in computational power while achieving good levels of unintelligibility of bitstream [14]. But privacy concerns still needs to be addressed for mobile video chat. Using diary and interview techniques, [15] points that privacy during video chat is a key concern of the end-users. The concern stems from the fact that the user at the other-end may make obscene gestures or the end-user when using video chat at highly crowded place wants to keep it private.

Blurring techniques have been used to make a given scene appropriate for users at other end to view. Though video blurring techniques can work appropriately as well as can protect users privacy in both contexts, they are not appropriate for the new generation of video chat services [16]. This is because the main function of the new generation of video chat services is to bring a user face-to-face via smartphones with another person from another corner of the world in his/her background. If the services provide users with all blurring faces or user background, the users who use these services will gradually lose their interest in talking to others.

Other attack models and countermeasures for video chat over smartphones needs in-depth analysis. The basic-level of privacy can be provided by the application provider by maintaining a public and a private profile of the end-users [16]. On the end-user preference, the application can display corresponding profile at the other end. Also its application provider responsibility to secure the end-user video chat accounts.

5 CONCLUSION

With the increased proliferation of smartphones and the advancement in communication technologies, the demand for video chat has increased tremendously. To catch up with this trend and exploit the economical incentive associated with it, both application providers and mobile carriers have significant roles to play in terms of improving the quality of video chat. Though there has been some approaches suggested in the literature, applicability of these solutions over cellular networks for high-quality end-to-end video chat has not yet been studied in detail so far. In this article, we lay down the limitations and summarize the challenges faced by video chat. We also discuss the possible solutions and their incompleteness. We believe that after addressing these challenges, the mobile video chat can be developed into its full momentum.

REFERENCES

[1] S. Higginbotham, "Can You See Me Now? The Future of Video Chat." [Online]. Available: http://gigaom.com/2010/06/07/can-you-see-me-now-the-future-of-video-chat/

[2] M. G. Ames, J. Go, J. Kaye, and M. Spasojevic, "Making love in the network closet: the benefits and work of family videochat," in *Proceedings of the 2010 ACM conference on Computer supported cooperative work*, 2010, pp. 145–154.

[3] J. Scholl, P. Parnes, J. D. McCarthy, and A. Sasse, "Designing a large-scale video chat application," in *Proceedings of the 13th annual ACM international conference on Multimedia*, ser. MULTIMEDIA '05, 2005, pp. 71–80.

[4] P. Seeling, F. H. P. Fitzek, G. Ertli, A. Pulipaka, and M. Reisslein, "Video network traffic and quality comparison of VP8 and H.264 SVC," in *Proceedings of the 3rd workshop on Mobile video delivery*, 2010.

[5] 3rd Generation Partnership Project, "ETSI TS 122 105 V10.0.0; Digital cellular telecommunications system (Phase 2+); LTE; Services and service capabilities (Release 10)," Tech. Rep., May 2011.

[6] J. Nurminen, O. Karonen, L. Farkas, and T. Partala, "Sharing the experience with mobile video: A student community trial," in *6th IEEE Consumer Communications and Networking Conference (CCNC)*, 2009, pp. 1–5.

[7] S. Sharma, N. Zhu, and T. cker Chiueh, "Low-latency mobile ip handoff for infrastructure-mode wireless lans," *IEEE Journal on Selected Areas in Communications*, pp. 643–652, 2004.

[8] 3rd Generation Partnership Project, "3GPP TR 36.819 v11.0.0, Coordinated Multi-Point Operation for LTE, 3GPP TSG RAN WG1," Tech. Rep., 2011.

[9] O. Boyaci, A. Forte, and H. Schulzrinne, "Performance of video-chat applications under congestion," in *IEEE International Symposium on Multimedia (ISM)*, 2009, pp. 213–218.

[10] J. Wang and D. Katabi, "Chitchat: Making video chat robust to packet loss," Tech. Rep., july 2010.

[11] K.-T. Chen, C.-C. Tu, and W.-C. Xiao, "Oneclick: A framework for measuring network quality of experience," in *INFOCOM 2009, IEEE*, 2009, pp. 702–710.

[12] Z. Wang and A. C. Bovik, *Modern Image Quality Assessment*, ser. Synthesis Lectures on Image, Video, and Multimedia Processing. Morgan & Claypool Publishers, 2006.

[13] A. Pande and J. Zambreno, "Poly-dwt: Polymorphic wavelet hardware support for dynamic image compression," *ACM Trans. Embed. Comput. Syst.*, vol. 11, no. 1, pp. 6:1–6:26, Apr. 2012.

[14] A. Pande, P. Mohapatra, and J. Zambreno, "Using chaotic maps for encrypting image and video content," in *IEEE International Symposium on Multimedia (ISM)*, dec. 2011, pp. 171–178.

[15] K. O'Hara, A. Black, and M. Lipson, "Everyday practices with mobile video telephony," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '06, 2006, pp. 871–880.

[16] X. Xing, Y.-L. Liang, H. Cheng, J. Dang, S. Huang, R. Han, X. Liu, Q. Lv, and S. Mishra, "Safevchat: detecting obscene content and misbehaving users in online video chat services," in *Proceedings of the 20th international conference on World Wide Web*, ser. WWW '11. ACM, 2011, pp. 685–694.