Location Based Advanced Phone Dialer

A mobile client solution to perform voice calls over internet protocol

Jorge Duda de Matos

Superior Institute of Technology (IST)

Lisbon, Portugal

Abstract

Mobile communication systems on a near future will become Internet Protocol (IP) based using wireless access points at public hot spots as an alternative to today's high prices offered by telecommunication companies.

This thesis examines and presents the necessary steps for the development of an application capable of conducting a normal conversation between users using a Pocket PC with access to an internet connection on a peer-to-peer communication.

The main purpose is to provide a generic solution so that users of handheld devices can utilize their personal devices to enable internet telephony audio environment, voice messages and voice contextualization messages in an environment where internet connection may have congestion and packet loss at different levels affecting the quality of service and raising concerns about security and privacy.

This solution called Location Based Advanced Phone Dialer (LBAPD) is integrated with the Wizi web location based application (http://wizi.com) that provides the location functionality.

An approach to the open-source Session Initiation Protocol (SIP) client in a peer-to-peer telephony setting, ported to a Pocket PC environment was used to accomplish the task.

The second part of the paper describes the architecture and explores some of the implementation's limitations.

Keywords: Telephony, peer-to-peer, SIP, Internet, Pocket PC.
1 Introduction

Today there are many applications that allow voice calls over the Internet at a low cost between users on different points of the globe. The industry of handheld devices is growing and technological advancements in wireless software applications and platforms, and improvements in ease of use makes this kind of devices the perfect candidates to handle voice calls over the Internet.

wireless networks provides some mobility to users and their increasing use both at home and at public hotspots, raises the interesting possibility of using VoIP in mobile devices enabling full mobility for the users. Of course the use of such a service is limited compared with the mobility provided to the mobile telecommunication networks because there not as many hotspots as mobile cells.

The development of VoIP applications for mobile devices brings some challenges for bandwidth and speed which doesn’t allow the same measure of performance achieved for example by laptops or desktop computers.

2 System architecture

2.1 Background

VoIP stands for voice over the internet protocol IP and applications for VoIP typically comply with several protocols some standard such as SIP [1] (Session Initiation Protocol) or H.323 [2] for setting up the call and RTP [3] for real time audio exchange and there are also other proprietary protocols such as the ones used for Skype.

2.2 Components

![Figure 2: Main components.](image.png)

The VoIP system is divided three main components as shown in Figure 2 each with two functions, one when sending data and another when receiving data.
Sending audio data

- The first component responsible for capturing sound from the microphone, converting it from analog to digital and buffering it.

- The second component collects audio block data processed by the first module, encodes it as RTP packets, compresses, and add encryption to be ready for the next step.

- The third component is responsible for sending RTP audio data blocks over the network through UDP packets.

Receiving audio data

When a packet is received the reverse process is done and audio data blocks are sent to the speaker.

- The first component converts data from digital-to-analog and sends the audio to the speakers.

- The second removes encryption, decompresses and decodes the packets so that the first component is able to process it.

- The third component collects packets from network so that the second component can process it.

2.3 The SIP protocol

![System architecture using SIP](image)

Figure 1: System architecture using SIP.
SIP is an application-layer signaling protocol working in the third module allowing creation, modification, and termination of multimedia sessions. SIP infrastructure can be used as the basis for creating peer-to-peer (P2P) applications running on mobile devices.

P2P nodes (e.g. Duda and Miguel from Figure 1) use SIP message to communicate and they include REGISTER to make a registration to the server, INVITE to make an invitation for a call session, BYE to end a conversation, ACK to acknowledge a BYE or INVITE messages. There are other types of SIP messages in the specification but are beyond the purpose of this paper.

As this architecture uses a web service for the SIP server, the server only responds to users requests. In figure 1, Duda wants to make a call to Miguel.

First it contacts the Server telling he wants to talk to Miguel and needs the actual address. The Server responds with Miguel's public IP address. Duda receives the address starts an invitation to Miguel which once upon receipt of the INVITE message sends back a RINGING message to inform Duda that he received the invitation and eventually shall answer the call. When Miguel answers the call sends back an ACK message and the session is established. When the conversation ends, the nodes exchange a BYE/OK message. During call session, Duda and Miguel exchange RTP packets.

### 2.4 The RTP protocol

RTP defines a standardized packet format for delivering audio and video over the Internet. The RTP header is responsible for keeping information about packet sequence number, the type of CODEC and a timestamp to allow the receiver to order the packets and to calculate network latency.

### 2.5 CODEC

CODEC is a technical name for Coder/Decoder or Compress/Decompress. The Codec used is G.711 [4] and allow us to compress an audio data block up to a half of the original size. There are two main algorithms defined in the standard for G.711. The μ-law algorithm that is used in North America and Japan and the A-law algorithm that is used in Europe and other countries.

This codec requires very low processing power and it needs a minimum of 64 kilobits per second (Kbps) for one-way communication. It's suitable for Pocket Pcs. There are other codec’s that require more processing power but do better compression (e.g. G.723.1 with 5.3Kbps).

Choosing a voice codec is always a compromise. You have to choose between low bit rate versus low complexity or audio quality versus low bit rate etc.
2.6 Firewall / NAT

Firewall and NAT [5] (Network Address Translation) are important issues when a peer is behind a Firewall/NAT. The protocols used for this design does not handle the NAT Traversal very well.

When a user logs in, the Server determines the public side of the address used by the user’s NAT. this is the address given when the server receives an INVITE SIP message. The information about the user’s address is updated every time the users logs in or contacts the server about any other reason.

3 Quality of service

Quality of service is concerned about several issues regarding the problem of ensuring a smooth conversation between two peers. To determine the quality of service there are a few parameters to be aware of: Latency, Jitter, echo and Packet loss.

3.1 Latency

Services like VoIP are time sensitive. It’s very important to keep the delay within certain limits or the end result will be an intelligible voice packet for the end user. End to end conversations should not exceed 150 milliseconds [6].

3.2 Jitter

Jitter is defined as the standard deviation in the inter arrival time of packets. For example if a packet is expected to arrive at the destination at a predetermined time, but it does not due to a change in the transmission and the packet arrives a millisecond early or late, this would cause the voice stream to sound fragmented [6].

3.3 Echo

Echo is the effect of hearing the voice transmitted after a short period of time. It’s caused by some of the voice energy being reflected back to the caller. This can be very distracting and annoying in a normal conversation. An echo cancelation algorithm may be a solution to the problem.

3.4 Packet loss

Packet loss can occur for a variety of reasons, like high levels of congestion, link failure and depending on how many packets are lost it may cause degradation in voice quality.
4 The application

The Location Based Advanced Phone Dialer (LBAPD) provides three kinds of services:

- VoIP phone calls
- Voice messages
- VoiceTwitts messages

VoIP phone calls are made between users of LBAPD using wireless networks or GPRS. Voice packets are compressed using G.711 Codec. The call is set up in a peer-to-peer mode. The session is established sending and receiving RTP voice packets. Every ten seconds of call a synchronization signal is sent for delay adjustments.

Voice messages enable users to exchange voice message. Messages are sent to the web server with a destination user address so that the owner may collect it (Figure 3).

![Figure 3: Voice Message Processing.](image)

![Figure 4: VoiceTwitts Processing.](image)
We can see how it works through Figure 4. Duda records his voicetwitts message and sends it to the web server. The server sends the message address to http://twitter.com in HTTP format (http://voicemessageaddress.wav) so that users may click it in order to hear the message.

http://twitter.com it’s a website that enables users to post messages concerning one simple question, “what are you doing now?”.

5 Conclusions

This paper studied the possibility of implementing a VoIP system on a Pocket Pc with a low memory and low processing power.

The ability to record voice and make connection to Internet or Intranet using GPRS or using a wireless network enables a handheld device to VoIP. GPRS has a high latency for VoIP call but for voice messages it works well.

P2P SIP based telephony gives us the ability to build powerful services like exchanging voice messages and it would be interesting studying the possibility of exchanging also video.

6 References