Virtual Audio Chat: User Interaction and Audio Quality Evaluation

Lea Skorin-Kapov
Saška Mateša

Presentation Outline

- Introduction
- Multi-user *Virtual Audio Chat* (VAC): user interactions
- Audio quality evaluation: comparison of three coders
  - previously recorded speech and music
  - live audio chat
- Conclusions

Introduction

- Virtual Audio Chat: interactive VR/Web application that enables real-time audio streaming
- Speech quality and network requirements depend in great deal on type of coder being used
  - Different data rates and Mean Opinion Scores (MOS)
- Measurements of network throughput and subjective speech quality for three different coders: **PCM, GSM, G.723.1**

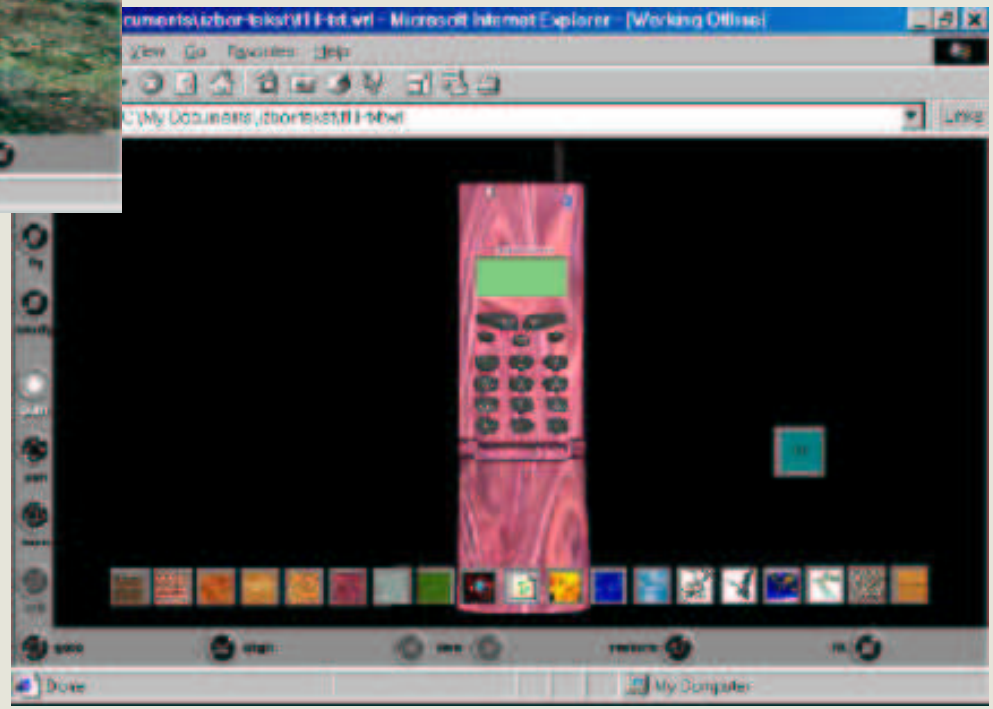
Multi-user Virtual Audio Chat

- Audio conference with (desktop) virtual reality interface

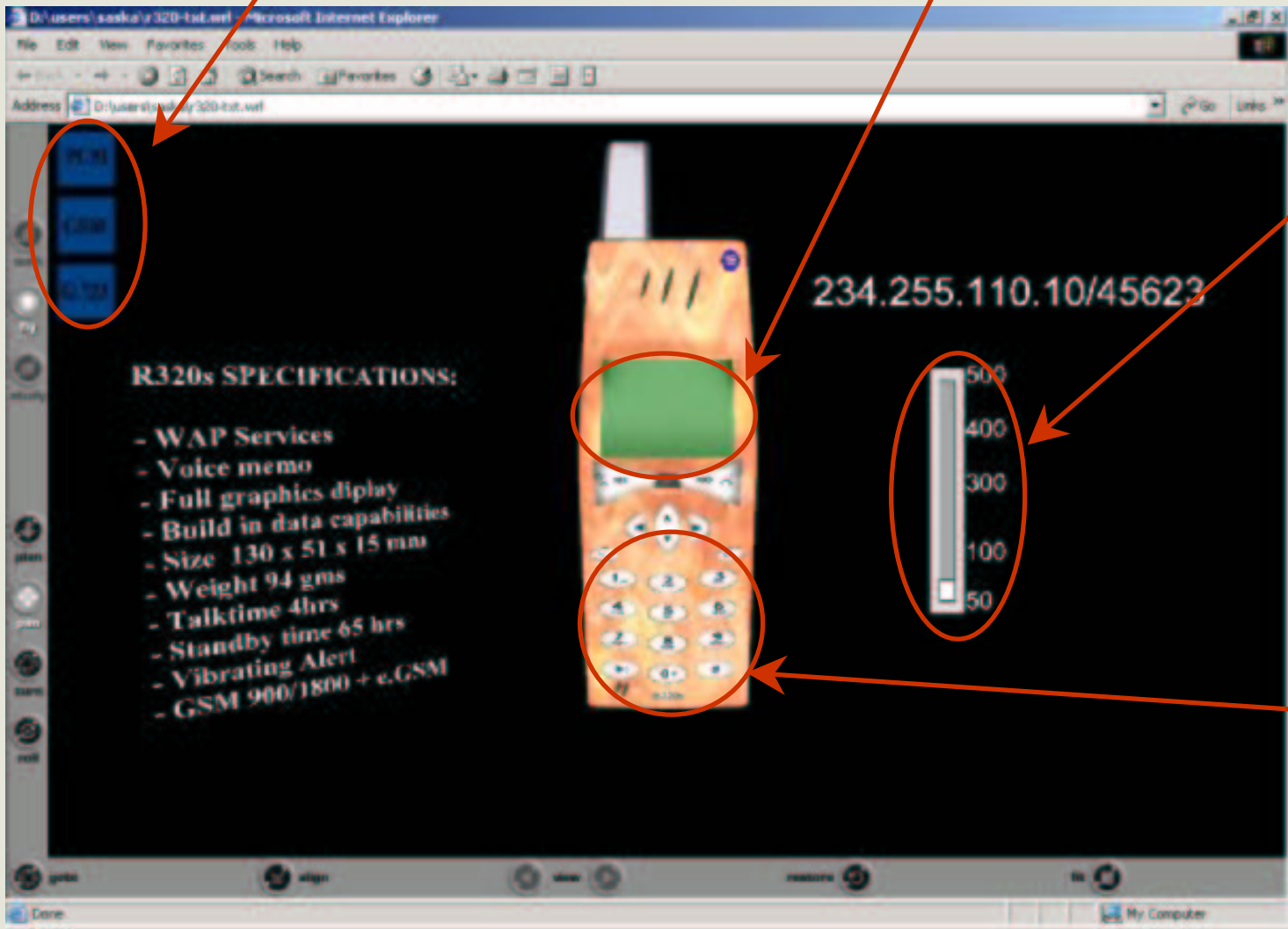
Three basic components:

- VRML mobile phone model - user interactions
- Java applet opening RTP based audio conference
  - Java Media Framework (JMF) API used
- Modified *Session Directory* (sdr) tool – enables user to schedule and announce multimedia session

User navigates through gallery and chooses mobile phone



User chooses which texture to apply


choice of coder
display
virtual
media-data
buffer
virtual keys
234.255.110.10/45623
R320s SPECIFICATIONS:
- WAP Services
- Voice memo
- Full graphics diplay
- Build in data capabilities
- Size 130 x 51 x 15 mm
- Weight 94 gms
- Talktime 4hrs
- Standby time 65 hrs
- Vibrating Alert
- GSM 900/1800 + e.GSM
500
400
300
100
50

Real-time data transfer using RTP/RTCP

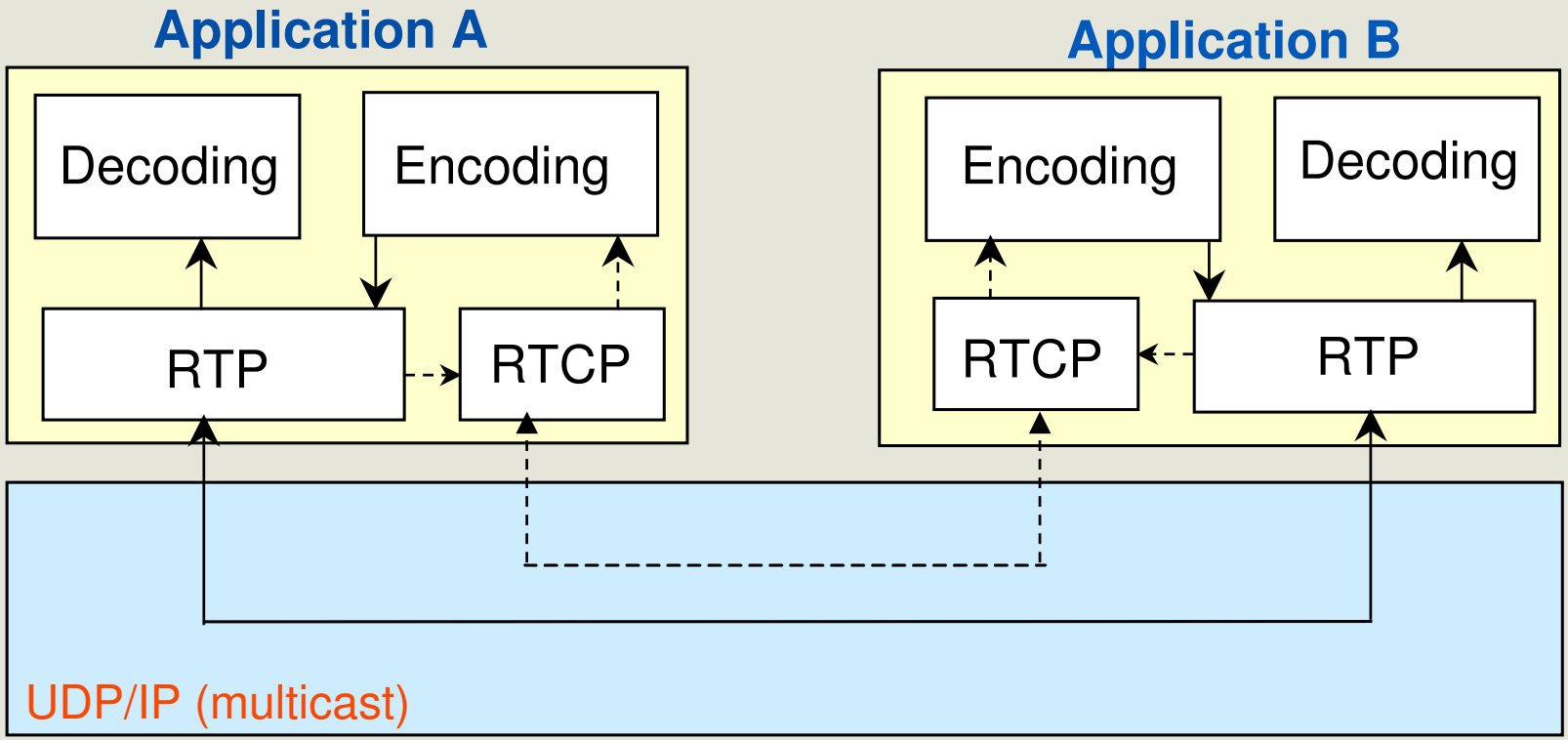
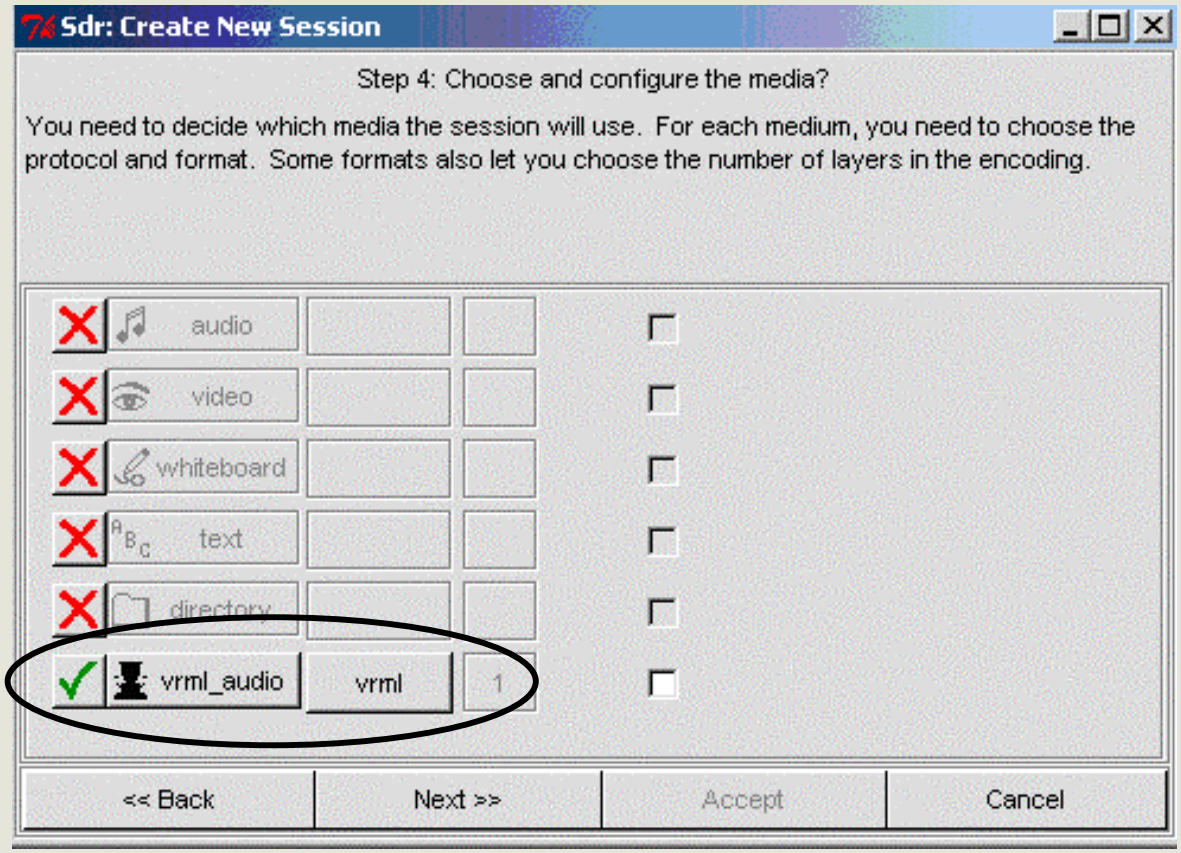


Figure adapted from “Internet Protocols for Multimedia Communications”, T. Braun , IEEE MultiMedia, 1997

Modified session directory (sdr) tool

- *Sdr* – scheduling/announcing multimedia sessions on the MBone

• Modified by enabling new media type: *vrml_audio*

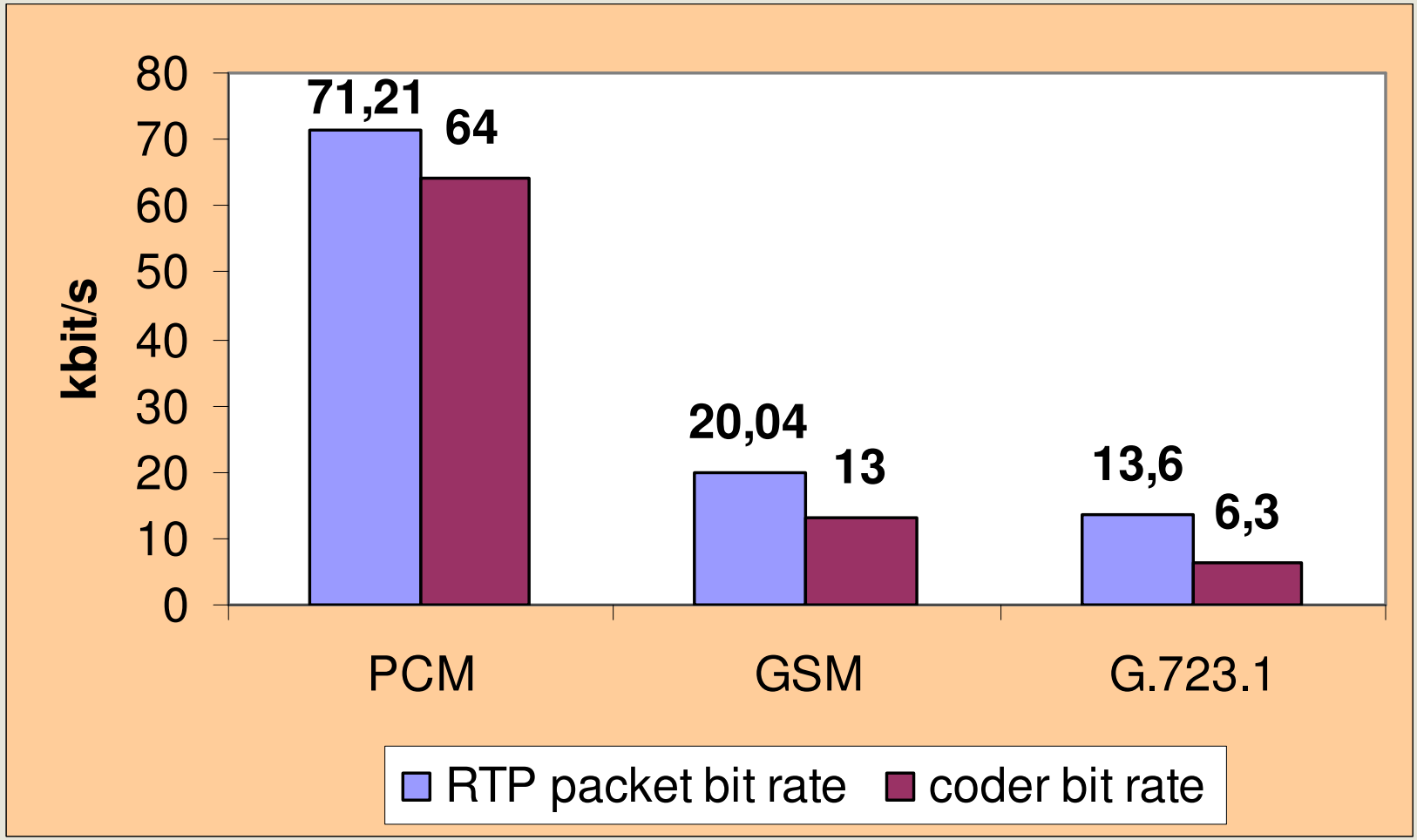


Audio quality evaluation

Measurements conducted for three different coders:

- PCM - ITU-T G.711 standard
  - simple waveform coder based on the encoding of voice samples
  - MOS score of 4.3, small algorithm delay
  - 64 kbit/s, not suitable for use on the Internet
- GSM - ETSI standard
  - vocoder - transmits parameters relating to model of source signal
  - MOS score of 3.5, 13 kbit/s
- G.723.1 - ITU-T G.723.1 standard
  - vocoder, MOS score of 3.8
  - 5.3 and 6.3 kbit/s: low bandwidth usage - suitable for Internet use

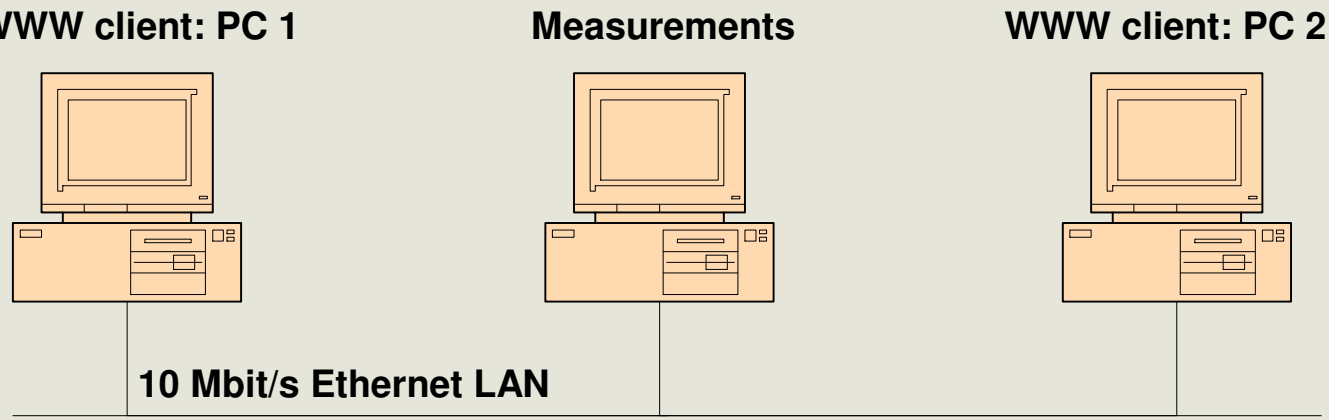
Test 1: Previously recorded speech and music



Average bit rates of RTP and RTCP packets:

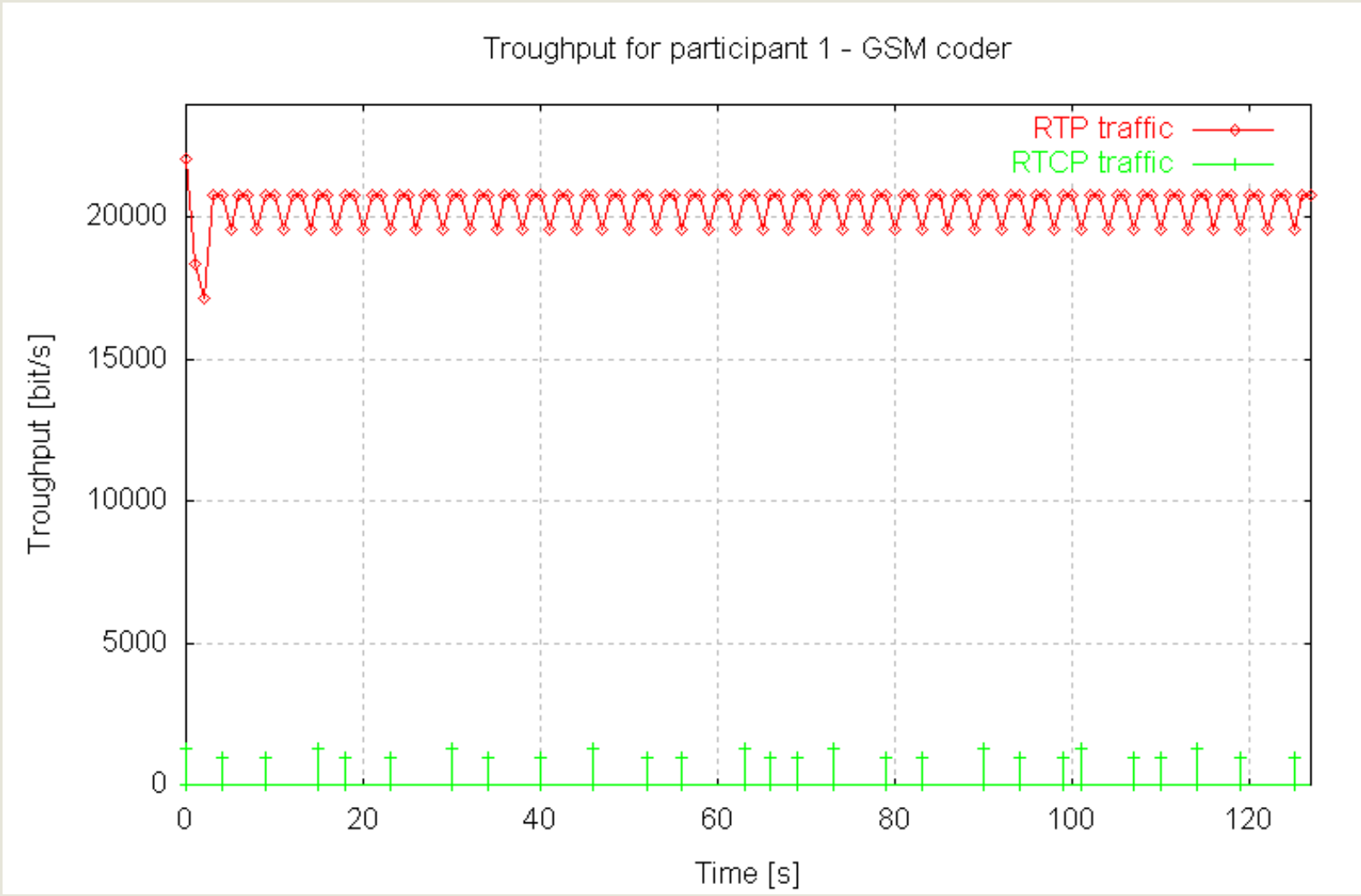
Coder	RTP packet bit rate [bit/s]	RTCP packet bit rate [bit/s]		
		male voice	female voice	music
PCM	72919.04	298.53	312.88	313.55
GSM	20520.96	282.16	291.39	279.58
G.723.1	13926.40	287.84	287.39	283.79

Test 2: Live audio chat



		average bit rate of RTP packets [bit/s]	average bit rate of RTCP packets [bit/s]
Scenario 1: PCM coder	Participant 1	71238.26	253.12
	Participant 2	71227.33	235.42
Scenario 2: GSM coder	Participant 1	20376.22	244.74
	Participant 2	20404.76	232.26
Scenario 3: G.723.1 coder	Participant 1	13602.75	269.08
	Participant 2	13601.80	232.40

Measurements of network throughput



Subjective measurements

- Subjective estimation of speech quality and comprehension of decoded signal

	male voice	female voice	music
PCM	good quality	good quality	good quality
G.723.1	fair quality (worse than PCM, approx. like GSM)	fair quality (worse than PCM, slightly better than GSM)	fair quality (worse than PCM, slightly better than GSM)
GSM	fair quality (worse than PCM, approx. like G.723.1)	fair quality (worse than PCM, little worse than G.723.1)	fair quality (worse than PCM, little worse than G.723.1)

- Decoded tones:
  - PCM - slightly higher
  - GSM and G.723.1 smother and deeper
- Music: poorer quality then decoded speech
- Quality of speech:
  - PCM - highest
  - GSM and G.723.1 – similar; poorer than PCM
- Comprehension good for all coders

Conclusions

- *VAC:* Web application with interactive VR interface
- *A*pplication enables measurements of:
  - network requirements
  - quality of service from the user side
- Performance analysis using three different coders:
  - lowest network requirements – G.723.1
  - best subjective quality – PCM
  - optimum – GSM or G.723.1.