Implementing Voice and Video over IP for iPhones and Smartphones

Roar Hagen, CTO Global IP Solutions

Introduction

The Apple iPhone promised to create a whole new end-user experience for multimedia mobile communications—and it has. In particular, the recent release of the iPhone 3G has created a groundswell of interest in applications for mobile voice and video over IP. However, mobile networks, which introduce delay and packet loss, as well as devices with limited CPU and are prone to high levels of acoustic echo, present challenges to developing high quality VoIP and video applications to even seasoned media processing experts. Given the difficulties associated with mobile IP communications, it is crucial to the success of mobile VoIP applications that developers understand how best to deal with the network and device limitations that can affect call quality and ultimately the bottom line.

This paper addresses these issues and explains how applications can be VoIP-enabled to take advantage of the unique capabilities of smart phones such as the iPhone. It discusses some of the inherent difficulties that mobile IP networks present, as well as issues such as echo and OS tuning that are unique to mobile devices. Finally, we will present how these obstacles can be overcome to provide a high-quality end-user experience.

Mobile VoIP

The widespread adoption of the Internet in the 1990s presented an opportunity to utilize IP networks for communications. In its earliest incarnation, voice over IP (VoIP) was seen as a telephony replacement providing a low-cost alternative to traditional PSTN communications. However, at the time most audio codecs were designed for existing TDM networks, and video codecs were primitive at best. The result was inconsistent, almost unusable, voice and video communications. As technology evolved, voice processing solutions were developed specifically for IP networks, drastically improving sound quality. The result was an explosion in popularity of PC applications, such as Skype and Google Talk that leveraged technology like wideband codecs to produce better than PSTN quality, while offering free or inexpensive calls. A similar improvement in video technology has further expanded the interest in these types of applications, as IP networks allow for a rich, multimedia communication experience.

While the PC has proven to be a suitable platform for IP communications, the ubiquity and convenience of mobile devices makes them ideally suited to deliver the next generation of applications for voice and video over IP. Mobile VoIP was initially enabled by Pocket PC devices such as the iPaq which ran Windows CE and had WiFi access to IP networks. As demand for IP communications grew, traditional mobile phones enabled access to IP networks and included application operating systems, creating what are known as smartphones. These smartphones usually run Symbian or Windows Mobile and offer “dual mode” capabilities with both traditional mobile network access in addition to WiFi network access. Smartphones present the most promising opportunity for the growth of the mobile VoIP market, especially as solutions such as the iPhone and Google’s Android platform are made available.

Implementing Voice and Video over IP for iPhones and Smartphones, *continued*

Despite the opportunity presented by smartphones, providing VoIP on a mobile device is a technical challenge. First, mobile phones are small devices with limited processing capabilities, and can be fairly expensive with short battery lives. Therefore, the VoIP processing software must be very efficient in terms of memory usage and computational complexity. Some degree of optimization to the processor is usually necessary and most devices require fixed point implementation of the signal processing algorithms. These tasks are quite demanding and require not only programming skills, but also signal processing experts.

In addition to device limitations, the quality of the network connection presents a unique challenge to mobile VoIP application development. Mobile devices, by definition, operate over wireless links. For a VoIP call the typical wireless access methods are WiFi, 3G or Edge packet data connections. Many factors can cause disturbances in the radio transmission of these wireless links. For instance, low signal strength when the device is far away from radio base stations or access points can lead to a low signal-to-noise ratio as the background noise level becomes significant. Interference from other users communicating in the same or neighboring frequencies is another major factor degrading reception quality. This can lead to bit errors that trigger discarding of packets. In addition, for real-time communication such as VoIP to be effective, delay must be minimized, eliminating the viability of retransmission, thus leading to packet loss. In order to compensate for the high levels of packet loss associated with mobile VoIP, the receiver software needs to mask the error in the output signal through packet loss concealment (PLC) techniques. For mobile solutions, the PLC methods must be able to handle quite high peak packet loss rates that could be as high as 30%.

Another area of consideration for application development covers the bottlenecks that tend to exist in wireless VoIP. Too many users sharing the same frequency band or access point, especially for WiFi, lead to variations in the throughput and transmission times. The resulting phenomenon is known as jitter and is evident when the receiver experiences irregular intervals between packet arrivals. Figure 1 below shows an example of jitter measured for a WiFi channel.

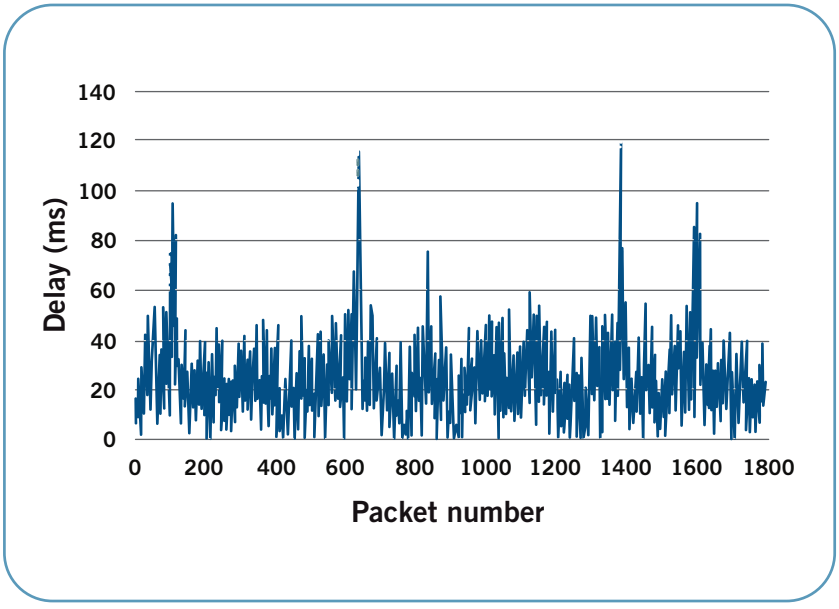

Figure 1. Delay variations – jitter – from a WiFi channel.

The best way to combat jitter is to add a buffer at the receiver side to smooth out the delay variation. The amount of buffering should be adaptive and roughly follow the envelope of the jitter curve. The best and most adaptive solutions should be able to minimize delay while accounting for jitter, yielding the clearest and most consistent conversations.

Implementing Voice and Video over IP for iPhones and Smartphones, *continued*

The iPhone as a VoIP Platform
The launch of the iPhone introduced ground-breaking design and multimedia capabilities to the mobile world. The iPhone's audio and video handling capabilities make it an extremely attractive platform for real-time IP communication. The phone's open APIs and flexible design mean that any developer familiar with the Mac OS can now enable high-quality voice calls, with the possibility of adding real-time video in the future. The IP network access consists of both WiFi and 3G/Edge channels thereby allowing VoIP calling in virtually any location, bringing the promise of ubiquitous IP communications one step closer.

While these capabilities are exciting, devices such as the iPhone present additional complications to designing VoIP applications. For instance, in addition to the traditional jitter problems found in mobile devices, smartphones like the iPhone introduce extra jitter when they need to simultaneously perform several tasks for the end user. For example, the iPhone may try to check for email during a call, leading to delays in the processing of the VoIP call. The effect experienced by the VoIP software is the same as network jitter, further emphasizing the need for a solid jitter buffer. Also, since users tend to want access to the iPhone screen while talking, hands-free operation is a typical usage scenario. The further a user moves from the microphone, and the more amplified the speaker volume, the greater the acoustic echo becomes. This heightens the already lofty requirement for efficient Acoustic Echo Cancellation (AEC) posed by most mobile devices. AEC is a tough problem on most platforms that is made even tougher on a mobile device with limited processing power.

These unique requirements demonstrate why having high-quality voice and video processing capabilities is so important. Because iPhone application developers are typically not audio or video experts, it is Global IP Solutions' goal to deliver real-time IP multimedia communication to the iPhone developer community.

Global IP Solutions for the iPhone
Global IP Solutions has been developing voice processing solutions since 1999, and has extensive experience with mobile VoIP. The company was founded on the principle that IP networks demand unique solutions to problems such as packet loss, delay and jitter. In order to combat these issues and deliver the best voice-quality possible, Global IP Solutions developed such innovative solutions as the iLBC (an IETF standard) and iSAC codecs, and its NetEQ jitter buffer. In 2002, the company continued this innovation by launching the first version of its voice processing package called VoiceEngine. VoiceEngine was initially launched for the PC and Windows and provided a complete package of VoIP sound processing that handles all the voice processing complexity for application developers. Soon thereafter, the first prototype of VoiceEngine Mobile for Windows CE on the iPaq was developed. This was a breakthrough because it demonstrated very high-quality VoIP over WiFi with a mobile device. VoiceEngine has since been a wildly popular product enabling such applications as Skype, Google Talk and Yahoo Messenger. It is a platform-independent product that has been available for the Mac OS for many years. The long mobile history of VoiceEngine, as well as Global IP Solutions' MAC experience made porting VoiceEngine to the iPhone a natural next step. Global IP Solutions also offers a VideoEngine Mobile product, and as the iPhone APIs become more open, developers will soon be able to create real-time voice and video solutions for the iPhone.

The unique requirements of smartphones, and the iPhone in particular, mean that VoiceEngine's ability to handle network impairments and device limitations are even more important for the mobile environment. The platform independent code base combined with a library of optimized code components for the ARM processor used in the iPhone make it possible to quickly release a very efficient high-performance software package for the iPhone without having to worry about CPU constraints. In addition, VoiceEngine contains a highly adaptive jitter buffer called NetEQ to manage the network impairments and additional jitter introduced on smartphones. NetEQ uses novel jitter handling methods as well as packet loss concealment techniques. It adapts very quickly to handle the required jitter and at the same time minimize buffering delay. Finally, VoiceEngine utilizes advanced AEC and noise suppression solutions to minimize the effects of echo and background noise that are so common in mobile environments. All of these capabilities were optimized specifically for the iPhone, making VoiceEngine the ideal solution for developers of mobile VoIP applications.

Implementing Voice and Video over IP for iPhones and Smartphones, *continued*

VoiceEngine Mobile for iPhone also provides a rich API that enables developers to integrate very high-quality VoIP into their target applications, while delivering the flexibility to adjust a variety of parameters, such as echo cancellation. Features include:

- Selection of voice codecs
- Echo cancellation
- NetEQ advanced jitter buffer
- Noise reduction
- Automatic Gain Control

Conclusions

As the market for real-time voice and video communications on mobile devices grows, developers need to be aware of the unique demands presented by producing applications for mobile devices. Factors such as jitter, delay, acoustic echo and CPU limitations must all be taken into consideration in order to produce a high-quality voice and video experience. Global IP Solutions is uniquely positioned to offer a suite of products that have been developed to maximize efficiency, and deliver interoperability, enabling developers to offer cutting-edge applications to their customers.

For more information on Global IP Solutions, please visit *www.gipscorp.com*

Global IP Solutions, Inc. Corporate Headquarters • 642 Harrison Street, Second Floor • San Francisco, CA 94107 USA
Phone: +1 415 397 2555 • Fax: +1 415 397 2577 • Email: info@gipscorp.com

GLOBAL IP SOLUTIONS
See the Difference. Hear the Difference.