Introduction

- Up to now: M/D/1 models for performance assessment (delay) of packetized voice streams

- Problem: can M/D/1 always be applied?

- Generic Markov modulated queueing model that can capture more details of the traffic streams (active/passive periods, silence detection, signalling traffic ...)

- Purpose: comparison of D-BMAP/D/1 and M/D/1 for ‘realistic’ scenarios
Packetized voice transport

one-way Mouth-to-Ear (M2E) delay
overall distortion (codec & packet loss)

Encoding and packetisation stage
Packet transport stage
Dejittering and decoding stage

Dejittering delay

Deterministic delay
Stochastic delay

Pr[w > W]

Minimal delay
Delay of first packet
Mouth-to-ear delay

Packet loss P

Worst case packet loss, if first packet is fastest possible
Dejittering delay (2)

- Literature: calculations based on M/D/1 model
- Networking 2000 conference: heuristic formula for calculating the dejittering delay based on M/G/1:

\[ W_d = \mu + \alpha(P)\sigma \quad \mu = \text{mean delay} \]
\[ \sigma = \text{standard deviation} \]

\[ \alpha(P) = \begin{cases} 1 & P = 10^{-10} \\ \frac{2N}{1+2N} & P = 10^{-9} \\ \frac{N}{2} & P = 10^{-8} \\ \frac{N}{4} & P = 10^{-7} \\ \frac{N}{8} & P = 10^{-6} \\ \frac{N}{16} & P = 10^{-5} \\ \frac{N}{32} & P = 10^{-4} \\ \frac{N}{64} & P = 10^{-3} \\ \frac{N}{128} & P = 10^{-2} \\ \frac{N}{256} & P = 10^{-1} \end{cases} \]

\[ \sigma = \sqrt{\frac{\mu^2}{N}} \]

Dejittering delay (3)

M/D/1 queue; 16 nodes
Generic queueing model

- Is M/D/1 model always valid?

- Consider generic queueing model:
  heterogeneous D-BMAP/G/1 (1 node)
  » Sources are grouped into $K$ traffic classes; $N_k$ sources per traffic class
  » 1 transmission line; no interruptions
  » Infinite storage capacity
  » Packet arrivals generated by class-dependent D-BMAP (special case: MMBP)
  » Packet transmission times of class $k$ described by general distribution

- Result: generating-functions approach
  \[ \Pr[w_k > W] = \sum_j C_j \ z_j^W \] (multiple-poles approximation)
  \[ W : 5 - 10 \text{ ms} \]

Traffic scenario

- Multiplex $N$ identical packetized voice streams

- Packetized voice stream:
  \[ R = \text{link rate (512 \text{ kbps} - 10 \text{ Mbps})} \]
  \[ L_p = \text{packet length (160 payload + 40 header = 200 bytes)} \]
  \[ T_a = \text{time to transmit 1 packet (3.125 ms - 390.6 ms)} \]
  \[ I_a = \text{IAT between packets} \]
  \[ \text{encoding rate} = 16-64 \text{ kbps} \Rightarrow I_a = 80-20 \text{ ms} \]
  \[ T_a = \text{Length active period} = 120 \text{ s} \]
  \[ p = \Pr[\text{source is active}] \text{ (e.g., 3 calls/hour} \Rightarrow p = 0.1 \]
  \[ T_p = \text{Length passive period} \Rightarrow T_p = \frac{p}{1-p}T_a \]
Arrival model

- \( N \) sources, each described by 2-state Markovian arrival process:

\[
\begin{align*}
\begin{array}{c}
\text{A} \\
\text{P} \\
\text{P} \\
\text{A}
\end{array}
\end{align*}
\]

\[
\begin{align*}
p_{11} & \quad p_{12} \\
p_{21} & \quad p_{22}
\end{align*}
\]

- Probability generating matrix \( Q(z) = \begin{bmatrix} p_{11}G(z) & p_{12}F(z) \\ p_{21}G(z) & p_{22}F(z) \end{bmatrix} \)

- \( T_a \) and \( T_p \) = mean length of geometrically distributed ON and OFF period respectively

\[
\Rightarrow p_{12} = 1/T_a, \quad p_{11} = 1 - 1/T_a
\]

\[
p_{21} = 1/T_p, \quad p_{22} = 1 - 1/T_p
\]

- No packets during passive period \( \Rightarrow F(z) = 1 \)

Bernoulli arrivals during active period \( \Rightarrow G(z) = 1 - \gamma + \gamma z \) with \( \gamma_p = 1/I_a \)

Maximum load versus link rate

Source activity \( p = 0.1 \)

\[
\begin{align*}
\Pr[w > 6.25 \text{ ms}] < 1E-4 & \quad \Pr[w > 12.5 \text{ ms}] < 1E-4
\end{align*}
\]

Source bit rate

- 64 kb/s
- 32 kb/s
- 16 kb/s
- M/D/1
Maximum load versus link rate

Source activity $p = 0.025$

![Graph showing maximum load versus link rate for different bit rates and source activity $p = 0.025$.](image)

Multiple poles approximation

$$\Pr[w_k > W] = \sum_j C_j z^{-jW}$$

$p = 0.01$

$W = 6.25$

$N = 14000$

![Graph showing the approximation of the multiple poles.](image)
Impact of variance active periods: arrival model

- \(N\) sources, each described by a 3-state Markovian arrival process:

\[
\begin{pmatrix}
    p_1 G(z) & 0 & p_{23} \\
    0 & p_{22} G(z) & p_{23} \\
    p_3 G(z) & p_{32} G(z) & p_{33}
\end{pmatrix}
\]

- \(T_{a1}, T_{a2}, T_{p}\) = mean length of geometrically active 1, active 2 and passive period

\[p_{31} = p/T_p, \quad p_{32} = (1-p)/T_p, \quad p_{33} = 1 - 1/T_p\]

\[
\frac{1}{\mu_{p1}} = I_a \left(1 + \sqrt{\frac{1-p}{2p}(L-1)(1-1/I_a)}\right) \quad \frac{1}{\mu_{p23}} = I_a \left(1 - \sqrt{\frac{p}{2(1-p)}(L-1)(1-1/I_a)}\right)
\]

\[\text{Var[active period]} = L \times I_a (I_a - 1)\]

- Bernoulli arrivals during active period \(\Rightarrow G(z) = 1 - \gamma_a - \gamma_c\) with \(\gamma_c = 1/I_a\)

Impact of variance active periods (2)

Source Activity \(p = 0.1, \Pr[w > 6.25 \text{ ms}] < 1E^{-4}\)
Traffic scenario (silence detection)

- Multiplex $N$ identical packetized voice streams

- Packetized voice stream:

```
  +-------------------------+          +-------------------------+
  |                         |          |                         |
  | talking                 |          | listening               |
  +-------------------------+          +-------------------------+
```

- $R =$ link rate (512-4096 kb/s)
- $L_p = $ packet length (160 payload + 40 header = 200 bytes)
- slot length = time to transmit 1 packet (3.125 ms - 390.6 $\mu$s)
- $I_a =$ IAT between packets when talking
  
  codec rate = 16-64 kb/s $\Rightarrow I_a = 80-20 ms$
- $T_a =$ Length active period = 120 s
- $T_t =$ Length talking period = 1 s
- $T_l =$ Length listening period = 1.5 s
- $p =$ Pr[source is active]
- $T_p =$ Length passive period $\Rightarrow T_p = \frac{p}{1-p}T_a$

Arrival model (silence detection)

- $N$ sources, each described by 3-state Markovian arrival process:

```
  +---+     +---+     +---+
  | T |     | P |     | L |
  +---+     +---+     +---+
```

- Probability generating matrix $Q(z) = [p_{11}G(z) \quad p_{12} \quad 0]

- $T_r, T_a, \text{ and } T_p =$ mean length of geometrically distributed talking, listening and passive period

\[
p_{12} = \frac{1}{T_r} \quad , \quad p_{11} = 1 - p_{12} \\

p_{31} = \frac{1}{T_p} \quad , \quad p_{33} = 1 - p_{31} \\

p_{22} = 1 - \frac{1}{T_t} \quad , \quad p_{23} = \frac{(T_t + T_l)}{(T_a T_t)} \quad , \quad p_{21} = 1 - p_{22} - p_{23}

- Bernoulli arrivals during active period $\Rightarrow$ $G(z) = 1 - \gamma + \gamma z$ with $\gamma_a = 1/I_a$
Impact of silence detection

Source Activity $p = 0.1$, $Pr[w > 6.25 \text{ ms}] < 1E-4$

![Graph showing impact of silence detection on source activity and bit rate](image)

Conclusions

- Calculation of dejittering delay

- Validity of M/G/1 model: comparison with D-BMAP/G/1
  - maximum load determined by encoding and output link rate
  - negligible impact of
    - active period distribution
    - silence detection
    - signalling traffic
    - ...

COST 257 Final Seminar 29/09/2000 : 18
Future work

- Alternative scenarios (packet lengths, …)
- Heterogeneous voice streams
- Constant packet rate during active periods
- End-to-End calculations
- ...

COST 257 Final Seminar 29/09/2000 : 19