National Center for Scientific Research
N.C.S.R. 'Demokritos'
Athens, Greece

Statistical Analysis of Sign Language VideoConference Traffic in Multipoint IP Sessions

S. Kouremenos, D. Kouremenos,
S. Domoxoudis and A. Drigas

The Fourth Conference on Videoconference and Emerging Technologies
Gallaudet University - Washington D.C.
April 2004

VideoConference Traffic Modeling

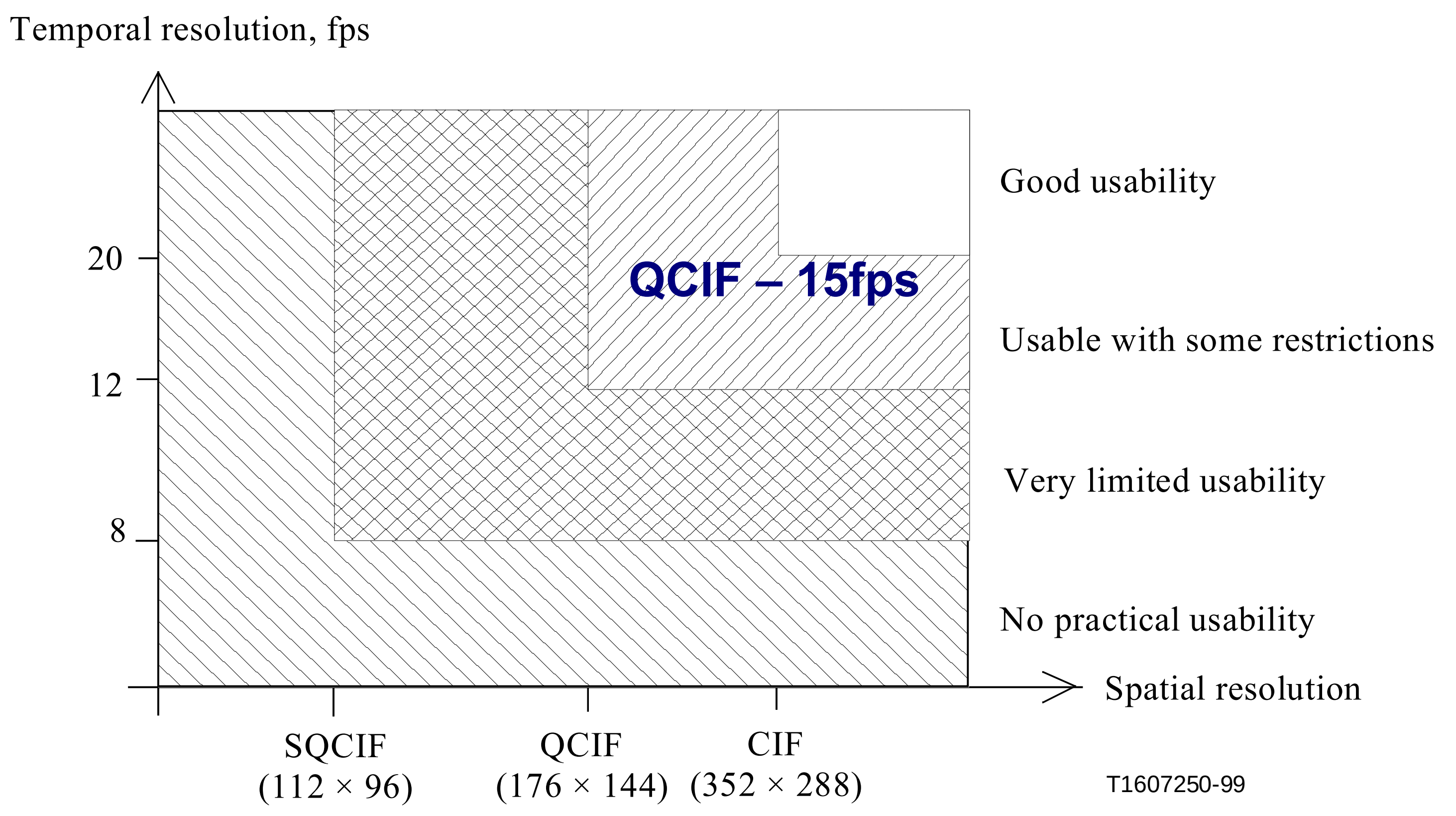
- A problem of major importance.
- Valuable insights about the resulting network load.
- A theoretical assessment of the network performance.
- Extensively studied in literature - Full Theoretical Models have been proposed
- Complex Procedure
  - Variation of videoconference session parameters
    - Number of participants
    - Target Video bit rate
    - Target Frame rate
  - Variation of videoconference content (H&S, Movies, Sign Language)
  - Different Versions and Implementations of Video Codecs (H.261, H.263, H.263+, H.264)

Why Sign Language VideoConference Traffic Modeling is a separate research thread?

- Increasing availability of affordable communication channels (ISDN, ADSL) for the end-users (signers).
- Readily available videoconferencing software (MS NetMeeting).
- Established video coding standards (such as H.261 and H263).
- Exchange of bandwidth demanding qualitative video information - minimum video bit rate of 384 kbs and frames per second are at least 15 reported although 30 is ideal.
- Videoconferencing traffic modeling research has been tested ONLY on head&shoulders or movies.

Sign Language VideoConference Requirements

Official application profile document of ITU

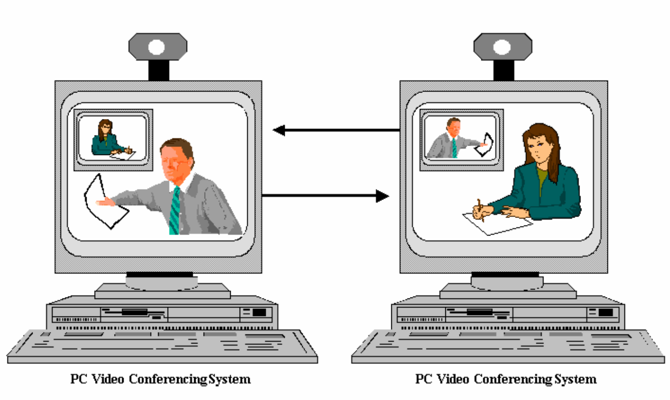


VideoConference Topologies

CLIENT TO CLIENT

(One-point Communication)

- More Flexible
- Less QoS capabilities

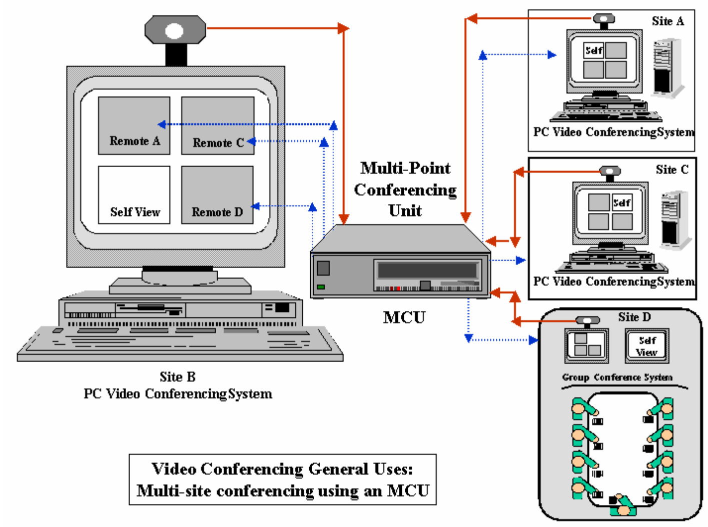


Video Conferencing General Uses:
One - to - One Meeting

CLIENT TO MCU

(Multipoint Communication)

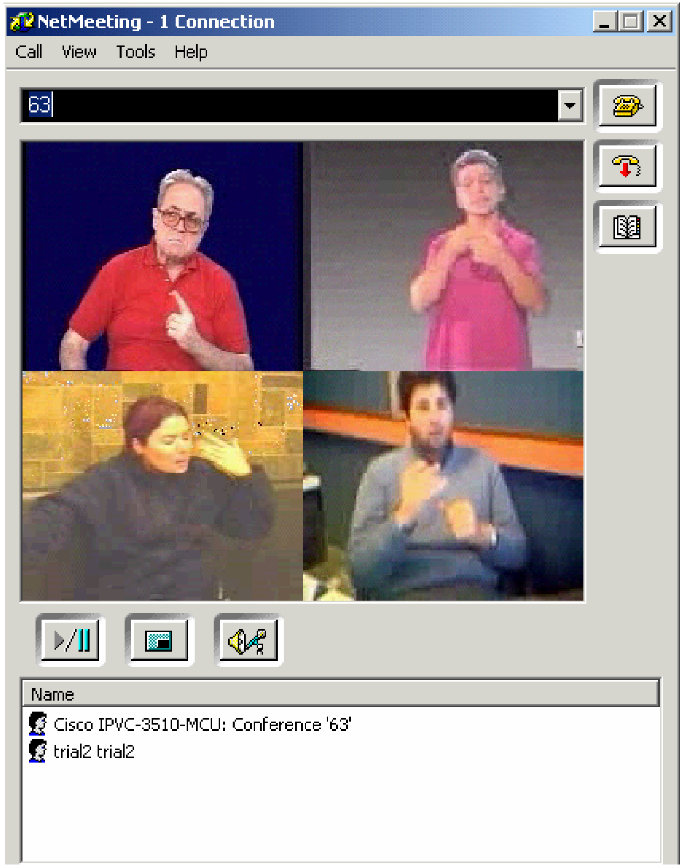
- Better Synchronization, Control and QoS
- Demand of large bandwidth for Continuous Presence



Video Conferencing General Uses:
Multi-site conferencing using an MCU

Experiments Description

- Multipoint - Continuous Presence Video Conference Sessions between Native Greek Signers
- CISCO MCU 3510 in High Quality Mode (CIF)
  - Target Video Bit Rate = 320KBits/sec
  - Target Frame Rate = 15fps
- MS NetMeeting (to ensure the direct usefulness and applicability of our results)
  - QCIF H.261 and H.263 encoded video – Best Quality
- 1h Duration

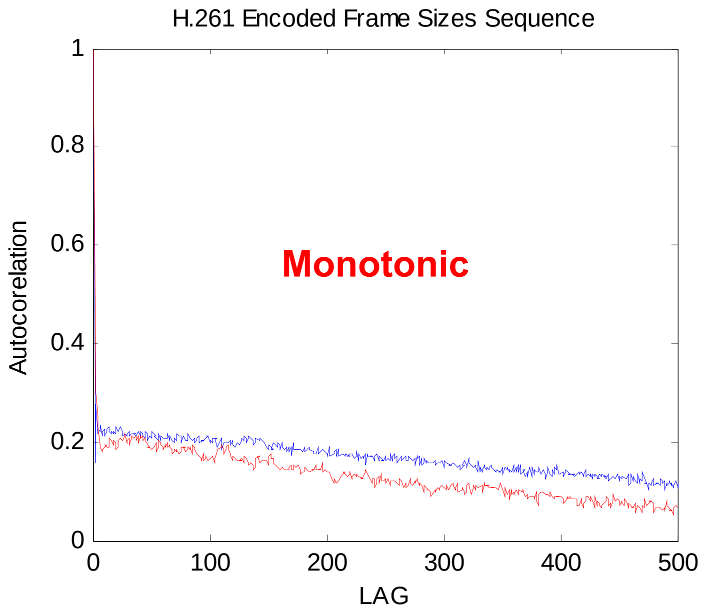


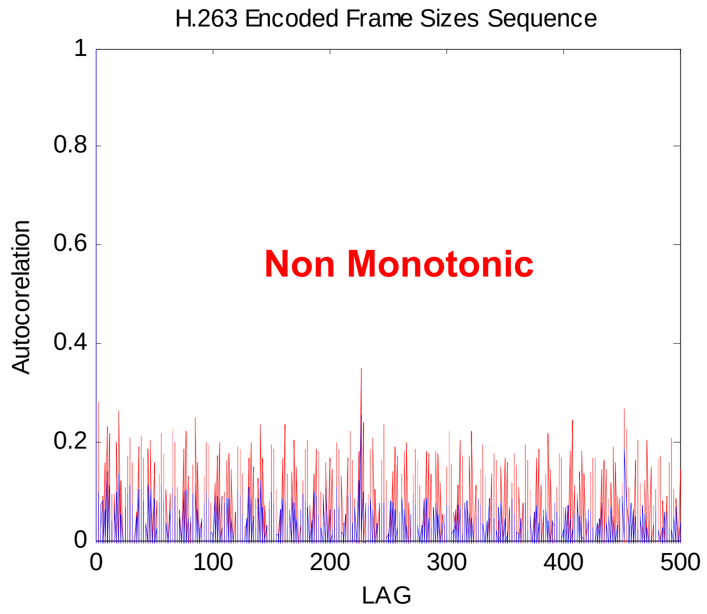
Experiments Quantities at the Frame level

Exp	1					2				
Terminal	VC1	VC2	VC3	VC4	MCU	VC1	VC2	VC3	VC4	MCU
Scenario	Continuous Presence - H261					Switched Presence - H263				
Target Video Bit Rate (For Terminals) (KBits/sec)	320					320				
Target Video Bit Rate (For the MCU) (KBits/sec)	1280					320				
Target Frame Rate (fps)	15					15				
Duration (sec)	1800					1800				
Video Bit Rate (Kbits/sec)	215.95	215.68	216.79	217.64	867.63	206.59	217.22	208.36	207.69	208.36
Frame Rate (fps)	7	6	8	9	9	15	7	15	15	15
Average Frame Size (Bytes)	3877	4371	3543	2941	11724	1729	3966	1727	1724	1727
Variance	309660	246830	175860	126850	10834000	87849	225250	47428	42892	47428

Sign Language VideoConference Traffic Analysis (1)

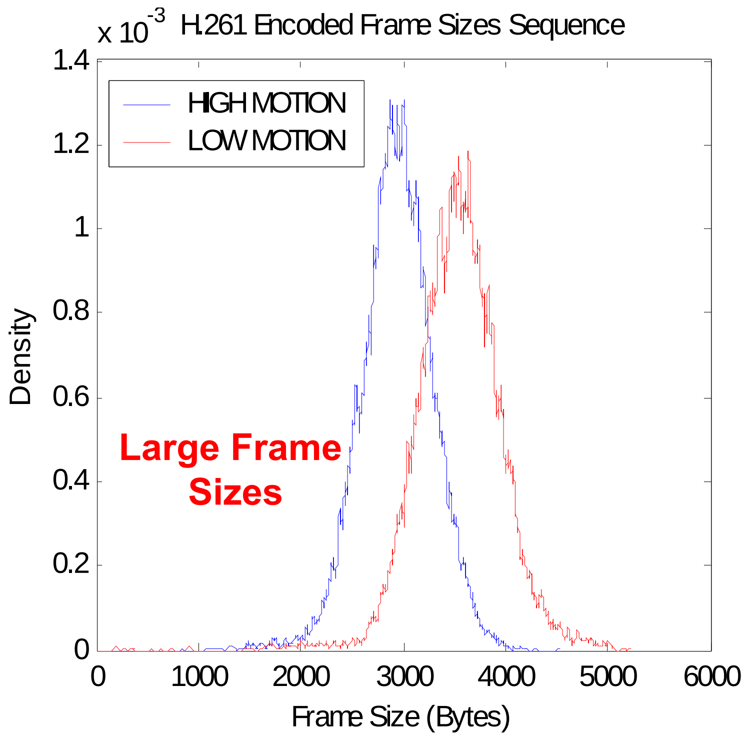
- **The frame sizes sequence is a Stationary Stohastic process with an AutoCorrelation Function Exponentially decaying and a Gamma-like Distribution.**

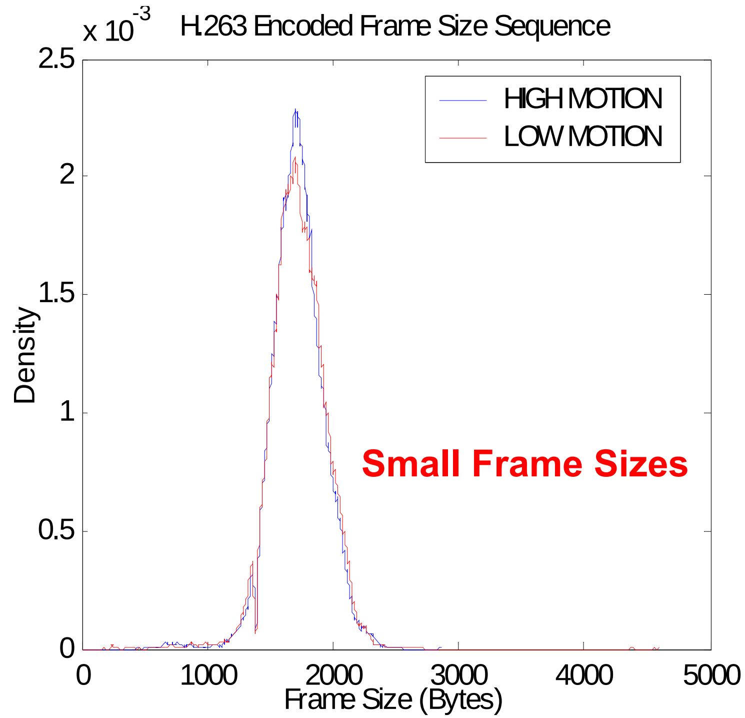




Sign Language VideoConference Traffic Analysis (2)

- The frame sizes sequence Distribution exhibits a Symmetrical Gamma-like Distribution similar in all cases

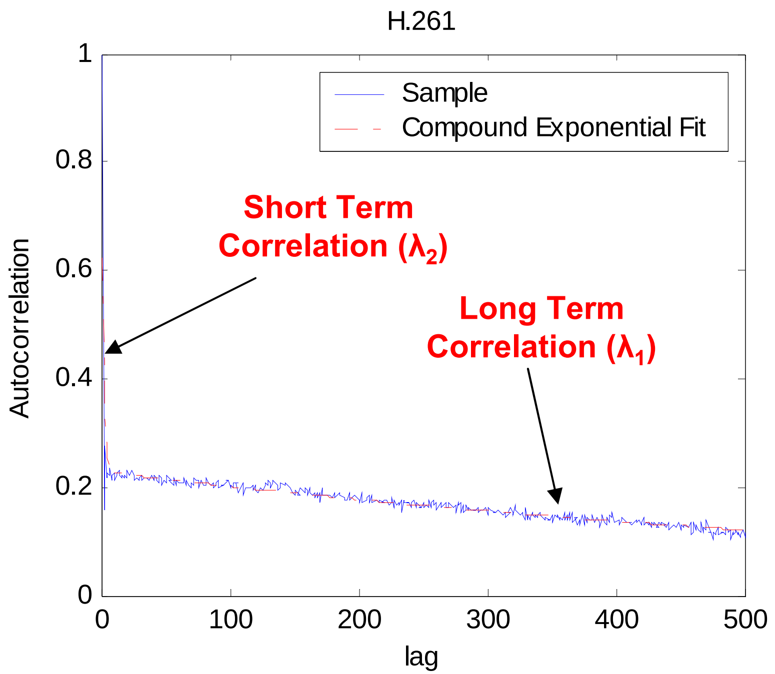


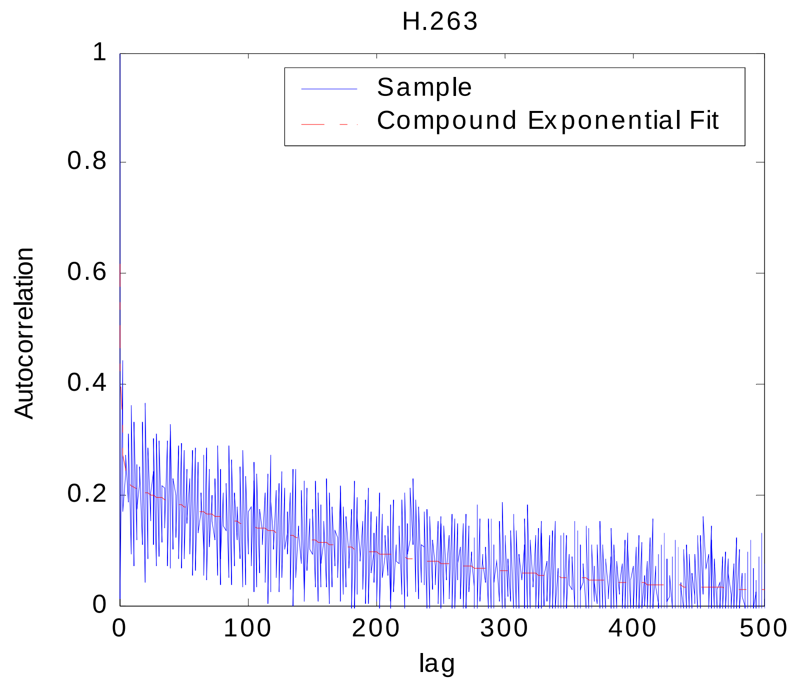


Traffic Modeling (AutoCorrelation Function)

Compound Exponential Fit

$\varrho_{\varkappa} = w\lambda_1^{\varkappa} + (1-w)\lambda_2^{\varkappa}$, with $|\lambda_2| < |\lambda_1| < 1$





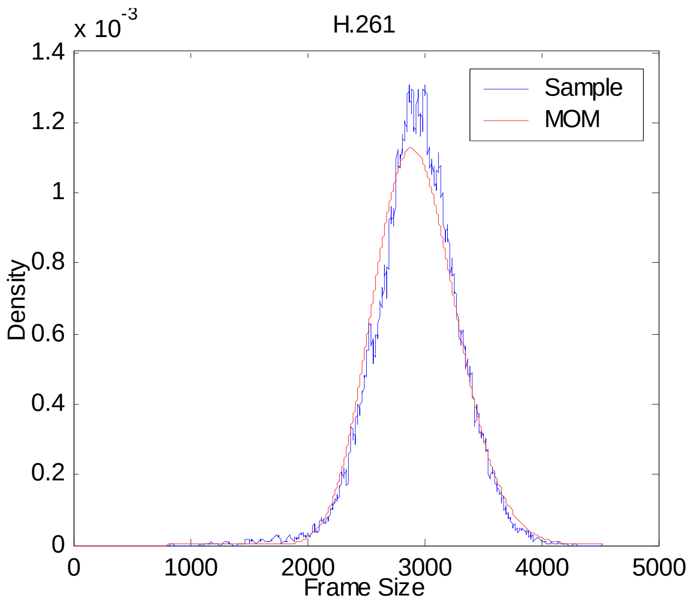
What matters is the λ1 parameter

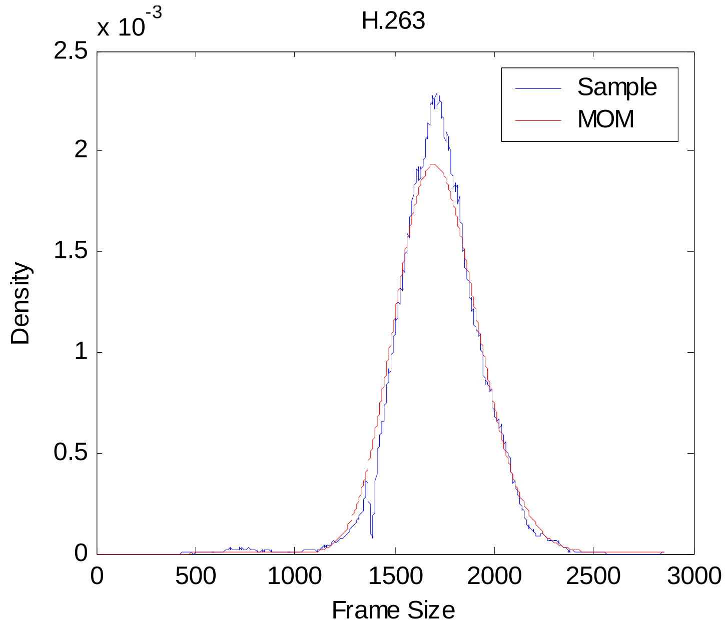
Traffic Modeling (Probability Density Function)

Gamma Density Function

$$f(x) = \frac{1}{\Gamma(p)} \frac{1}{\mu} \left(\frac{x}{\mu} \right)^{p-1} e^{-\frac{x}{\mu}} \qquad \mu, p > 0, x \geq 0 \qquad \Gamma(p) = \int_0^\infty u^{p-1} e^{-u} du$$

MOM Method: $p = \frac{m^2}{v}$ **and** $\mu = \frac{v}{m}$





Full Theoretical Models in Literature Markov Chain Models

- **DAR(1) Model (D.P. Heyman)**

$$P_{dar} = \rho I + (1-\rho)Q$$

  - **ρ is the autocorrelation coefficient at lag-1**
  - **Q is a rank-one stochastic matrix with all rows equal to the probabilities resulting from the negative binomial density corresponding to the Gamma fit for the frame size distribution**

- **C-DAR(1) Model (S. Xu, Z. Huang, and Y. Yao)**

$$P_{cdar} = f(P_{dar} - I)$$

$$f = \frac{\ln \rho}{(\rho - 1)} T$$

The continuous version of DAR(1), where T is the frame rate of the videoconference traffic

Our generalization of C-DAR(1) model for Sign Language

Contribution of Results for simple and accurate modeling when using Sign Language

- **$\varrho = \lambda 1$ (close to 0.998) – Conservative Choice**
- **Q is constructed via the MOM Method with p and μ parameters**
  - **$68 < p < 72$ and $43 < \mu < 50$ for H.261**
  - **$62 < p < 70$ and $24 < \mu < 28$ for H.263**

Queuing Analysis via the C-DAR model and the fluid-flow method

Assume our model has M states, and the V is the rate vector $V=(V_1, V_2, \ldots, V_M)$, where V_i is the video bit rate in state i. Then the traffic can be expressed as (Q, V), where Q is the Transition Rate Matrix derived from the C-DAR model.

The queue occupancy is a continuous random variable x, $0<x<K$, where K is the queue buffer capacity. Define the steady-state probability distribution function $F_i(x)$ as the joint probability that the buffer occupancy is less than or equal to x, when in the i state of the source model.

If:

$$\vec{F}(x) = [F_1(x), F_2(x), \ldots, F_M(x)]^T \qquad D = diag(d_1, d_2, \ldots, d_M)$$

where $d_i = V_i - C$

Then we have:

$$D \cdot \vec{F}(x) = Q^T \cdot \vec{F}(x)$$

And the frame sizes overflow probability is:

$$Poverflow = \sum_i (V_i - C) P_i / \overline{V}$$

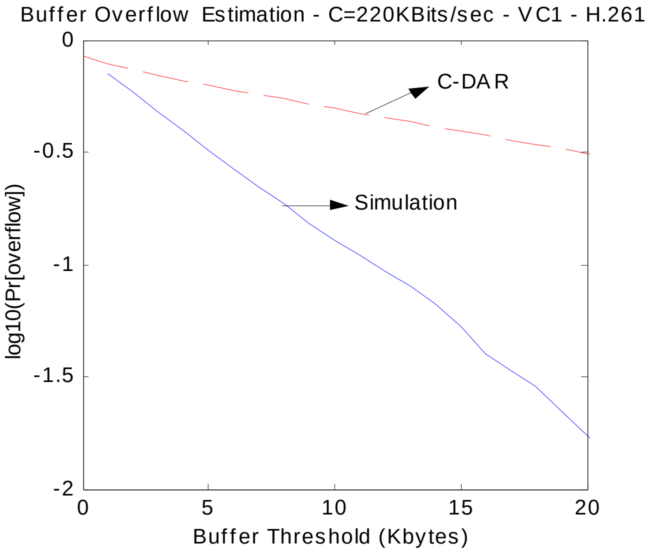
where

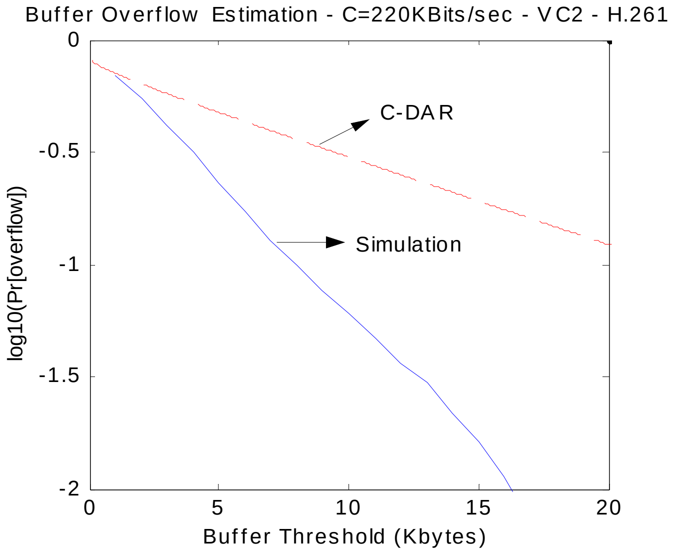
$$\overline{V} = \sum_{i \in S} \pi_i \cdot V_i \qquad Pi = Pi\{q(t) = K\} = \begin{cases} 0 & i \in \Omega_- \\ \pi_i - F_i(K), & i \in \Omega_+ \end{cases}$$

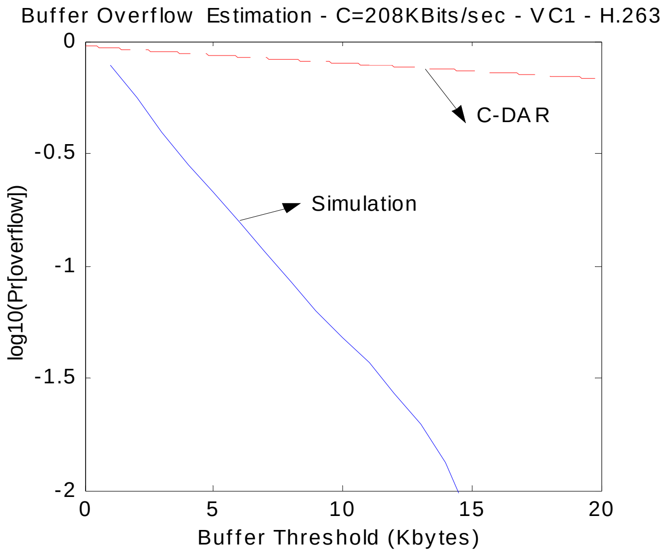
$$\Omega_+ = \{i \mid V_i > C\}$$

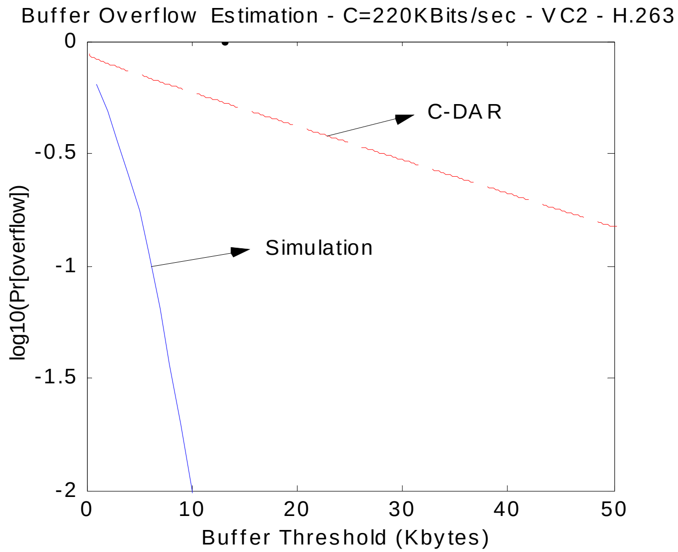
$$\Omega_- = \{i \mid V_i < C\}$$

Analysis Results vs Trace Driven Simulation









Further Work

- Experiments with different clients (CuSeeMe, VCON)
- Modeling Analysis of the new codec H.264
- Analysis of the traffic from the MCU in continuous presence mode

Thank you