Effect of WiFi systems on multimedia applications

ZOLTÁN GÁL, ANDREA KARSAI, PÉTER OROSZ

University of Debrecen, Center of Informatics zgal@unideb.hu, kandrea@fox.unideb.hu, oroszp@unideb.hu

Keywords: IPv4, IPv6, IEEE 802.11b/g, IEEE 802.11a, WiFi, L2/L3 roaming, VoIP, codec, QoS, H.323.

At WiFi hot-spot deployment an essential question is whether 802.11b or 802.11g and/or 802.11a system should be installed. For this decision an efficiency analysis is needed beyond the economic and rational considerations. As IP phone systems are increasingly available in university environment, the analysis of practicability of WiFi phones during movement appears as an obvious object in indoor and outdoor environment as well. In this paper we focus on the analysis of the properties of multimedia applications (video, streaming, IP phone) operated over IEEE 802.11b/g/a WiFi systems. (In: 2006/6, pp.15–23.)

1. Introduction

Wireless data transmission mechanisms that belong to the IEEE 802.11 family are spreading widely in indoor and outdoor environments as well due to their mobility feature. When deploying hotspots several considerations are required to be taken about the applied technology. This is more than just an economical or rational issue, it requires efficiency analysis as well. The WiFi system is based on the ISM (Industrial-Science-Medical) frequency range that allows service providers to deploy and operate multiple hotspots in the same physical area independently from each other. In outdoor environment the different service providers use radio channels usually in uncoordinates way. Since ETSI standards are applied to the radiated microwave power the densely deployed systems may cause interference with each other.

In company or academic environment network users set up more and more requirements toward mobile WiFi devices (such as laptops, palmtops, intelligent mobile phones) to support multimedia services. Since IP phone systems are dynamically spreading in academic environment, we need to analyze the usability of WiFi phones during physical movement in indoor and outdoor environment. In the 2.4GHz ISM range the voice transmission feature of the WiFi IP phone greatly depends on the applied voice encoding mechanism. The 5GHz WiFi transmission has a special channel coding mechanism that is more effective than that of the 802.11g. Its transmission rate is very sensitive to the distance between the access point and the client. During physical movement the high compression ratio data transmission standards are more sensitive to the radio cell roaming than the algorithms with lower compression ratio.

We know from previous analyses that the quality of multimedia services on mobile terminals is heavily affected by the velocity of the terminal during physical movement [1]. On mobile terminals the quality of multimedia applications depends strongly on data-link layer events.

2. Overview of multimedia coding/decoding algorithms

Spectacular development of DSP (Digital Signal Processing) architectures in the last few years and researches on human speech recognition affected the improvement of voice encoding and decoding technologies [2]. The new codecs beyond the usual AD/DA conversion apply internal patterns to analyze the input audio signal and forward it as a minimal bandwidth data stream.

PCM

The simple PCM (Pulse Code Modulation) audio is encoded according to the ITU-T G.711 standard [3]. The 64 kbps PCM voice is compressed by μ -law or Alaw procedures that convert the 12 or 13 bit sample to an 8 bit one using logarithmic scale. Benefits are simplicity, low complexity, low latency, good sound quality. Disadvantage is the high bandwidth requirement.

ADPCM

The ADPCM (Adaptive Differential Pulse Code Modulation) is also a common compression solution that is defined in the ITU-T G.726 standard. It uses 4 bit samples that are transmitted at 32 kbps. Unlike PCM these 4 bit words do not encode the amplitude of the voice rather the difference of amplitudes and the alteration rates. The algorithm applies a very simple linear estimate. Benefits are simplicity, low complexity, good sound quality, low latency, multiple coding rates. Disadvantages are relatively high bandwidth requirement, poor sound quality at lower bandwidths.

AMR-NB

The AMR-NB (Adaptive Multi Rate – Narrow Band) is mostly used on GSM and UMTS mobile networks. The algorithm supports eight compression ratios (4.75; 5.15; 5.90; 6.70; 7.40; 7.95; 10.20; 12.20 kbps). The algorithm can switch between these ratios at any time which is profitable on IP networks. The sender can alter the outgoing bandwidth at any time based on the real time statistics provided by RTP: according to the RTP signaling the encoder compresses the forthcoming samples at the newly changed rate and the decoder can decompress it in the same way. The 20 ms frames are encoded by ACELP algorithm using 5 ms *lookahead* value. Benefits are simplicity, relatively low complexity, low bandwidth requirement, good sound quality, low latency, multiple coding rates. Disadvantages are few implementations, lack of open source code.

AMR-WB

The AMR-WB mechanism is used by the G722.2 encoder that is optimized to wide bandwidth. It uses ACELP algorithm and encodes 7 kHz audio sampled at 16 kHz. It adaptively alters encoding rates (23.85; 23.05; 19.85; 18.25; 15.85; 14.25; 12.65; 8.85; 6.6 kbps). The encoder uses 20 ms frames and 5 ms *lookahead* buffer. Benefits are very good sound quality, low latency, multiple encoding rates. Disadvantages are high bandwidth requirement for good sound quality, moderate computing complexity.

RTP protocol

The RTP (Real-Time Protocol) provides point-to-point application layer transport service for real-time (audio, video) traffic, therefore it uses services like PDU identification, sequencing, time stamping and transmission control. It is used commonly over UDP using its multiplexing and checksum generation services and it is sometimes used over TCP as well. RTP cannot guarantee neither packet arrival nor correct packet sequence. RTP is optimized to a variable and overloaded network condition typical for IP networks. RTP transports content data to one direction while it uses the duplex channels of RTCP to relay control information that includes quality parameters as well. RTP can perform time recovering, sender identification, content identification, sequencing and loss detection. QoS, resource allocation, packet loss recovery and on-time delivery are not related to RTP.

Classification of voice encoders/decoders

PCM and ADPCM belong to the family of waveform codecs that use the redundant characteristics of the waveform. Compression techniques developed in the last 10-15 years are focusing on the voice source characteristics. These compression codecs create the simplified parameters of the original voice source that results in smaller bandwidth. They are called source codecs including LPC (Linear Predictive Coding), CELP (Code Excited Linear Prediction) and MP-MLQ (Multipurpose Multilevel Quantization) procedures. Advanced codecs substitute the human voice source with a mathematical model and they transmit the representation of the voice instead of the compressed voice. The most common telephone voice encoding and packet switched voice standards are the following:

- G.711: 64 kbps PCM voice encoding technique used in the conventional digital PBX centers and networks.
- G.726: Uses 40, 32, 24, 16 kbps ADPCM encoding.
 ADPCM voice transmission is recommended between packet switched and conventional PBX networks.
- G.728: Low latency fluctuation version of CELP that transmits voice at 16 kbps. CELP voice has to be trans-coded to public telephone format in order to set up communication with a public endpoint.
- G.729: By using CELP compression it converts the voice audio to an 8 bit data stream. Two subversions exist that significantly differ from each other in processing complexity both providing 32 kbps ADPCM quality voice.
- G.731: Compresses voice or multimedia audio consuming very low bandwidth. As part of the H.324 protocol family it operates at 5.3 kbps and 6.3 kbps. The former applies CELP, the latter uses MP-MLQ technology while both provide good voice quality and further flexibility for the system.
- GSM: The GSM (Global System for Mobile Communication) standard of ETSI I-30036 is widely used in European mobile networks for voice and low bandwidth data communication. Full rate GSM operates at 13 kbps by using RPE (Regular Pulse Excited) encoder at 8 kHz sampling rate. Half rate GSM requires 7 kbps bandwidth at 5 kHz sampling rate. The input voice is split up into 20 ms frames it makes eight short term approximations for each of them. Furthermore each frame is divided into 5 ms subframes where the encoder calculates latency and gain for the long term approximator. Finally it guantizes the rest of the signal in each sub-frame. The GSM encoder generates good quality voice however the G.728 (CELP) encoder outperforms it by its higher bandwidth. GSM encoder requires low processing time. Benefits: simplicity, relatively low complexity, low bandwidth, low latency, open source. Disadvantages: G.728 outperforms it in guality/bandwidth ratio.

Nullsoft Video protocol

The Nullsoft Video (NSV) format is a bit-stream that is able to provide joint packaging for audio and video data [4]. It complies with all major video and audio compression mechanisms. As it is a bit-stream format, it does not need to download the entire data file before playing. It can provide streaming service as reliable synchronization occurs at any point of the stream. Secondary data channels may provide multiple soundsubtitle- or data flows.

NSV file consists of two major parts: optional header and mandatory bit-stream. All multi-bytes integers are stored in LSB format where the least valuable byte is the leftmost byte. Therefore a 4 bit and a 20 bit number will be stored in 3 bytes long. The voice and video data packet are transmitted in one frame. The voice may follow or precede the video data. The number of channels for additional information (e.g. title, 16:9/4:3 screen ratio, secondary audio channel, etc.) is limited to 15.

3. Characteristics of the VoIP network

After compressing the voice and converting it to data the stream will be transmitted over the IP network using RTP protocol. In VoIP networks we must consider the latency and the bandwidth as well. Bandwidth requirements are critical as they depend on not just the selected codec but the overhead of the layer protocols (IP, UDP) [5]. The latency is affected by the propagation speed of the signal the buffer handling mechanism of the sender and receiver as well as the encapsulation delay.

Bandwidth requirements on VoIP network

The operation of voice conversation over IP networks is affected by several parameters. Bandwidth requirement of the applied codec can be in the range of 3...64 kbps. The voice PDU usually is not longer than 20 bytes while the L2 (Ethernet) and L3 (IP) layers add significant overhead. Therefore the effective bandwidth requirement is affected by these overheads [6]. In order to simplify the problem various solutions had been introduced. By using Voice Activity Detection (VOD) the sender can interrupt the stream if the signal of the local analogue source decreases to a certain threshold level. The bandwidth requirement is decreased to its half as in human conversations a person listens to the other in half of the times. This solution claims more attention in determining the appropriate moments for switching on and off otherwise it can cause content loss. However the total silence can be disturbing as well. Comfort noise is usually applied to eliminate this problem that is perceptible at the pair of the non-speaking person as locally generated white noise.

Advanced systems reproduce remote background noise in the silence period of the remote person. Another solution is the compression of the RTP PDU header. As the RTP PDU may contain duplicated or redundant information routers alongside the route compress the header therefore the bandwidth requirement of voice can be significantly decreased. In most common LAN/MAN technological environment the required physical bandwidth is shown in *Table 1*. IP/UDP/RTP generates 40 bytes and Ethernet makes 14 bytes overhead. Each voice connection means two call streams while video connection creates four or six call streams simultaneously.

Latency on VoIP network

In VoIP design a generally accepted rule is to keep the endpoint to endpoint latency under 150 ms. Transmission latency of today's media is not perceptive for human ear however together with handling latency it may cause perceptive distortion. On user side the latency tolerance is about 250 ms. Voice stream with higher latency cause interference with the natural voice stream therefore they may quench each other. Handling latency has effect on conventional line-switched telephone networks also but its importance is higher at packet switched transmission due to buffering. It should be kept under 150-200 ms at latency design. Latency of the G.729 standard is about 20 ms what was designed in consideration of future demands as well. A VoIP product generates a frame in every 10 ms on average then it orders them in pairs and puts into a packet therefore the value of latency will be 20 ms. In packet switched networks latency is generated by putting the current packet into the outgoing queue and the latency of the queue. These values are device-dependents and typically not exceed 30 ms.

VoIP applications are sensitive not only to latency but to its alteration as well. Unlike the conventional telephone networks in packet switched transmission the value of latency can fluctuate by the network traffic load. Jitter is a short term alteration of the latency that is the fluctuation between the expected and real packet arrival time. Devices compensate it by using *playout* buffer therefore gaps in the voice stream can be avoided. It increases the total latency of the system. The buffer can be fix-sized or adaptive at different devices. In case of VoIP jitter is the most significant quality obstructive parameter. Packet switched voice transmission usually passes through systems with different latencies and transfer parameters what results in poor quality. General feature of these applications is a large size receiver buffer with at least 1 second voice buffering.

QoS on VoIP network

In packet switched networks voice quality is largely determined by the latency and the jitter of the network therefore QoS parameters have high priority in network design. Further important issues are the segregation of the voice traffic from other network data and the protection of critical data against the possible high bandwidth of the voice traffic. Elements of the effective QoS design are: required bandwidth, packet loss, latency

Table 1. Comparison of the VoIP/channel bandw	vidth
---	-------

Algorithm	Voice bandwidth [kbps]	Codec latency [ms]	Voice PDU size [B]	Voice rate [PDU/sec]	L2 PDU size [B]	Physical bandwidth [kbps]
G.729	8.0	15.0	20	50	74	29.60
G.711	64.0	1.5	160	50	214	85.60
G.723.1	6.3	37.5	30	26	84	17.47
G.723.2	5.3	37.5	30	22	84	14.78

and jitter. These factors are ensured by the following techniques:

- Control strategy: traffic limitation that means packet drops when the traffic between two networking devices exceeds a given threshold. This parameter can be configured at input and output side as well. Typical example of such techniques is RED (Random Early Detection) and WRED (Weighted RED). These techniques identify packets that can be dropped when necessary.

- *Traffic design:* provides equal input and output packet rate for buffering. Unlike control strategy it tries to avoid packet drop whereas it increases the latency and jitter caused by buffering.

- *Call set-up control:* controls rejection of the bandwidth requirement of the applications. In VoIP networks RSVP (Resource Reservation Protocol) can be applied to reserve bandwidth for a call. The H.323 gatekeeper can limit this reserved bandwidth.

- *Queues/scheduling:* used at buffering by discovering the priority of the packets. One buffer can be configured for latency-sensitive packets and another for data packets. IP/RTP priority queue is common mechanism in VoIP.

- *Tagging:* There are different techniques to tag packets that require special treatment. In VoIP networks packets can be tagged by IP preference bits (IP header ToS field) for example. The packet tagging mechanisms are important to preserve the internetwork QoS parameters of the packets.

- Fragmentation: Further fragmentation of large packets can be enabled on some devices before transmit-

ting on a low bandwidth link. This feature protects voice packets against high latency required for the transmission of large packets. Therefore the voice packet can be put among the fragments of a large data packet.

4. Measuring environment and measured values

For our measurements we applied wireless devices (access points and mobile terminal) that support IEEE 802.11b/g and IEEE 802.11a transmission mechanisms as well. We analyzed the effects of L2 roaming that was occurred during on foot physical movement of the mobile terminal in our indoor test network as shown on Figure 1. Mobile terminal passed at 5-6 km/h (1.4-1.7 m/s) velocity parallel with the straight line that links the two access points. Within one measurement period (TS_i) the mobile terminal gets from the micro cell of AP₁ to the micro cell of AP₂ then backward it reaches to the cell of AP1 again. As a multimedia service video streaming and IP phone were run on the mobile terminal. MT was a laptop and WinAmp multimedia application was run for streaming (TCP) and SoftPhone application for voice conversation (UDP) analyses.

For TCP traffic the wired node was a streaming server where we downloaded Nullsoft (NSV) format multimedia contents from with different bandwidth. In order to measure UDP traffic SoftPhone application was run on the mobile terminal and the wired node as well and voice conversation was generated between them. Voice en-

Figure 1. The measurement environment





coding mechanism was selected on the Phone Center located inside the wired LAN. Bandwidth values of the streaming contents were the following: 80, 150, 300, 500 kbps. The applied voice encoding mechanisms were: G.728 (16 kbps), GSM (29 kbps), G.711 (80 kbps), Wideband (272 kbps). We set the *data retry* parameter of the access points to fixed 32 for TCP and UDP while we adjusted the *beacon period* between 20, 50 and 100 ms. Receiver buffer of the WinAmp application was fixed to 1000 ms. We performed twelve measurements for TCP, the same amount for UDP with each IEEE 802.11 standard, creating in this way seventy-two captured data files.

The two access points (AP₁, AP₂) were connected to the same VLAN on a L2 switch. Ethernet frames from the VLAN appear on the wired network were passed to a traffic capture node by mirroring to a dedicated physical port that used tcpdump to capture and store the data in libcap format file. Afterwards we analyzed the measured values with Ethereal v0.10.4 protocol analyzer therefore we can determine the time intervals that affect the quality of the applications. Radiated microwave power of the access points were set to 5 mW for IEEE 802.11b/g and 11dB for IEEE 802.11a. Physical distance between the two APs was 50 meters, traffic of the MT was unencrypted and association type was open. In each TS_i measurement (i = 1, 2, ..., 72) the MT started from point S then passed on B, S, A points and finally reached S again.

In order to determine the times that affect the quality of applications we identified the T_0 moment (*Figure 2*) in every capture data file. This is the arrival moment of the LIP (Last Important Packet) received by the wired node before L2 roaming. This is the last content data of the MT before roaming. Descriptions of LIP and FIP packets are shown in *Table 2*.

L2 roaming event occurs within the Tr time as we described in our previous paper [1]. Tr time is identified by the arrival of the roaming frames to the new AP. On the mobile terminal IPv6 was also used that can perceive the recovery of the data link layer of the protocol stack and it immediately initiates the discovery of the neighbor nodes. Tmt time denotes the relative moment when the MT can restart LLC transmission. We used this feature of IPv6 to identify the frames sent during accurate roaming event as we can experience multiple cell changes during the physical movement within indoor points S->B->S->A->S due to the Rayley fading effect. The Ts time is the operational latency of the multimedia connection. This is perceptible directly by the user and its high value may produce gaps in service and loss of the connection. FIP packets are used to identify the Ts value.

5. Analysis and explanation of measurement results

Comparison and analysis of the measurement results gave us the possibility to draw important conclusions. Different IEEE 802.11 standards show different behaviors at roaming events in indoor environment [7]. The roaming process is greatly depends on the beacon period (Tb) value which is a configuration parameter of the access point [8]. MT learns the beacon period of the AP [9] from a signal in the beacon. As soon as the MT does not receive eight consecutive beacons roaming event will be initiated [1]. By continuously monitoring the incoming beacon frames the MT perceives the loss of radio signal quality and initiates a roaming process.

Measurement results for TCP traffic are shown in *Figures 3, 4, 5* while UDP traffic values are displayed in *Figures 6, 7, 8.*

• When the beacon period was decreased step by step from 100 ms down to 50 ms and 20 ms the MT perceived more rapidly the alteration of S/N ratio, therefore it became more sensitive to the change of environmental conditions during the physical movement. In this way the roaming times rise up from 0.5-2.8 s to 0.1-14.9 s and then fall down to 1.5-2.9 s at each streaming technology. At first TCP dropout rises up from 2.5-17 s to 2.4-19.8 s than falls down to 1.8-7.9 s. Consequently the Tb=20 ms beacon period is better then the 100 ms value. It is a useful establishment about the beacon configuration. However very low Tb values may cause multiple roaming events in indoor environment due to the multipath signal propagation that induces repeated TCP dropout.

• In the case of IP phone communication, for a given voice coding technique the adjustment of beacon period from 100 ms to 50 and 20 ms involves the decrease of the cell change time from 0.1-44.5 sec to 0.1-12.5 sec while UDP dropout decreases from 0.2-49.8 sec to 0.2-19.9 sec that indicates the advantage of Tb=20 ms value.

• In indoor environment the IEEE 802.11 technologies show different behaviors regarding the beacon period. During streaming the IEEE 802.11a resumes the connection within a longer period of time. Regarding roaming performance it is followed by the IEEE 802.11b and the 802.11b that has the most advantageous features in case of indoor roaming.

• IEEE 802.11a produces very high latency therefore connection is dropped even in wide bandwidth voice communication while IEEE 802.11g shows the best reaction time where the connection dropout can be kept under 4 sec. This dropout is affordable be-

Table 2. Meaning	of	the	significant	packets
------------------	----	-----	-------------	---------

Transport layer	Important packet	Description
TCP	LIP	Last ACK packet (60 bytes) from MT sent to the server before L2 roaming
TCP	FIP	First ACK packet (60 bytes) sent to the server after L2 roaming
UDP	LIP	Last UDP packet (60 bytes) from MT sent to the wired node before L2 roaming
UDP	FIP	First UDP packet sent to the wired node after L2 roaming



Figure 3. The streaming behavior (Tb=20 msec)



Figure 4. The streaming behavior (Tb=50 msec)



Figure 5. The streaming behavior (Tb=100 msec)

cause it may be a relatively rare event in mobile IP phone systems if users are notified previously.

• Application dropout times have different characteristics depending on the bandwidth of the stream. NSV application at 150 kbps depends least on the radio technology at Tb=20 msec while it depends the most on the technology at Tb=50 ms. With Tb≤50 ms NSV programmes at 80 and 500 kbps are least affected by the radio technology.

• IP phone communication with Tb≤50 ms provides the least dropout by using GSM encoding that is optimized to mobile environment. GSM provides weaker sound quality than G.728 however it adapts more ef-



Figure 6. The IP Phone behavior (Tb=20 msec)



Figure 7. The IP Phone behavior (Tb=50 msec)



Figure 8. The IP Phone behavior (Tb=100 msec)

fectively to roaming situations. G.711 greatly depends on the roaming mechanism of the radio technology that comes from the features of the PCM designed for conventional wired environment.

• After roaming with high beacon period streaming resumes within a 4-11 sec period while with Tb \leq 50 the latency of the TCP connection is 3 sec that can be observed at the time difference Ts-Tr.

• At IEEE 802.11b/g latency of UDP traffic (IP phone) is under 0.5 sec for each voice encoding technique. At IEEE 802.11a the resuming of UDP traffic is significantly delayed (3-7 sec) after the MT arrives to the new radio cell.

• IEEE 802.11a technology performs its best at Tb = 50 ms. IEEE 802.11g provides the best data-link service for streaming at Tb=20 ms.

• For IP telephony IEEE 802.11g gives the best overall performance independently from the voice encoding mechanism and the beacon period. It is followed by the 802.11b and the 802.11a that shows the disadvantageous behavior.

• Flexibility order of the voice encoding mechanisms for data-link layer dropout (descending order): GSM, Wideband, G.711, G.728.

6. Conclusions

In this paper we analyzed the characteristic features of multimedia applications (streaming, IP phone) during indoor roaming events operated on IEEE 80211a, 802.11b and 802.11g transmission technologies. For TCP measurements NSV format streams were sent towards a laptop moving on foot in indoor environment. For UDP measurements IP SoftPhone were run on the laptop while it had a voice conversation with a wired IP phone. Measurements were performed for the common encoding mechanisms.

Based on these measurements we can establish that the beacon period of the access point has significant impact on the roaming process. Different IEEE 802.11 technologies show different behaviors for both TCP and UDP traffics. Beacon period also has effects on them.

For mobile indoor wireless streaming terminal IEEE 802.11g with <=50 ms beacon period provides the best performance. For mobile WiFi IP phones in slightly loaded 802.11g indoor environment good quality of service can be experienced. Roaming may cause 2-3 sec of dropout in best case that is an acceptable value for users if they are previously notified. GSM voice encoding flexibly adapts to dropouts in the data link layer, it is followed by the wideband encoding despite its higher bandwidth requirement compared to G.711 and G.728.

In the future transport layer services with low dropout (<250 ms) and fast roaming convergence are required to design in order to provide continuously acceptable operational quality for multimedia services in indoor mobile WiFi systems. Performance of the H.323/ SIP mobile IP video conference and web collaboration services in mobile environment have to be evaluated as well.

References

- Zoltán Gál, Andrea Karsai, Peter Orosz, "Evaluation of IPv6 Services in Mobile WiFi Environment", Selected Papers of Info-Communication-Technology, Volume LX., 2005, pp.47–54.
- [2] Dodd, Annabel Z., The Essential Guide to Telecommunications.
- [3] Newton, Harry, Newton's Telecom Dictionary, http://www.harrynewton.com/
- [4] Justin Frankel,
 "Nullsoft Video (NSV) Format 2.1 Specification", http://ultravox.aol.com/NSVFormat.rtf
- [5] Jonathan Davidson,
 "Voice over IP Fundamentals", http://www.ciscopress.com/
- [6] Cisco Documentation DVD Home Page: "Voice/Data Integration Technologies", http://www.cisco.com/univercd/cc/td/doc/cisintwk/ ito_doc/voicdata.htm
- [7] WFi Alliance: "Wi-Fi Certified for WMM- Support for Multimedia Applications with Quality of Service in Wi-Fi Networks", http://www.wi-fi.org/files/uploaded_files/ wp_1_WMM%20QoS%20In%20Wi-Fi_9-1-04.pdf
- [8] SpectraLink, Co.:
 "White paper: Deploying netlink wireless telephones, best practices", http://www.spectralink.com/files/literature/ NetLink_Best_Practices_122105_01.pdf
- [9] 3Com, Co.:

"White paper: Deploying 802.11 Wireless LANs", http://www.3com.com/other/pdfs/products/en_US/ wireless_lans_wp.pdf