Performance Evaluation of Rate-Control Strategies for real-time Systems Over a Rate-Varying Channel

Muhammad Ajmal, Peter Johansson, Robert Forchheimer
Division of Information Coding, Dept. of Electrical Engineering
Linköping University, SE-581 83 Linköping, Sweden
Email: {ajmal, pejoh, robert}@isy.liu.se

Abstract— We compare the performance of two real-time systems for communication of distortion-tolerant content such as audio and video over rate-varying channels. The basic difference between the systems is the placement of a buffer either before or after the rate-controlled encoder. For our evaluation we target signal quality versus end-to-end delay. We develop an analytical framework for their comparison, when the systems use a deterministic rate-varying channel. Furthermore, for stochastic rate-varying channels we use the Gilbert-Elliott channel model and suggest some heuristic rate-control strategies for the two systems when communicating over such a channel.
Contrary to intuition it is shown that one of the systems performs better than the other system. As the analysis uses approximations, performance is also validated through simulations.

Index Terms— Buffers, Interactive systems, Source coding, Rate-varying channel.

I. INTRODUCTION

INTERACTIVE communication is becoming common on present day networks. In most of the present communication networks/channels the available rate or throughput are highly dynamic, such as in Internet or IP based networks in which delays are neither bounded nor predictable and data is prone to packet losses due to congestion. Similarly, for the wireless channel, interference and fading may make the throughput vary severely with time.

Media Signals used in point-to-point interactive communication is often represented using lossy compression techniques, where the output rate and the accuracy of the representation at the decoder are chosen by the encoder. This provides a trade-off between channel utilization and signal quality. Moreover, it is common to use a transmit buffer to temporarily store the encoded data for the input signal, when the channel rate is limited and does not match with the encoded rate. As interactive communication imposes strict constraints on overall delay, this translates to a limited size of the transmit buffer. Ultimately the limited transmit buffer puts constraint on the encoding rate i.e. the signal needs to be encoded at the rate which guarantees the overall delay constraints. To

cope with this challenge various potential solutions have been proposed for video transmission based on source-rate control methods performed at frame level [1]-[6] and/or at macro block level [7]-[10]. These methods are applicable to interactive and pre-encoded video i.e. streaming and rely on the information about the state of the channel, transmitter buffer size and its fullness, and the amount of bits used to encode previous frames. These authors have used the concept of a buffer after the encoder to smooth the variation in bit rates produced by the encoder, and similarly before the decoder to keep the end-to-end delay constant.

However, [11] has used the concept of a smoother, which will pre-fetch some of the data before schedule by creating *look-ahead delay* for the streaming of pre-recorded video. The authors of [12] have used the concept of dynamic rate shaping to *shape* the video that is, to reduce, the bit rate by dropping part of the pre-encoded video, according to current network condition. The same authors have extended the concept of rate shaping by adopting MPEG-4 fine granularity scalability (FGS) for source coding in [13], and further taking into account the error concealment (EC) method in [14].

The last four papers have treated the concept of buffering before the rate-control decision device, but for streaming of pre-encoded video. To our knowledge, none of the authors have used this idea in an interactive communication system, nor have we found any performance comparison of the above mentioned systems for the same end-to-end delay. One obvious reason for the latter case may be, that it is a common perception that both systems are equivalent in their performance and that placement of the buffer either before or after the rate-control decision device causes no affect on the system performance.

In this paper, we look into the issue of the buffer position and its affect on system performance, by comparing two simple communication systems namely the input buffer system and the transmit buffer system. The input buffer system uses the (FIFO) buffer before the encoder and after the decoder while the transmit buffer system makes use of the (FIFO) buffer after the encoder and before the decoder. Performance is measured as SNR value versus end-to-end delay. The end-to-end delay is measured from the time when the data enters the system until the same

data leaves the system. First of all, we describe in detail both the system models. As the input buffer system stores the input signal in uncompressed form, the buffer size will be larger than the buffer size of the transmit buffer system, which stores the compressed bitstreams. The size ratio will depend on the compression factor. To compare them we formulate the framework, and initially analyze their performance for a deterministic rate-varying channel. Secondly, we suggest a model for a stochastic rate-varying channel and look at some source-rate control strategies for both communication systems. Finally, we evaluate the performance of these strategies both analytically and through simulation. We use the deterministic framework as a counterexample to the intuition that the system performance for the same end-to-end delay is independent of buffer location. Although highly simplified, the deterministic channel model behavior may occur in practical systems e.g. during periodic radio fading.

The paper is organized as follows. In section II, we describe the two communication systems in detail along with the framework for their comparison. Based on a deterministic rate-varying channel, analytical analysis is performed in section III. In section IV, we formulate a stochastic model for a rate-varying channel, and in section V we introduce some rate-control strategies for communication over the formulated channel. In section VI, one of the rate-control strategies is analyzed in detail for both the systems. Section VII provides the analytical and simulation performance results. Conclusions are drawn in section VIII.

II. COMMUNICATION SYSTEM MODELS

In this section we will describe in detail the two systems and will formulate an analytical framework for their comparison which will be used in further sections.

A. Input Buffer System

In the Input Buffer System the transmitter consists of an input buffer and an instantaneous lossy encoder as shown in Fig.1.

We consider a general "continuous" source. This could be an analog signal in which the buffer content is expressed in seconds, or a time discrete signal in which the buffer content is expressed in number of samples. A video source could be fit into this description if the video data is split into blocks (possibly frames) which should be small in comparison with the size of the input

Figure 1. Input buffer communication system model

buffer. Moreover, for simplicity we assume a memoryless Gaussian source with unit variance and use the mean squared error distortion measure. Gaussian sources are the most difficult to compress and represent a worst case scenario [15]. The Distortion-rate function for a Gaussian source with unit variance and squared distortion measure is given by

$$D(R) = 4^{-R/2W} \qquad (1)$$

Where R is the instantaneous rate in bits/s and W is the bandwidth of the analog input signal. In the rest of this paper we will assume that (1) holds also for the "operational distortion-rate" function of the used encoders. We further assume that the channel fluctuates between two states with transmission delay equal to zero[1]. In the bad state, with duration t_1, the channel transmits R_1 bits/s. In the good state, with duration t_2, the rate is R_2 bits/s. Moreover, R_1 and R_2 indicate the throughput of the channel during these states and R_1 is assumed to be less than that of R_2, e.g. due to addition of more redundant bits for error control. The times for the good and bad states are known for the deterministic rate-varying channel. The receiver unit consists of an instantaneous decoder followed by an output (play-out) buffer of the same size as the input buffer to compensate for the jitter produced by the transmitter. The total delay t_d for this system is given by the capacity of the input buffer, which needs to be measured in seconds (or samples) as the signal is still in its non-digitized or uncompressed form. In our work, the allowed system delay is assumed to be shorter than a full cycle $t_1 + t_2$ of the channel states. If we encode the input signal at the channel rate then the input buffer will remain empty with $t_d = 0$ seconds while if we encode the input signal at higher rate than that of the channel rate, after some time we will have filled the input buffer. This makes the system delay equal to the signal stored in the input buffer. At that instant we can either switch to the channel rate to avoid buffer overflow, or encode at a slower rate than the channel rate to empty the input buffer. The rate-control strategy that we adopt for this deterministic rate-varying channel is denoted *constant rate per state*. It uses different constant encoding rates for each state within the constraint imposed by the input buffer, and is discussed in the Appendix. The optimal constant rate for the bad state is to encode at a rate R_{s_1} that leaves the input buffer full at the end of the bad state. We use t' to denote the time index for the input signal. At time t = 0 seconds both times[2] will be zero, but at the end of the bad state t' will lag from channel time by t_d seconds

$$t_1 R_1 = t' R_{s_1} = (t_1 - t_d) R_{s_1} \qquad (2)$$

It is clear from (2) that at the end of the bad state only $t' = (t_1 - t_d)$ seconds of input signal have been encoded (and sent) during t_1 seconds of real time, while t_d seconds of input signal will be in the input buffer. In the good
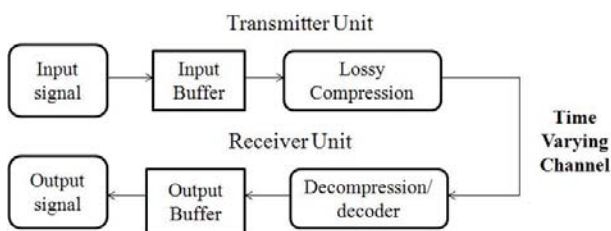
---

[1] Any other delay can be accounted for by shifting the final results accordingly.

[2] The channel time t and the input signal time t'.

state, we adopt a constant encoding rate R_{s_2} that empties the input buffer at the end of the good state and so t' will catch up with the channel time. For the good state we have

$$t_2 R_2 = t' R_{s_2} = (t_2 + t_d) R_{s_2} \qquad (3)$$

A similar rate-control strategy is suggested in [18,19] for streaming video over wireless channels, in which the playback buffer content builds up during the good states and shrinks during the bad states. Now to geometrically interpret the *constant rate per state* we take the curve of cumulative number of bits transmitted over the channel and shift it to the left by t_d seconds as shown in Fig 2. The accumulated number of bits produced at the encoder will be a curve within the corridor of Fig 2. The average distortion will be given as

$$D = 4^{-R_{s_1}/2W} \Big(\frac{t_1 - t_d}{t_1 + t_2} \Big) + 4^{-R_{s_2}/2W} \Big(\frac{t_2 + t_d}{t_1 + t_2} \Big) \quad (4)$$

To find the value of R_{s1}, we know that in the bad state, $(t_1 - t_d)$ seconds of input signal is transmitted. The number of bits available are $R_1 t_1$. Thus it is seen that

$$R_{s_1} = \frac{R_1 t_1}{t_1 - t_d} \qquad (5)$$

Similarly for the good state $(t_2 + t_d)$ seconds of input signal is sent. The available numbers of bits are $R_2 t_2$. This requires that

$$R_{s_2} = \frac{R_2 t_2}{t_2 + t_d} \qquad (6)$$

It should be noted that these choices of rates are only valid as long as $R_{s_1} < R_{s_2}$. At the point where they become equal due to a sufficiently large input buffer, a common fixed rate will be optimal. It can be seen that this happens when the allowed delay is

$$t_d^* = \frac{t_1 t_2 (R_2 - R_1)}{t_1 R_1 + t_2 R_2} \qquad (7)$$

corresponding to the *cross-over* rate,

$$R_{s_1} = R_{s_2} = \frac{(t_1 R_1 + t_2 R_2)}{t_1 + t_2} \qquad (8)$$

At t_d^*, the corridor will be wide enough, thus enable the system to sustain a constant encoding rate for the two states, and will be independent of the channel throughput variation. At higher allowed delays, the optimal rate depends on the relation between the input buffer size and the total transmission time $T_{session}$. If we consider the total transmission time to be *infinite*, the above rate is the best that can be used. Any further allowed delay will not improve the performance for the *constant rate per state* strategy.

As explained earlier the receiver unit consists of an instantaneous decoder followed by an output (play-out) buffer to overcome the jitter produced by the transmitter. For jitter-free play-out we need to avoid both overflow and underflow. To avoid output buffer underflow the initial
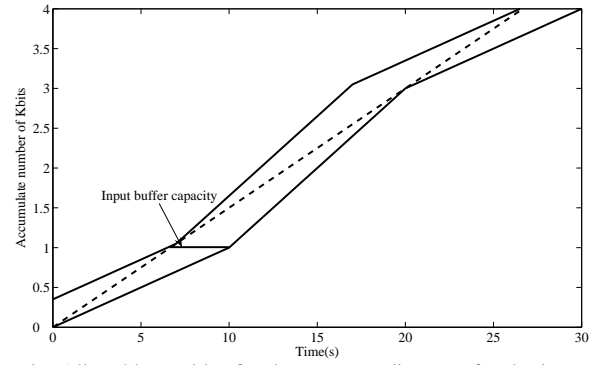

Figure 2. Allowable corridor for the source coding rate for the input buffer system model

play-out delay δ has to be selected such that at any time instant t we have at least $(t' - \delta)$ seconds of the input signal available for output. i.e.,

$$\delta \; \epsilon \; \Re \; : \quad t \; \geq \; (t' - \delta), \quad \forall t \leq T_{session} \qquad (9)$$

The left-hand side of (9) defines the time index for the channel signal for t seconds received by the instantaneous decoder while the right-hand side defines the output (play-out) signal for $(t' - \delta)$ seconds. Moreover, (9) should hold for the whole transmission time i.e. $T_{session}$. The value of δ is chosen by right-shifting the encoding rate curve by t_d seconds shown in Fig.2 which is equal to the capacity of the input buffer.

This means that the total delay for the input buffer system is equal to the capacity of the input buffer which we have already stated. Moreover, to avoid output buffer overflow, the size of the output buffer must be equal to $max\{t - (t' - \delta)\}$. The maximum difference occurs when the input buffer becomes empty (and the output buffer fills fully). This means that the size of the output buffer must be equal to the input buffer size.

B. Transmit Buffer System

In the Transmit Buffer System the transmitter unit consists of an instantaneous coder followed by an output buffer shown in Fig.3. The receiver unit consists of a receive buffer followed by an instantaneous decoder. In contrast to the input buffer case, the size of the transmit buffer is expressed in bits. We use the same source coder and channel model as we assumed for the input buffer system. The transmit buffer is used for storing encoded data when the encoding rate is higher than that of the channel rate. Due to the real-time constraint we will use a limited size transmit buffer B_t of bits. Similarly, the receive buffer of size B_r bits is used to cope with the jitter produced by the transmitter. For the transmit buffer system, the *constant rate per state* strategy will encode the input signal at R_{s_1} rate and will fill the transmit buffer at the end of the bad state.

Thus,

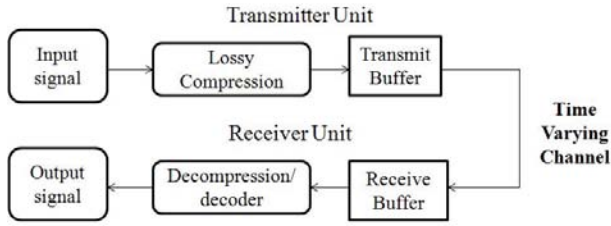$$t_1 R_{s_1} = t_1 R_1 + B_t \qquad (10)$$

Figure 3. Transmit buffer communication system model

Similarly, for the good state the encoding rate R_{s_2} is chosen such that to empty the transmit buffer at the end of the good state.

$$t_2 R_{s_2} = t_2 R_2 - B_t \qquad (11)$$

Thus,

$$R_{s_1} = \frac{B_t}{t_1} + R_1 \qquad (12)$$

$$R_{s_2} = R_2 - \frac{B_t}{t_1} \qquad (13)$$

To geometrically interpret the *constant rate per state* strategy for the transmit buffer system we take the curve of cumulative number of bits transmitted over the channel and shift it upwards by B_t bits as shown in Fig.4. The accumulated numbers of bits produced at the encoder will generate a curve within the corridor of Fig.4 while the average distortion will be equal to

$$D = 4^{-R_{s_1}/2W}\left(\frac{t_1}{t_1 + t_2}\right) + 4^{-R_{s_2}/2W}\left(\frac{t_2}{t_1 + t_2}\right) \quad (14)$$

The system delay can be deduced as follows. Consider the case when the transmit buffer is full. As we have already stated this happens when the channel has been in the bad state and the transmit buffer is filled with the input signal coded at R_{s_1} bits/s. This means that the transmit buffer contains B_t/R_{s_1} seconds of the input signal. At that instant the receive buffer is empty (in steady state) and the receiver will not add any additional delay. Thus

$$t_d = \frac{B_t}{R_{s_1}} \qquad (15)$$

By replacing B_t above in the expressions for R_{s_1} and R_{s_2}, we obtain

$$R_{s_1} = \frac{R_1 t_1}{t_1 - t_d} \qquad (16)$$

$$R_{s_2} = R_2 - \frac{t_d t_1 R_1}{t_2(t_1 - t_d)} \qquad (17)$$

Inserting the values of R_{s_1} and R_{s_2} in (14) we can finally express D as a function of delay t_d.

It should be noted again that these choices of rates are only valid as long as $R_{S_1} < R_{S_2}$. At a point where they become equal due to a sufficiently large transmit buffer, a common fixed rate will be optimal. It can be seen that this happens for the transmit buffer size

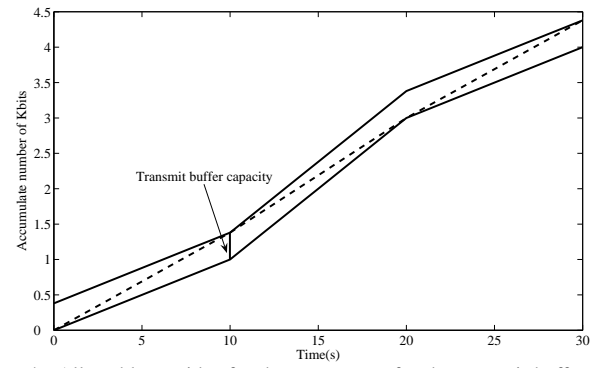$$B_t^* = \frac{t_1 t_2 (R_2 - R_1)}{t_1 + t_2} \qquad (18)$$

Figure 4. Allowable corridor for the source rate for the transmit buffer system

corresponding to the rate

$$R_{s_1} = R_{s_2} = \frac{t_1 R_1 + t_2 R_2}{t_1 + t_2} \qquad (19)$$

This is the same cross-over rate as for the input buffer case. Inserting this rate yields the *cross-over* delay

$$t_d^* = \frac{B_t^*(t_1 + t_2)}{R_1 t_1 + R_2 t_2} = \frac{t_1 t_2 (R_2 - R_1)}{R_1 t_1 + R_2 t_2} \qquad (20)$$

Comparing with (7) shows that both systems achieve the same performance for this value of the delay. It should be noted that the receive buffer will not have the same size as the transmit buffer. Let us consider the situation at the end of the good state. The transmit buffer is now empty and the receive buffer is full. The latter will now hold a signal segment coded at R_{s_2} bits/s. This means that the signal stored in the receive buffer corresponds to B_r/R_{s_2} seconds. Since the system cannot allow different delays at different times we find that,

$$t_d^* = \frac{B_t}{R_{s_1}} = \frac{B_r}{R_{s_2}} \qquad (21)$$

Thus the size of the receive buffer needs to be larger than the transmit buffer in order to hold the larger amount of data produced in the good channel state for the same time duration of the input signal.

III. A PERFORMANCE COMPARISON FOR THE DETERMINISTIC CHANNEL

In order to compare the two communication systems we take the parameters values as $t_1 = 0.1$ seconds, $t_2 = 1$ seconds, $R_1 = 64$ Kbit/s, $R_2 = 192$ Kbit/s and $W = 12.5$ KHz, typical for speech or lower quality audio signal. For the cross-over delay we find that $t_d^* = 0.0645$ seconds at the rate $R_{s_1} = R_{s_2} = 180.36$ Kbit/s. We further find that the *cross-over* transmit buffer size for the transmit buffer system becomes $B_t^* = B_r = 11.6$ Kbits. Fig.5 shows the performance of the two systems. As the value of t_d increases from zero, the performance difference increases accordingly, and reaches a maximum of 1.75 dB at $t_d = 0.04$ seconds. The reason for the better performance of the input buffer system is that the lower

rate R_{s_1} is used during a shorter time $(t_1 - t_d)$ compared to the transmit buffer system (t_1). For higher channel rates this performance difference would increase further. However, further increase in the value of t_d will result in diminishing the performance difference until $t_d = t_d^*$, where the performance of the two systems will again become equal. Increasing t_d beyond t_d^* will not contribute any improvement in performance.

IV. STOCHASTIC RATE-VARYING CHANNEL

In the previous section, we considered a deterministic rate-varying channel where we had full knowledge of the channel duration and rate in each state. Many practical communication channels behave randomly and their durations in any state varies over time. To take account of the random nature of the channel, we assume a Gilbert-Elliot model [16] that fluctuates between two states, the bad state with rate R_1 bits/s and the good state, with rate R_2 bits/s. Moreover, R_1 and R_2 indicate the throughput of the channel in these states. The time spent in each state is exponentially distributed, with parameters λ_1 and λ_2 respectively. Thus

$$P(T_i \le t) = \int_0^t \lambda_i e^{-\lambda_i T} dT \qquad (22)$$

The average time in the bad state is thus $1/\lambda_1$ seconds and for the good state it is $1/\lambda_2$ seconds.

V. RATE CONTROL STRATEGIES

In this section we will discuss two rate control strategies for the two systems when using the above channel model. Our objective is to compare the resulting performance of these systems to see if the previously found difference between them is true also for the stochastic rate-varying channel.

A. One Rate per State

The *one rate per state* rate-control strategy is similar to the *constant rate per state* strategy i.e. when the channel
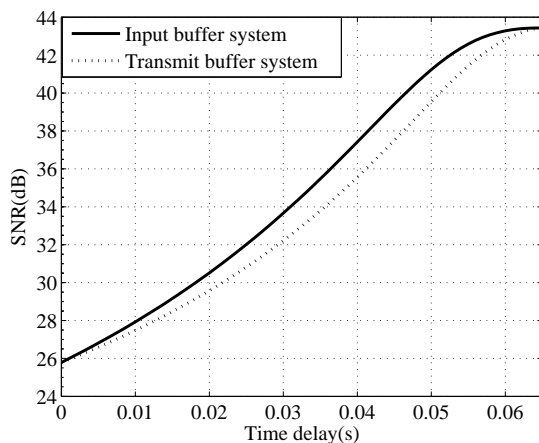

Figure 5. Performance comparison of the two communication systems. SNR=10log(1/D)

switches state the algorithm will make a choice of rate during that state and will pursue that rate within the constraint[3] imposed by the input buffer or transmit buffer. The algorithm will guess a length τ_i for the time that the channel will stay in the current state i. For the transmit buffer system at the beginning of the bad state interval, the chosen coding rate will be equal to

$$R_{s_1} = \frac{B_t}{\tau_1} + R_1 \qquad (23)$$

If the channel duration in the bad state exceeds τ_i (guess duration for the bad state) the coding rate will switch to that of the channel rate R_1 to avoid overloading the buffer. It is also possible that the channel switches to the good state before τ_1, thus without having filled the transmit buffer completely. In the good state, the chosen coding rate will be

$$R_{s_2} = R_2 - \frac{B_t}{\tau_2} \qquad (24)$$

Several different cases may be possible, depending on the channel duration in the good state:

a) : If the transmit buffer is full and the channel stays in the good state for τ_2 (guess duration of the good state) then there will be one coding rate (24) for the entire good state τ_2.

b) : If the transmit buffer is full and the channel remains in the good state for less duration than τ_2, then the coding rate (24) will not be able to completely empty the transmit buffer.

c) : If the transmit buffer is partially full due to earlier switching from the bad state, and the channel stays in the good state for τ_2 then the coding rate (24) can not be sustained for the entire good state (due to partial fullness of the buffer) and the coding rate will switch to R_2 after emptying the transmit buffer.

d) : Similarly, if the transmit buffer is partially full and the channel remains in the good state for a duration less than τ_2, then either it empties the transmit buffer at coding rate (24), or will leave some bits before switching to the bad state. It is also possible that the transmit buffer becomes empty earlier and at that time the encoding rate will follow R_2.

For the input buffer system we will have also the same possibilities with coding rate $R_1 t_1/(\tau_1 - t_d)$ in the bad state and $R_2 t_2/(\tau_2 + t_d)$ in the good state. The coding rate will switch to R_1 in the bad state upon filling the input buffer and will follow R_2, when it empties the input buffer in the good state.

B. The Engineer's Algorithm

The idea behind the *Engineer's algorithm* scheme resemble the rate-control scheme described in [20],

---

[3]Buffer becomes full or empty completely.

where the encoding rate is adjust for every T seconds, on the basis of channel information and play back buffer occupancy. The essence of the strategy is to make a general algorithm, that adjusts the coding rate on short time instant basis, so as to allow for a broader class of channels (more than two states). The algorithm will guess a length τ_i for the time that the channel will stay in state i, but will take a new decision about the rate at every time instant by sensing the occupancy of the buffer. At every instant, the algorithm will aim to fill the buffer during the next τ_1 seconds, when it considers the channel to be in bad states (e.g. worse than the estimated average rate). Consequently this will lead to an exponentially increasing filling pattern and the buffer will never be filled completely. The same strategy will be followed for emptying the buffer during the good states. Unlike the *one rate per state* strategy, the *Engineer's algorithm* will not utilize the entire buffer. However, for the transmit buffer system communicating over a highly unreliable/variable channel, full utilization of the transmit buffer may adversely affect the system performance. For example, when the channel rate all of a sudden becomes worse and the transmit buffer is full, then all the data in the buffer needs to be discarded. For the input buffer system, however, in such a situation the encoding rate will switch to the channel rate without discarding any data.

For the two-state channel we will get the following behavior.

1) Transmit Buffer System: The rate at which the transmit buffer will be filled is given by

$$b'(t) = R_{s_1}(t) - R_1 = \frac{B_t - b(t)}{\tau_1} \qquad (25)$$

Where $b(t)$ is the number of bits in the transmit buffer. This will lead to,

$$b(t) = B_t + (b_1 - B_t)e^{-(t-t_1)/\tau_1} \qquad (26)$$

Here b_1 is the number of bits in the buffer at the beginning of the time interval t_1. Initially the value of b_1 will be equal to zero, but onward, that is, for steady state channel condition it will have some other nonnegative values which will depend on the channel duration in these states. Similarly for the good state the rate at which the buffer will be emptied becomes $b'(t) = R_2 - \frac{b(t)}{\tau_1}$ so that

$$b(t) = b_2 e^{-(t-t_2)/\tau_2} \qquad (27)$$

where b_2 is the number of bits in the transmit buffer at the time t_2 when the system switch to the good state. The value of b_2 will also depend on the previous channel duration in these states.

2) Input Buffer System: The ratio of input buffer fullness at any time instant in the bad state will be

$$T'(d) = \frac{t' - t}{t_d} \qquad (28)$$

Where t' and t denotes the time index of the input signal and the channel time respectively as explained in section

II-A while t_d denote the size of the input buffer. This will lead to input buffer filling

$$T_d = t_d + (t'_1 - t_d)e^{-(t-t_1)/\tau_1} \qquad (29)$$

t'_1 shows the fullness of the input buffer at the beginning of the bad state at time t_1. Initially the value of t'_1 will be equal to zero, but onward i.e. when the channel is in the steady state it will have some other nonnegative values which will depend on the channel duration in these states. Similarly for the good state, the buffer will be emptied as

$$T_d = t'_2 e^{-(t-t_2)/\tau_2} \qquad (30)$$

t'_2 indicates the fullness of the input buffer filled at time interval t_2 when the channel switch to the good state. Its value will also depend on the previous channel duration in both states.

VI. ANALYTICAL ANALYSIS OF THE ONE RATE PER STATE STRATEGY

A. Transmit Buffer System Model

We will analyze the *one rate per state* strategy by assuming that the buffer is empty ($b = 0$) at the start of the bad state and is full ($b = B_t$) at the beginning of the good state. The time-average of the expected distortion for this algorithm is computed by first conditioning on the time interval T being longer than τ_1

$$E[D/Bad\ state, T > \tau_1] =$$
$$\int_{\tau_1}^{\infty} \left(\tau_1 D(R_1 + \frac{B_t}{\tau_1}) + (T - \tau_1)D(R_1) \right) P(\mathrm{d}T) =$$
$$\tau_1 D(R_1 + \frac{B_t}{\tau_1})e^{-\lambda_1 \tau_1} + \frac{1}{\lambda_1}e^{-\lambda_1 \tau_1}D(R_1) \quad (31)$$

And then conditioning on the time interval being at most τ_1

$$E[D/Bad\ state, T \leq \tau_1] =$$
$$\int_0^{\tau_1} \left(TD(R_1 + \frac{B_t}{\tau_1}) \right) P(\mathrm{d}T) =$$
$$\frac{1}{\lambda_1}\left(1 - (1 + \lambda_1 \tau_1)e^{-\lambda_1 \tau_1} \right)D(R_1 + \frac{B_t}{\tau_1}) \quad (32)$$

For the bad state we thus get

$$E[D/Bad\ state] =$$
$$e^{-\lambda_1 \tau_1}D(R_1) + (1 - e^{-\lambda_1 \tau_1})D(R_1 + \frac{B_t}{\tau_1})$$
$$(33)$$

Similarly for the good state

$$E[D/Good\ state] =$$
$$e^{-\lambda_2 \tau_2}D(R_2) + (1 - e^{-\lambda_2 \tau_2})D(R_2 - \frac{B_t}{\tau_2})$$
$$(34)$$

we use (15) for the relation between B_t and t_d. Now bringing the results together gives the time-average of the expected distortion

$$E[D] = \frac{1}{\frac{1}{\lambda_1} + \frac{1}{\lambda_2}}\Big(\frac{E[D/Bad\ state]}{\lambda_1} + \frac{E[D/Good\ state]}{\lambda_2}\Big) \tag{35}$$

We can use (35) to find the values of τ_1 and τ_2 that will minimize the expected distortion, as a suggestion of good values to use for those parameters in the algorithm for a given buffer size.

B. Input Buffer System

For the input buffer system we assume that the input buffer will be empty $(T_d = 0)$ at the start of the bad state and is full i.e. $(T_d = t_d)$ at the beginning of the good state. The time-average of the expected distortion for this algorithm is computed by first conditioning on the time interval T being longer than τ_1

$$E[D/Bad\ state, T > \tau_1] =$$
$$\int_{\tau_1}^{\infty}\Big(\tau_1 D(\frac{R_1\tau_1}{\tau_1 - t_d}) + (T - \tau_1)D(R_1)\Big)P(\mathrm{d}T) =$$
$$\tau_1 D(\frac{R_1\tau_1}{\tau_1 - t_d})e^{-\lambda_1\tau_1} + \frac{1}{\lambda_1}e^{-\lambda_1\tau_1}D(R_1) \tag{36}$$

And then conditioning on the time interval being at most τ_1

$$E[D/Bad\ state, T \leq \tau_1] =$$
$$\int_0^{\tau_1}\Big(TD(\frac{R_1\tau_1}{\tau_1 - t_d})\Big)P(\mathrm{d}T) =$$
$$\frac{1}{\lambda_1}\big(1 - (1 + \lambda_1\tau_1)e^{-\lambda_1\tau_1}\big)D(\frac{R_1\tau_1}{\tau_1 - t_d}) \tag{37}$$

For the bad state we thus get

$$E[D/Bad\ state] =$$
$$e^{-\lambda_1\tau_1}D(R_1) + (1 - e^{-\lambda_1\tau_1})D(\frac{R_1\tau_1}{\tau_1 - t_d}) \tag{38}$$

Similarly for the good state

$$E[D/Good\ state] =$$
$$e^{-\lambda_2\tau_2}D(R_2) + (1 - e^{-\lambda_2\tau_2})D(\frac{R_2\tau_2}{\tau_2 + t_d}) \tag{39}$$

Combining the two results and taking the time duration of each state, the time-average of the expected distortion becomes

$$E[D] =$$
$$\frac{(\tau_1 - t_d)}{(\tau_1 + \tau_2)}E[D/Bad\ state] + \frac{(\tau_2 + t_d)}{(\tau_1 + \tau_2)}E[D/Good\ state] \tag{40}$$

From (40) we see that, as the capacity of the input buffer increases, the distortion due to the bad state will decrease accordingly, but this is valid only when $\tau_1 > t_d$.
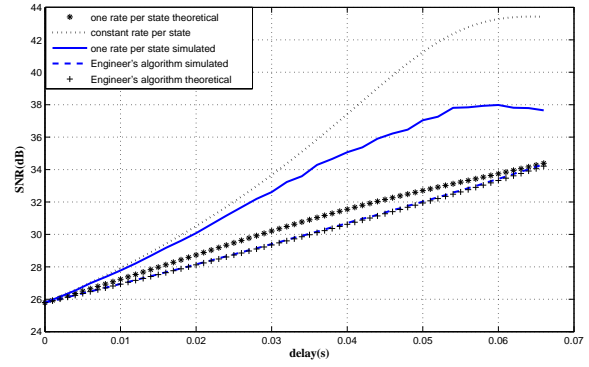
Figure 6. Performance of different rate-control strategies for the input buffer system
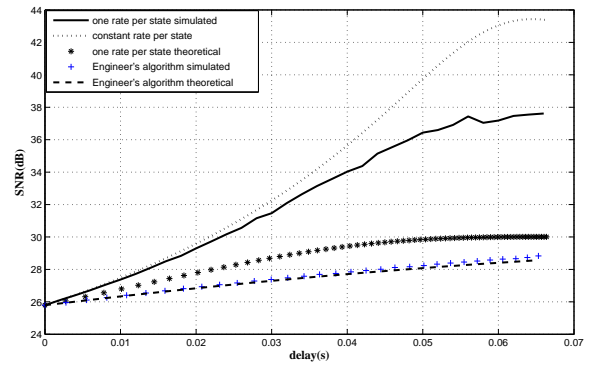

Figure 7. Performance of different rate-control strategies for the transmit buffer system

VII. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed rate-control strategies for each system model, we plot in Fig.6 and Fig.7 the performance results using the expected distortion derived in section VI-A and VI-B and corresponding results obtained through computer simulations of the two systems. The parameters used are $\tau_1 = 10$, $\tau_2 = 1$ corresponding to an average time of 0.1 seconds in the bad state and 1 second in the good state. We used these time durations as guess times of the channel in the two states respectively, for both rate control strategies. The rates R_1 and R_2 are 64 Kbit/s and 192 Kbit/s respectively while the signal bandwidth is $W = 12.5$ KHz. For simulation we used Matlab programming, where for each value of system delay the simulation experiment ran for 800 cycles[4]. We varied the values of channel duration in the bad and good states for each cycle randomly, such that for 800 cycles their mean values are approximately equal to 0.1 second and 1 second respectively. Finally we computed the average of the whole session[5] which gives the average distortion for that value of system delay. Similarly the analytical performance of the *Engineer's algorithm* can be derived for the steady state channel condition. For comparison and verification with the deterministic case in section III,

---

[4]A cycle consists of one bad state and the ensuing good state.
[5]The session consists of 1600 transitions of the channel state.

we also plotted its result for both systems. This can be viewed as an upper bound on the performance for any strategy based on a real-time constraint in the form of a limited buffer at the encoder. The simulation result of the *one rate per state* strategy displays better performance than the corresponding analytical result for both systems. The reason for this difference is the approximation we made that the input/transmit buffer will be full at the end of an interval when the channel is in the bad state, and similarly will be empty at the end of the good state. This will only happen if the interval has length larger than τ_i. But the probability of the above assumption to work out is only 0.368. Moreover, the *one rate per state* strategy outperforms the *Engineer's algorithm*. The latter scheme does not utilize the input buffer efficiently i.e. it will be only partially filled and so the average encoding rate will be lower as compared to the *one rate per state* strategy. For high values of t_d the effect of not occupying the buffer will be diminished. Consequently the performance gap will be reduced for higher values of t_d. Fig.8 shows the comparison of the two systems. In contrast to the deterministic rate-varying channel, which gives the same performance for both the systems at the *cross-over* delay i.e. at $t_d = 0.0645$ seconds, the analytical performance for the transmit buffer system is approximately 4 dB lower

than the input buffer system. The reason for this better performance of the input buffer system is due to the fact that for the analytical case we have two encoding rates for the bad state, namely R_{s1} and R_1 of ratio 0.632 and 0.367 respectively. Similarly, we have two encoding rates for the good state i.e. R_{s2} and R_2 with ratio 0.632 and 0.367 respectively. Although at *cross-over* delay R_{s1} becomes equal to R_{s2}, the expected distortion in the bad state will not be equal to the value for the good state. Furthermore, the input buffer system uses the bad state expected distortion for a duration equal to $(t_1 - t_d)$ seconds in the total expected distortion shown in (40), as compared to the transmit buffer system which is using it for fixed t_1 seconds shown in (35). So, as the value of t_d increases, the performance difference will increase accordingly. Finally, if we look at the simulation results of the *one rate per state* strategy for the two systems, it is clear that the input buffer system outperforms the transmit buffer system with up to 1 dB. For the *crossover* delay $t_d = 0.0645$ seconds the simulation performance of the two systems becomes the same. Fig.9 demonstrates the performance difference of the two systems for the *Engineer's algorithm*. The same justification can be used for this performance gap as explained for the analytical case of the *one rate per state* strategy. From these results,
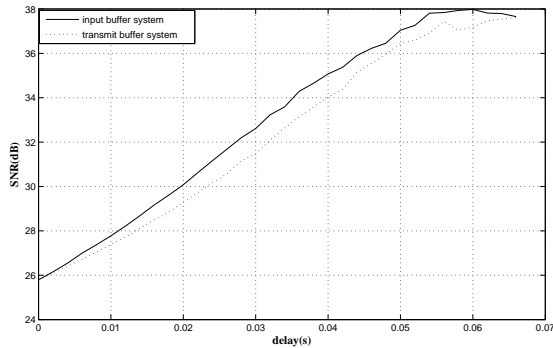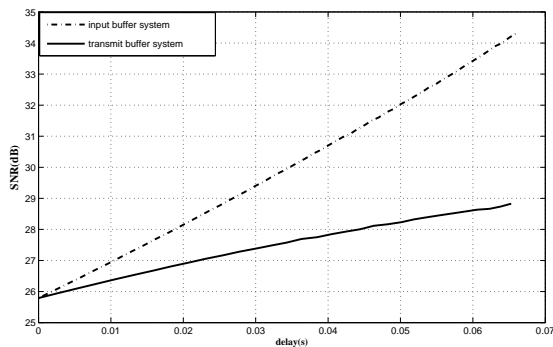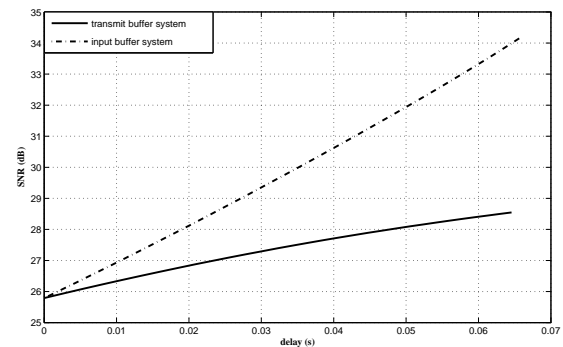

(a) Simulated results
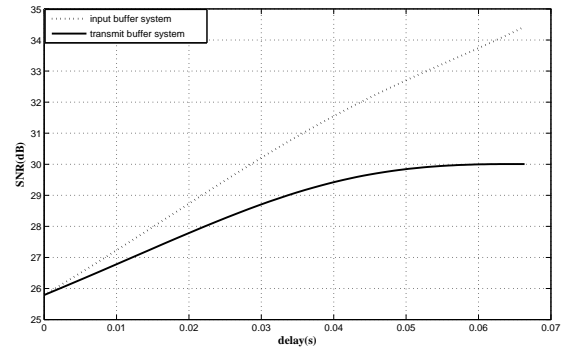
(b) Analytical results

Figure 8. Comparison of the "one rate per state" performance

(a) Simulated results

(b) Analytical results

Figure 9. Comparison of the "Engineer's algorithm" performance

we thus observe that positioning the buffer before the rate control device leads to better performance also for the stochastic rate-varying channel. Secondly, we notice that the *one rate per state* strategy which effectively utilize the buffer for the proposed channel model, gives lower distortion as compared to the *Engineer's algorithm*.

VIII. CONCLUSION

In this paper, we evaluate the performance of two similar real-time communication systems with a fundamental difference of utilizing a buffer before and after the encoder respectively, for quality in terms of SNR versus delay. Each system measures the buffer content in different unit, so we develop an analytical framework for their comparison. Moreover, we show analytically that the buffers of the input buffer system will be symmetric to each other at variable bit encoding while for the transmit buffer system the buffers will not be the same size due to storing of differently encoded signals of the same duration. The comparison is made in terms of system delay and the results shows that the input buffer system is able to use a slightly higher rate. From the periodic channel case analyzed initially, it is seen that a larger portion of the signal can be coded at a higher rate whenever the channel is in the good state compared to the transmit buffer case. We next introduce some practical rate-control strategies when the communication channel switches randomly between a good state and a bad state. We capture this randomness of the channel by taking the Gilbert-Elliott channel model, and apply the rate-control strategies to both the communication systems. Contrary to the intuitive perception, the performance of the system using the buffer before the encoder is better than the system that makes use of the buffer after the encoder, for both the rate-control strategies. However, both the systems become equal in term of performance for a specific end-to-end delay value, when the encoding rates can be chosen the same for both states. At that point the buffers of the transmit buffer system become symmetric to each other. Using these rate-control strategies beyond that end-to-end delay will not improve the performance anymore. The strategies used for the stochastic channel are not necessarily the best one. It can even argued that there may exist some strategy that makes the output buffer case meet the performance of the input buffer case. However, the result from the deterministic channel case gives a strong indication that there is indeed a performance difference between the two systems.

APPENDIX
OPTIMAL SOURCE CODING RATE FOR THE PERIODIC CHANNEL

In this paper we have assumed[6] that the optimal strategy is to keep a constant coding rate until the buffer is full or empty, then change rate. Using this result for the channel which varies rate periodically between two states, the
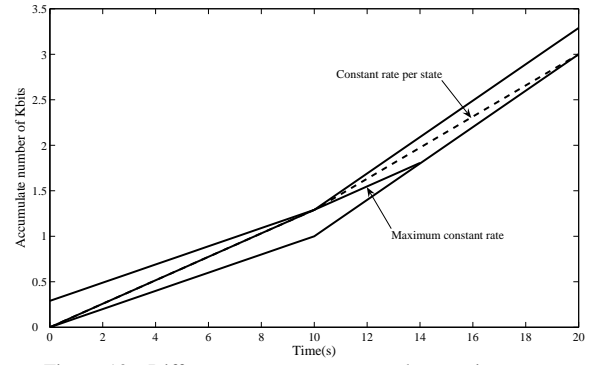
---

[6]This assumption is based on [17].

Figure 10. Different constant rate-control strategies

maximum constant rate for $R_s(t)$ will be the one that touches the upper bounding curve in the accumulated plot at the end of the *bad* state. For the *good* state there are two possibilities; either continue that constant rate (*maximum constant rate* strategy) until we touch the lower boundary, or switch to another constant rate for that state (*constant rate per state*) shown in the Fig.10.

To sort out which constant rate strategy among these gives better performance that is, minimum distortion, we calculate the distortion for these strategies in the *good* state as both will have the same distortion in the *bad* state. For the *maximum constant rate* strategy

$$ D = \frac{T'}{t_2} 4^{-R_{s_1}/2W} + \frac{(t_2 - T')}{t_2} 4^{-R_2/2W} \qquad (41) $$

where T' is the time duration for R_{s_1} in the *good* state, while R_2 is the channel rate in the *good* state. For the *constant rate per state* strategy the distortion for the *good* state is given by

$$ D = 4^{-R_{s_2}/2W} \qquad (42) $$

From (13) we have $R_{s_2} = R_2 - \frac{B_t}{t_2}$, so (42) becomes as

$$ D = 4^{-(R_2 - \frac{B_t}{t_2})\frac{1}{2W}} = 4^{-R_2/2W} \cdot 4^{B_t/2t_2W} \qquad (43) $$

Now if we subtract (43) from (41) we get the difference equal to

$$ 4^{-R_2/2W}\left(1 - 4^{B_t/2t_2W}\right) + \frac{T'}{t_2}\left(4^{-R_{s_1}/2W} - 4^{-R_2/2W}\right) $$
$$ (44) $$

For practical parameter values, the value of $1 - 4^{(B_t/2t_2W)}$ will be some positive value. Similarly $4^{-(R_{s_1}/2W)} - 4^{-(R_2/2W)}$ (as $R_{s_1} < R_2$) will give some positive value, therefore the net distortion difference will be a positive value. This means that the *maximum constant rate* inflicts more distortion as compared to the *constant rate per state*, and from which follows that the optimal rate-control strategy for our assumed channel is the *constant rate per state*.

REFERENCES

[1] H. Wang, A. Ortega, "Rate-distortion optimized schelduling for redundant video representations", *IEEE Trans. on Image processing,* vol.18, No.2, pages 225-240, Feb. 2009.

[2] S. Zhou, J. Li, J. Fei, Y. Zhang,"Improvement on rate-distortion performance of H.264 rate control in low bit rate", *IEEE Trans. on circuits and systems for video tech,* vol. 17, No.8, pages 996-1006, Aug. 2007.

[3] Z. Zhang, G. Liu,H. Li and Y. Li, "A novel PDE-based rate-distortion model for rate control", *IEEE Trans. on circuits and systems for video tech,* vol. 15, No.11, pages 1354-1364, Nov 2005.

[4] H. Lee, T. Chiang, and Y. Zhang, "Scalable rate control for MPEG-4 video", *IEEE Trans. On circuits and systems for video Tech,* vol.10, No.6 pages 878-894, Sept 2000.

[5] H. Song, J. Kuo, "Rate control for low-bit-rate video via variable-encoding frame rates",*IEEE Trans. On Circuits and Systems for video Tech,* vol.11, No.4, pages 512-521, Apr. 2001.

[6] C. Yuan, A. Ortega, M. Khansari, "Rate control for robust video transmission over burst-error wireless channels", *IEEE Trans. On selected areas in communications Tech,* vol. 17, No.5 , pages 756-773, May 1999.

[7] H. Wang, S. Kwong, "Rate-distortion optimization of rate control for H.264 with adaptive initial quantization parameter determination", *IEEE Trans. On circuits and systems for video Tech,* vol.18, No.1, pages 140-144, Jan. 2008.

[8] C. An, T. Q. Nguyen, "Iterative rate-distortion optimization of H.264 with constant bit rate constraint", *IEEE Trans. on Image processing,* vol.17, No.9, pages 1605-1615, Sept. 2008.

[9] J. R. Corbera, S. Lei, "Rate control in DCT video coding for low delay communications", *IEEE Trans. On circuits and systems for video Tech,* pages 172-185, Feb. 1999.

[10] T. Wiegand, M. Lightstone, D. Mukherjee, T. Cambell, S. K. Mitra, "Rate-distortion optimized mode for very low bit rate video coding and emerging H.263 standard", *IEEE Trans. On circuits and systems for video Tech,* vol.6, No.2 pages 182-190, Apr. 1996.

[11] P. Thiran, J. Boudec, F. Worm, "Network calculus applied to optimal multimedia smoothing", *in Proc. INFOCOM '01,* vol.3, pages 1474-1483, Apr. 2001.

[12] T. P. Chen, T. Chen, "Adoptive joint source-channel coding using rate shaping", *in Proc. ICASSP '02,* Vol.2, pages 1985-1988, May 2002.

[13] T. P. Chen, T. Chen, "Fine-grained rate shaping for video streaming over wireless networks", *in Proc. ICASSP '03,* Vol.5, pages 688-691, Apr. 2003.

[14] T. P. Chen, T. Chen, "Error concealment aware rate shaping for wireless video transport", *EURASIP Journal on applied signal processing, special issue',* pages 889-905, 2003.

[15] M. Quirk, "Diversity coding for communication systems", *Bell labs. Engineer's notes,* Dec. 1979.

[16] E. N. Gilbert, "Capacity of burst noise channel", *Bell sys. Tech. J.,* Vol.39, No.9, pages 1253-1265, Sept. 1960.

[17] P. Johansson, R. Forchheimer, "Optimal source coding rate for limited buffer size", *Technical Report* LiTH-ISY-R-2776, Feb. 2007 available at http://www.icg.isy.liu.se/publications/en/Techreports.html.

[18] M. Hassan, L. Atzori, M. Krunz, "Video transport over wireless channels: a cycle-based approach for rate control", *in Proc. of 12th annual ACM international conference on multimedia 04,* pages 916-923, 2004.

[19] L. Atzori, M. Krunz, M. Hassan, "Cycle-based rate control for one-way and interactive video communications over wireless channels", *IEEE transactions on multimedia,* vol.9, pages 176-184, Jan. 2007.

[20] Carta M.T., Onali T., Atzori L., "Window-based rate control approach for video streaming over wireless networks", *in proc. ICC '03,* pages 1772-1777, June 2007.

Muhammad Ajmal is currently a Ph.D student at ISY department, Linköping University, Sweden. He received his Master degree in Electrical Engineering from Blekinge Institute of Technology, Sweden in 2009. While he did his Bachelor in Electrical Engineering from UET Peshawar, Pakistan in 2005. His research interests include overlay optical networks and energy-efficient design of WDM networks.

Peter Johansson received his M.Sc degree in Electrical Engineering at Linköping university in 1992. He received his Licentiate of Engineering in Applied Mathematics at Linköping university in 1997, and joined the Image Coding Group in 1999. His research activities include video coding with dense motion field estimation and IP-video.

Robert Forchheimer received the M.S. degree in electrical engineering from the Royal Institute of Technology, Stockholm in 1972 and the Ph.D degree from Linköping University in 1979. During the academic year 1979 to 1980 he was a visiting research scientist at University of Southern California. Since 1981 he has been with Linköping University. He has further worked at The Royal Institute of Technology in Stockholm and at the University of Hannover, Germany.

Forchheimer's research areas have involved data security, packet radio transmission, optical computing, integrated circuits for image processing, image coding, organic electronics and bioinformatics. He has authored or coauthored more than 160 papers in these areas and also holds 10 patents. He is the cofounder of five companies employing more than 500 people. Forchheimer is currently in charge of the Information Coding Group at Linköping University. There, his main work concerns algorithms and systems for networked real-time communication, organic electronics and bioinformatics.