

# Quantifying Skype User Satisfaction

Kuan-Ta Chen, Chun-Ying Huang, Polly Huang, and  
Chin-Laung Lei

*SIGCOMM'06*, Sep 2006, Pisa, Italy.

# Goals and Motivation

- How can we evaluate quality of service in real-time P2P multimedia services like Skype?

Traditional telephony QoS metrics:

- Mean Opinion Score (MOS): human surveys of sample audio clips
- Average Call Duration (ACD)

Decentralized P2P architecture, packet encryption, and dynamic network conditions make evaluation very difficult.

# Contributions

- Two QoS metrics:
  - *User Satisfaction Index (USI)*: based on call duration
  - *Interactivity measures*: based on voice activity
- Ciphertext voice activity detection
- Chokepoint P2P service tracing
- Statistical methodology

# Trace collection

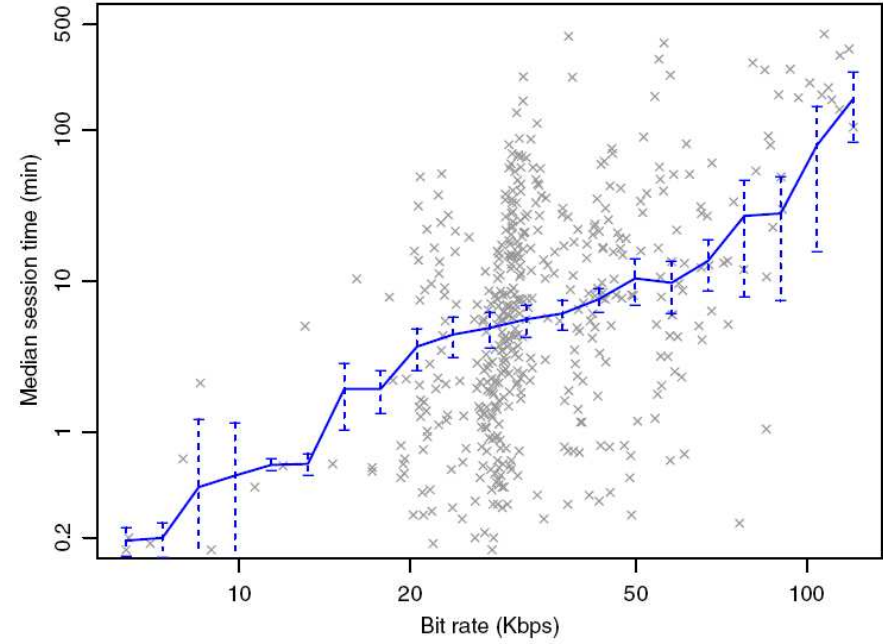
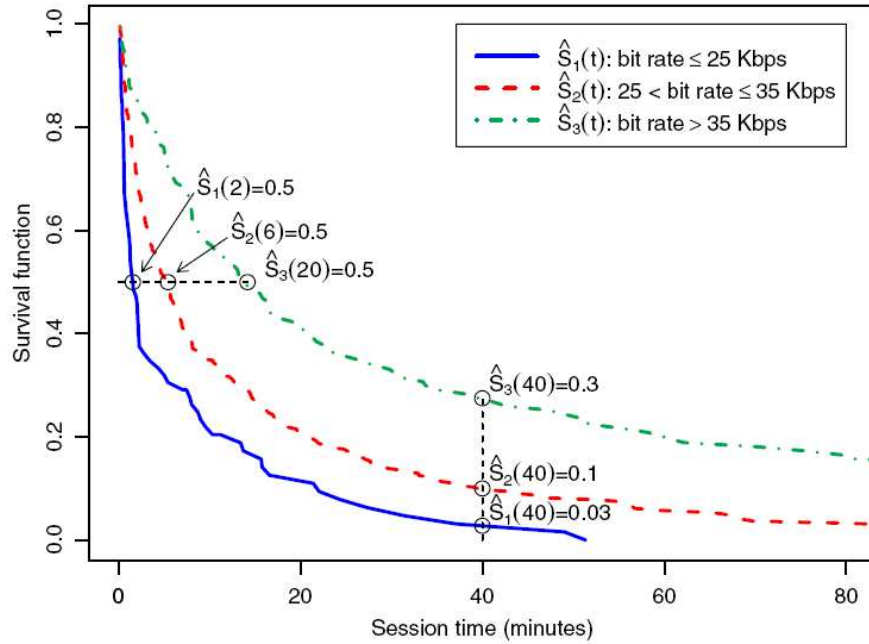
- Packet sniffer installed on campus's WAN uplink.
- Two stage packet filtering:
  - *online*: HTTP request to `ui.skype.com` indicates a Skype login and gives the host's Skype port. (captures 1-2% of packets)
  - *offline*: Skype packets are partitioned into sessions (calls). Orphan packets are discarded.
- Skype hosts are ping-ed for RTT (network latency) every 2 minutes, at 1Hz.
- 462 usable sessions recorded, 45% relayed

# Call duration as a QoS indicator

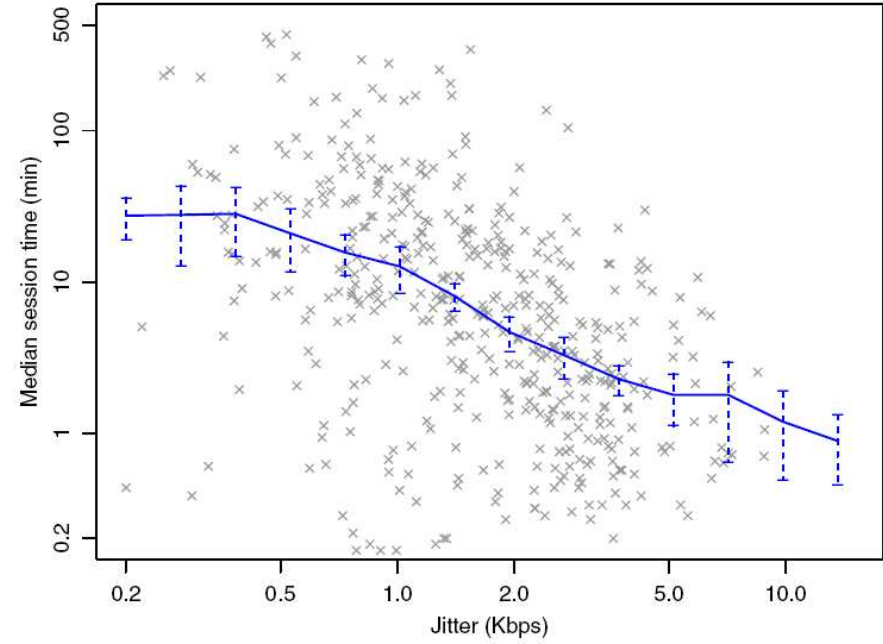
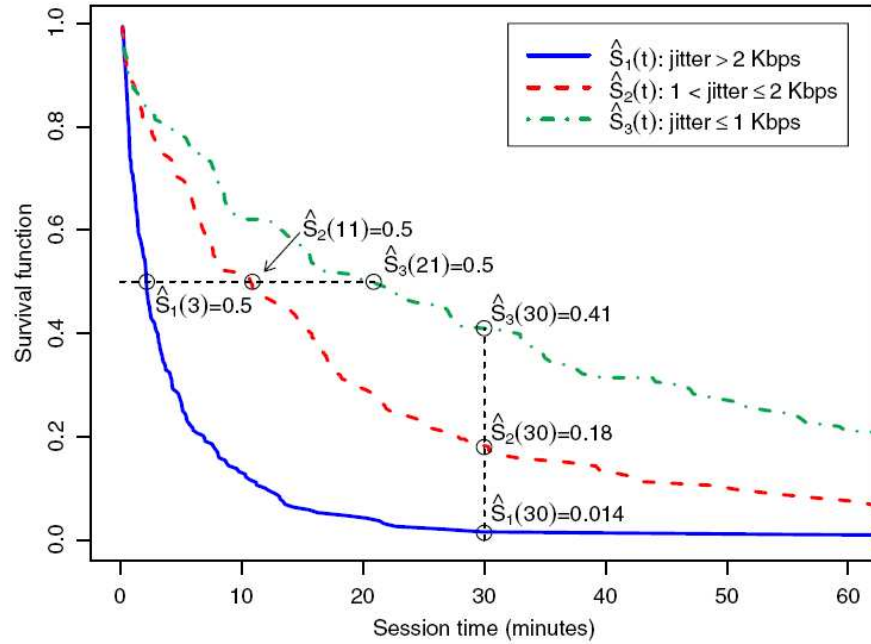
They use statistical analysis to determine the correlation to call duration of:

- Bitrate
- Network conditions:
  - Round Trip Time (RTT) – as measure by pings
  - Jitter ( $\delta(\textit{bitrate})$ ) – sample bitrate each sec
  - Packet loss – is factored into jitter estimate

# Results: bitrate correlation



# Results: jitter correlation



# User Satisfaction Index

Regression analysis is used to determine separate impact to USI of each factor:

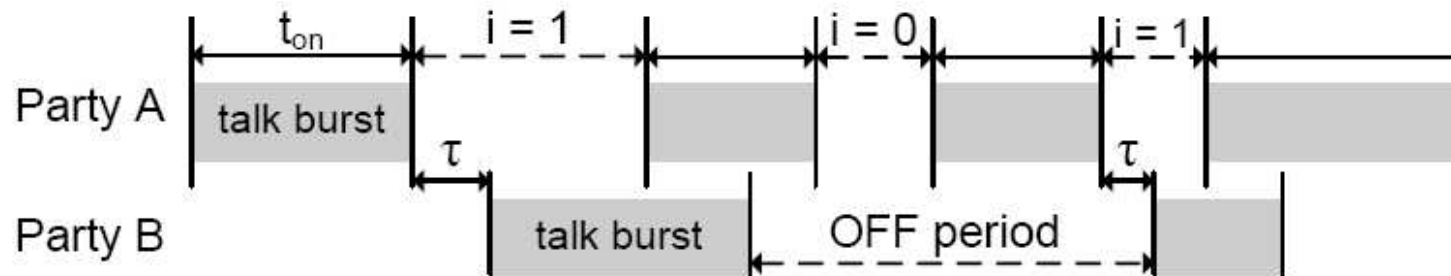
bitrate/jitter/RTT = 46% / 53% / 1%

$$USI = 2.15 * \log(\textit{bitrate}) - 1.55 * \log(\textit{jitter}) - 0.36 * RTT$$

Conclusions: improve user satisfaction by *increasing bitrate* and *decreasing jitter*, even at the expense of latency.

# Interactivity as a QoS indicator

- *Responsiveness* – users should alternate speaking
- *Response Delay* – short delay means better comprehension
- *Talk Burst Length* – poor quality leads to slower speech (limited 10 sec)



Index of Interactivity:  $\text{count}(i = 1) / \text{count}(i = 1 \text{ or } i = 0)$

Avg. Response Time:  $\text{mean}(\tau)$

Avg. Burst Length:  $\text{mean}(t_{on})$

So, we need a way to detect voice activity in the encrypted media stream...

# Skype's *two* voice codecs

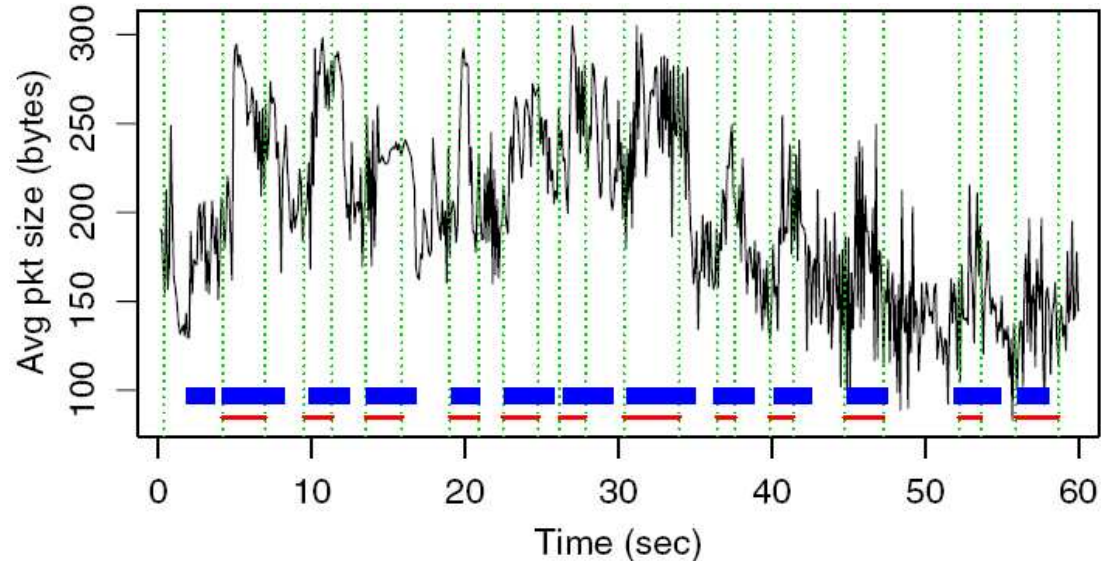
- *iSAC*: wideband codec, 10-32kbps Variable Bit Rate
- *iLBC*: 13.3kbps Constant Bit Rate
  
- Packet frequency (frame size) is relatively constant 30ms. This can be increased to reduce packet header overhead and thus bandwidth at the expense of latency.
- Bandwidth numbers are per-channel (direction)
- Packet header is 20 bytes IP + 8 bytes UDP => *7.5 kbps header overhead*
- What is the effect of AES Encryption on packet length?
- Traces showed 20-64 kbps (likely due to header and encryption, *not* higher bitrate codec)

# Bitrate analysis

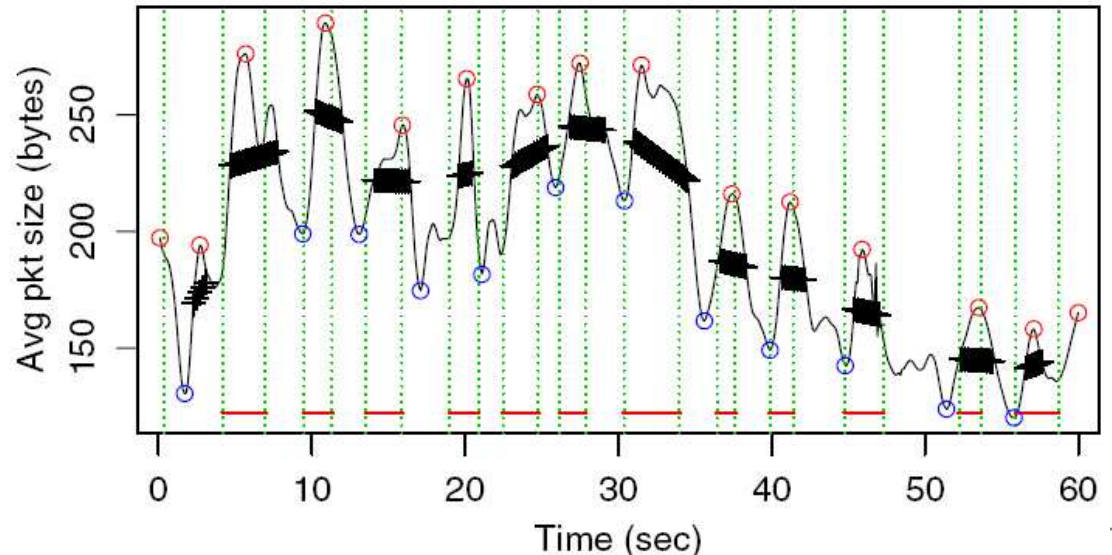
Bitrate factors:

- *Volume level*
- *microfluctuations* due to varied syllable entropy.
- *macrofluctuations* due to network conditions, CPU load.

Filter out variations below the 1 sec and above 5 sec scale.



(a) Original packet size process



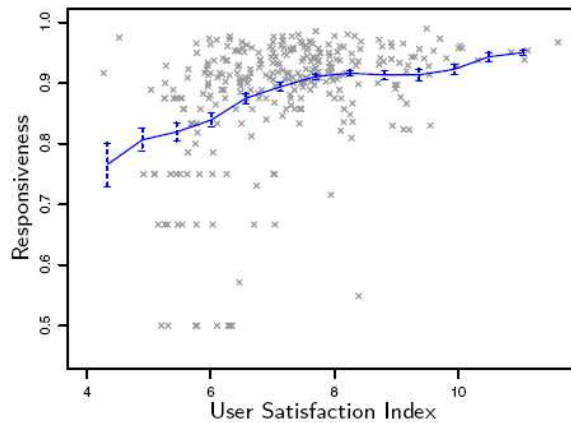
(b) Wavelet denoised process with estimated ON periods

# Interactivity results

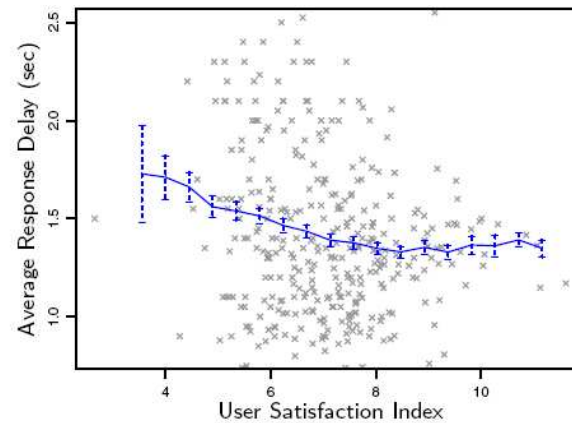
## Feasibility:

- Filtering yielded 73-92% accuracy in voice activity detected on a synthetic waveform and a voice sample.

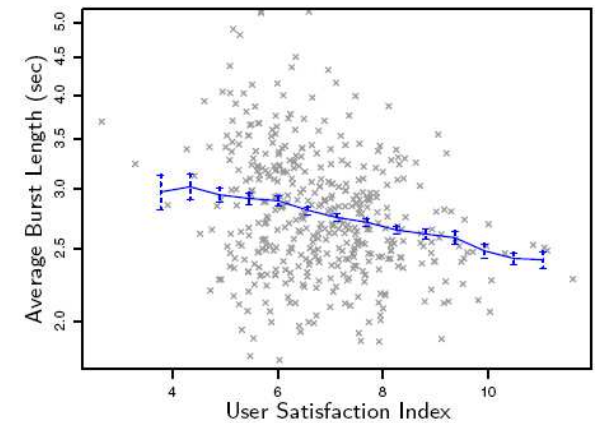
## Correlations with USI:



(a) Responsiveness vs. USI



(b) Average response delay vs. USI



(c) Average talk burst length vs. USI

# Future work

- Use better network instrumentation
- Other applications:
  - Traditional client-server VoIP
  - Other P2P apps
- Real-time integration for dynamic bitrate adjustment

# Problems

- Interactivity is not shown to be correlated with QoS (call duration) and intuition behind interactivity is weak.
- Experience shows that latency *is* important QoS factor.
- Network effects analysis is flawed:  $\Delta$  bitrate is inherent in VBR codec, not indication of jitter & packetloss.
- Generally, few variant data points in statistical analysis
- Better benchmark for USI is needed: perhaps MOS
- Ignored Skype's CBR iLBC codec