Optimal Multicast Feedback

Jörg Nonnenmacher Ernst W. Biersack
Institut EURECOM
B.P. 193, 06904 Sophia Antipolis, FRANCE
{nonnen,erbi}@eurecom.fr

Abstract

We investigate the scalability of feedback in multicast communication and propose a new method of probabilistic feedback based on exponentially distributed timers. By analysis and simulation for up to 10^6 receivers we show that feedback implosion is avoided with good latency performance obtained. The mechanism is robust against the loss of feedback messages and robust against homogeneous and heterogeneous delays. We apply the feedback mechanism to reliable multicast and compare it to existing timer-based feedback schemes. Our mechanism achieves lower NAK latency for the same performance in NAK suppression. It is scalable, the amount of state at every group member is independent of the number of receivers. No topological information of the network is used and data delivery is the only support required from the network. It adapts to the number of receivers and leads therefore to a constant performance for implosion avoidance and feedback latency.

Keywords: Feedback, Multicast, Reliable Multicast, Performance Evaluation, Extreme Value Theory.

1 Introduction

With the deployment of Multicast in the Internet and with the increasing number of satellites multicast communication is gaining in importance. A major challenge in multicast communication is the *feedback implosion* that occurs when a large number of receivers sends feedback to the sender.

In this paper we investigate feedback of groups from 1 up to 10^6 receivers towards a single sender, as needed for:

- *Reliable multicast*: Reliable multicast guarantees the delivery of data from the sender to every receiver. Feedback messages (FBMs) are needed in order to signal the loss (NAK), or the reception of data (ACK).
- *Estimation of the number of receivers*: is required to (i) perform fair congestion control between multicast and unicast, (ii) and stop transmission, when no receivers are listening, (iii) adapt scalable protocols to the number of receivers, e.g. by adjusting the amount of FEC [1], or to adjust the period of periodic control message sending.

The amount of potential feedback increases linearly with the number of receivers and leads to a high traffic concentration at the sender, wasted bandwidth, and high processing requirements. Feedback implosion imposes high requirements to the mechanism for *feedback implosion* avoidance. Several solutions exist for implosion avoidance based on hierarchies, timers, tokens and probing (see section 5 on related work).

Very little work [2, 3] was done on the analysis of timer-based schemes for multicast feedback. We give the analytical foundation of timer-based feedback, where the timer choice, the sender-receiver delays and the delays between receivers can be modeled by arbitrary distributions. The analysis allows to compute:

- The expected number $E(X)$ of FBMs returned to the sender.
- The expected feedback delay $E(M)$ due to the timers.

We propose a new probabilistic feedback method for multicast based on exponentially distributed timers and show by analysis and simulation for up to 10^6 receivers that feedback implosion is avoided. We show the robustness of our mechanism to loss of FBMs, to homogeneous delays, and to heterogeneous delays.

We further evaluate our mechanism in the context of reliable multicast with respect to NAK implosion avoidance and with respect to the NAK latency. A comparison of our mechanism to existing timer-based feedback schemes shows that the feedback latency of our mechanism is lower, for the same performance in NAK suppression.

Our mechanism requires very little state and has a low computational complexity at every receiver – independent of the group size. No knowledge about the network topology, nor support from the network is required to allow for implosion avoidance.

Using an estimate of the number R of receivers, our feedback mechanism allows to adjust the average number of returned FBMs to any value > 1 via a tradeoff with feedback latency. In order to estimate the number of receivers two methods are compared: one based on probing, the other based on the feedback of our scheme.

The remaining part of the paper is organized as follows. In section 2 the analysis for timer-based feedback schemes is given. In section 3 the performance is evaluated for reliable multicast feedback. Section 4 shows the robustness of timer-based feedback for loss and heterogeneous delays. Section 5 discusses the work in the context of related work and section 6 concludes the work.

2 Timer-based Feedback

Consider the case where a sender needs to receive at least one FBM from R receivers and where the total number of returned FBMs should be as small as possible in order to avoid *feedback implosion*.

sion when every receiver delays its feedback sending by a random time. A receiver that receives a FBM of another receiver can suppress its own feedback sending, referred to as *feedback suppression*.

Our timer-based feedback mechanism works as follows:

1. The sender multicasts a **request for feedback** (I, λ, T) to the R receivers. I is the identification for the feedback round.

2. Receiver i, receives the **request** at time d_i and schedules a **exponentially distributed timer** z_i in the interval $[0, T]$. The parameter for the truncated exponential distribution is λ. When the timer expires receiver i:

   - sends the feedback message FBM(I, z_i) back to the sender if no other FBM(I, z_j) was received by i.

   - suppresses its feedback, if a FBM(I, z_j) of some other receiver j was received before (see figure 1 for the suppression of i's feedback); this requires that j sends its feedback earlier than i and that the delay $d_{i,j}$ between receiver i and receiver j is such that:

$$d_i + z_i > d_j + z_j + d_{i,j}$$

3. On the receipt of the FBMs, the sender computes an estimate $\hat{R}$, for the number of receivers, using the knowledge about the timer settings of all receivers i that returned feedback: z_i, λ, T (see [4] for details).

4. The sender computes T and λ for the next **request for feedback** based on $\hat{R}$ and its requirement for the tradeoff between feedback latency and the mean number of FBMs it wants to receive.

Please note that feedback suppression is possible when all receivers are connected to the sender via a multicast feedback channel, but feedback suppression is possible also in the case where receivers return feedback via unicast as long as the sender multicasts the information about the received feedback to all receivers.

The SRM protocol [2] uses a similar mechanism for the sending of NAKs, with two differences: First, SRM uses a **uniform distributed timer choice** z_i from an interval that depends on the sender-receiver delay d_i. Second, SRM prevents loss of FBMs by scheduling a second request via an exponential back-off in a larger interval in the future.

In the following we analyze the expected number $E(X)$ of FBMs returned to the sender from R receivers and the expected feedback latency $E(M)$ due to timers, when FBMs are not subject to loss. In section 4 we investigate the performance under loss of FBMs. First we introduce the following random variables:

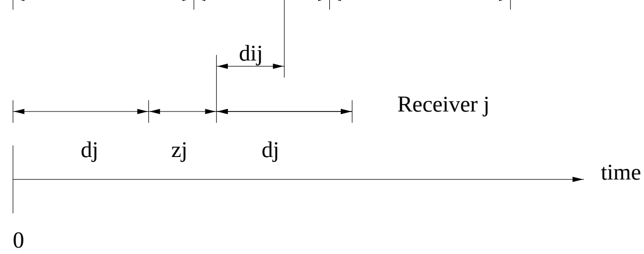

Figure 1: The timing for the feedback and the suppression of receiver i's FBM.

D_i	- one-way delay between the sender and receiver i and vice versa.
Z_i	- time receiver i delays its feedback.
$V_i = D_i + Z_i$	- the time between the sending of the request for feedback and the time the timer expires at i.
$D_{i,j}$	- one–way delay between receiver i and receiver j and vice versa.
$W_{i,j} = V_j + D_{i,j}$	- time between the sending of the request for feedback and the reception of j's feedback at i.
X_i	- Bernoulli, describes the number of FBMs from receiver i.
$X = \sum_{i=1}^{R} X_i$	- total number of FBMs received at the sender from the group of receivers.

Let the delay d_i of receiver i to the sender and the delay $d_{i,j}$ between two receivers i, j be described by the densities $f_{D_i}(d_i)$ and $f_{D_{i,j}}(d_{i,j})$. Different timer choices, also timer choices dependent on the source-receiver delay d_i, can be compared in their performance, when the density for the timer choice is kept general:

$$f_{Z_i|D_i}(z_i|d_i) \tag{1}$$

Then, the density of $V_i = D_i + Z_i$ can be calculated by a transform changing variables [5, ch. 6.3], resulting in:

$$f_{V_i}(v_i) = \int_{-\infty}^{\infty} f_{D_i}(s_i) \cdot f_{Z_i|D_i}(v_i - s_i|s_i) ds_i \tag{2}$$

The same way the density of $W_{i,j} = D_{i,j} + V_i$ can be derived. Since $D_{i,j}$ and V_i are independent the joint density is given by:

$$f_{D_{i,j},V_i}(d_{i,j}, v_i) = f_{D_{i,j}}(d_{i,j}) \cdot f_{V_i}(v_i)$$

Such that the density of $W_{i,j}$ using the transform in [5, ch. 6.3] is given by:

$$f_{W_{i,j}}(w_{i,j}) = \int_{-\infty}^{\infty} f_{D_{i,j},V_i}(s_{i,j}, w_{i,j} - s_{i,j}) ds_{i,j} \tag{3}$$

Since only the first timer setting is considered, the Bernoulli random variable X_i describes, whether **the** FBM from receiver i is sent ($X_i = 1$) or not. Receiver i sends feedback, only when no the probability:

$$P(X_i = 1) = \int_0^\infty f_{V_i}(v_i) \prod_{j=1, i \neq j}^{R} (1 - F_{W_{i,j}}(v_i))\, dv_i \qquad (4)$$

The analysis of the timer settings given above is valid for arbitrary delay distributions of D_i and $D_{i,j}$.

For a better understanding of the timer mechanism and the feedback suppression we will first consider the case, where the delays are homogeneous: All receivers $i = 1, \ldots, R$ have the same delay $d_i = c$ from the sender and the same delay $d_{i,j} = c$ to any other receiver j:

$$f_{D_i}(d_i) = \delta(d_i - c) \qquad f_{D_{i,j}}(d_{i,j}) = \delta(d_{i,j} - c) \qquad (5)$$

In section 4.1 we analyze the timer mechanism for heterogeneous delays.

Furtheron we consider the case, where **all** receivers $i = 1, \ldots R$ choose a timer out of an interval $[0, T]$ – independent of the delay d_i between sender and receiver:

$$f_{Z_i|D_i}(z_i|d_i) = f_{Z_i}(z_i) \qquad , z_i \in [0, T] \qquad (6)$$

We are especially interested in the minimal timer, which is the one expiring first. Let $M = \min_{i=1}^{R}\{Z_i\}$ be the random variable describing the minimal timer. Since the Z_i are identically and independently distributed, the distribution of the minimal timer is given by [6, ch 2]:

$$F_M(m) = P(M \leq m) \quad = \quad 1 - (1 - F_{Z_i}(m))^R$$

Our performance measures of the timer mechanisms are:

- **The expected feedback latency** $E(M)$ **due to the timer mechanism**, given by the minimal timer:

$$E(M) = \int_0^T (1 - F_M(m))\, dm \qquad (7)$$

- **The expected number** $E(X)$ **of FBMs** at the sender given as:

$$E(X) = \sum_{i=1}^{R} E(X_i) = RP(X_i = 1) \qquad (8)$$

Using these two performance measures, different distributions for the timer choice are examined in terms of feedback suppression and feedback latency: The **uniform** distribution and the **exponential** distribution. We also investigated the **beta distribution** without performance improvement over the **exponential distribution**, for details see [4].

2.1 Uniform Distributed Timers

A uniform distributed timer choice out of the interval $[0, T]$ of every receiver i is given by the density:

$$f_{Z_i}(z_i) = \begin{cases} \frac{1}{T} & , 0 \leq z_i \leq T \\ 0 & , otherwise \end{cases} \qquad (9)$$

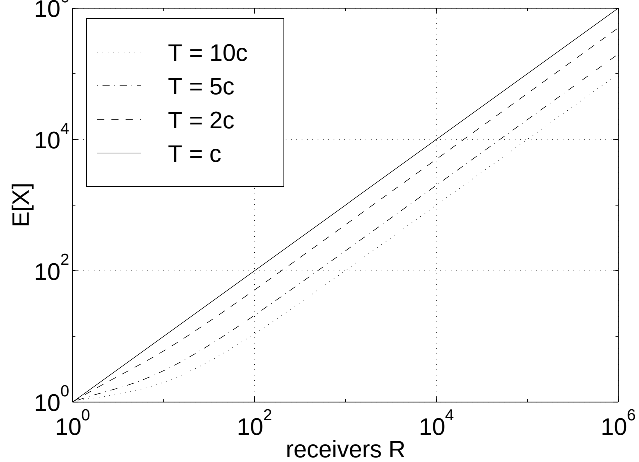

Figure 2: Expected number $E(X)$ of FBMs for **uniform distributed timer** choice from intervals of size $T = c, 2c, 5c, 10c$ for R receivers.

The expected number $E(X)$ of FBMs is:

$$E(X) = \begin{cases} R & , c \geq T > 0 \\ 1 + \frac{c}{T}R - \left(\frac{c}{T}\right)^R & , 0 < c < T \end{cases} \qquad (10)$$

The expected feedback latency $E(M)$ due to the uniform distributed timer choice is:

$$E(M) = \frac{T}{R+1} \qquad (11)$$

Let the interval size T be a multiple of the delay c between receivers. For large numbers R of receivers the expected number of FBMs is $E(X) \approx \frac{c}{T}R$ and thus increases linearly with the number of receivers, see figure 2. The feedback latency (11) on the other hand decreases with R. This means the tradeoff, already reported in [2], exists around $T = Rc$ between suppression and latency.

This is illustrated by figure 3; All receivers set independently a timer in the interval $[0, T]$. All k receivers that set their timer in the interval $[m, m + c]$ will send feedback, the other $R - k$ receivers with timers $z_i > m + c$ will suppress their feedback sending, since the FBM of the receiver with the minimum timer m reaches them before their timer expires.

The **only** way to adapt the feedback mechanism to the number R of receivers is to change the interval size T, which makes the scheme dependent on a correct receiver estimate:

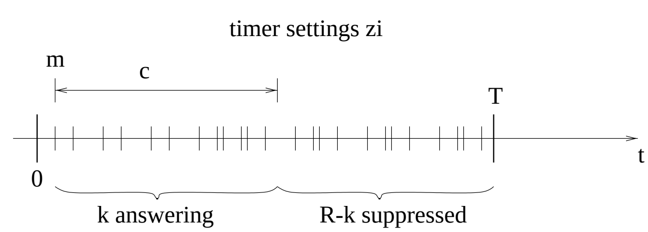

Figure 3: Timer Setting.

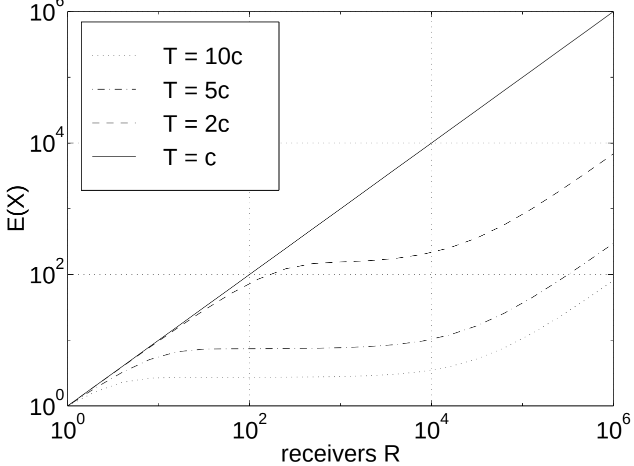

Figure 4: Expected number $E(X)$ of FBMs for **exponentially distributed timer** choice with parameter $\lambda = 10$ from intervals of size $T = c, 2c, 5c, 10c$ for R receivers.

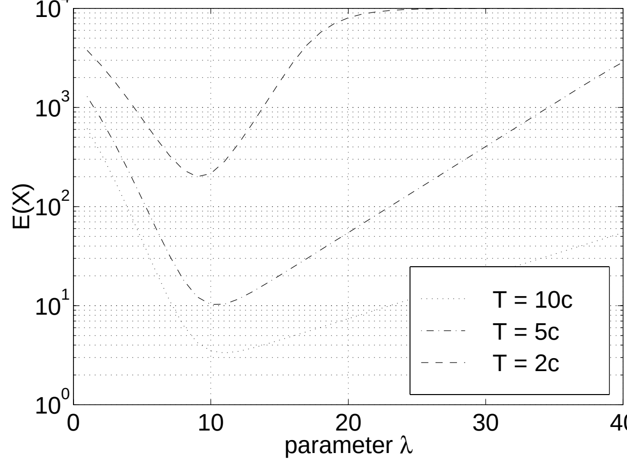

Figure 5: Expected number $E(X)$ of FBMs for **exponentially distributed timer** choice, dependent on parameter λ from intervals of size $T = 2c, 5c, 10c$ for $R = 10^4$ receivers.

- If the number R of receivers is overestimated, the interval size T will be chosen too large and a high feedback latency is encountered.
- If the number R of receivers is underestimated, the small interval size T will lead to a feedback implosion.

An alternative to the **uniform distributed timer choice** and to the intricacies coming with the change of the interval size T is to fix the interval $[0, T]$ and to change the shape of the distribution. Fixing the interval gives a bound on the feedback delay. In order to also achieve a low number of FBMs the minimal timer needs to be separated as far as possible from the mass of the timer settings. Therefore, the following properties are desirable for the density f_{Z_i} for the timer choice:

- The minimal timer is separated from other timers by enabling some few timers to be set on a broad range and by grouping most timer settings on a small range.
- Feedback suppression is not sensitive to errors in the receiver estimate.

We investigate the **exponential distribution** for the timer choice and evaluate its performance.

2.2 Exponentially Distributed Timers

A truncated exponentially distributed timer choice in the interval $[0, T]$ is given by the density:

$$f_{Z_i}(z_i) = \begin{cases} \frac{1}{e^{\lambda}-1} \cdot \frac{\lambda}{T} e^{\frac{\lambda}{T} z_i} & , 0 \le z_i \le T \\ 0, & , otherwise \end{cases} \quad (12)$$

The weight of the density shifts towards T with an increasing λ and results in a dense timer setting at high values. The expected number $E(X)$ of FBMs and the feedback latency are:

$$E(X) = R \qquad , c \ge T > 0$$

$$E(X) = R \frac{e^{\lambda \frac{c}{T}} - 1}{e^{\lambda} - 1} \qquad , 0 < c < T \quad (13)$$

$$- \; e^{\lambda \frac{c}{T}} \left(\left(\frac{1 - e^{-\lambda \frac{c}{T}}}{1 - e^{-\lambda}} \right)^{R} - 1 \right)$$

$$E(M) = T \int_0^1 \left(1 - \frac{e^{\lambda m} - 1}{e^{\lambda} - 1} \right)^{R} dm \quad (14)$$

Figure 4 shows that the exponentially distributed timer choice ($\lambda = 10$) outperforms the uniform distributed timer choice for feedback suppression, compare figure 2. For a wide range of numbers $10 < R < 10^4$ of receivers the expected number $E(X) < 4$ of FBMs stays constant ($T = 10c$).

The impact of the parameter λ on suppression is shown in figure 5. The expected number $E(X)$ of FBMs is convex in λ with a minimum at λ_o. For $\lambda > \lambda_o$ the timer settings, including the minimal timer m, are forced on a narrow range close to T, for $\lambda < \lambda_o$ the minimal timer is close to 0 and the other timers are not well separated from the minimal timer, resulting in feedback implosion in both cases. The latter can also be observed in figure 4 for $R > 10^4$ receivers: $\lambda = 10$ is too small for R receivers, the minimal timer is already close to 0, and an increasing number of receivers will fall in the interval $[m, m + c]$ as indicated by the increasing number $E(X)$ of FBMs.

Further it can be observed from figure 5 that the optimal λ_o is independent of the interval size T, if T is large enough. This allows to determine a λ_o for optimal suppression - dependent only on the number of receivers: $R \mapsto \lambda_o$. By fitting this function a very simple relation is found:

$$\lambda_o = \log(R) + 1 \quad (15)$$

Given the optimal λ_o and the number R of receivers the tradeoff between feedback latency (14) and the number of feedback messages (13) is solely determined by T. Given a desired expected number $E(X) = N$ of feedback messages, T can be adjusted

$$T = 1.2 \frac{\lambda_o}{\log(N)} \quad (16)$$

We can draw the following conclusions for feedback suppression dependent on the distribution:

- Exponentially distributed timers outperform uniform distributed timers for feedback suppression.
- It is possible to control feedback implosion with probabilistic timers by employing a parametric distribution on a small interval of size T.
- Adjusting the parameters λ_o and T to a number of receivers is possible by very simple expressions (15) and (16).

In this section we focused on the feedback suppression. In the next section the performance of the timer schemes for reliable multicast feedback is investigated with a closer look on the feedback latency.

3 Comparison for Reliable Multicast

In reliable multicast communication, negative acknowledgments (NAK) are shown to achieve higher throughput performance than positive acknowledgments (ACK) [7], if retransmissions are multicast. The number of receivers that are potential NAK senders out of a group of R receivers depends on the loss of data packets. For packet loss with probability p at each of the R receivers, the average number of potential NAK senders is $R_l = pR$. However, NAK implosion needs to be avoided for the worst case, where **all** R receivers lose the packet [8].

We examine the timer distributions for:

- The worst case for *NAK implosion*: All R receivers are potential NAK senders.
- The average case for *NAK latency*: R_l receivers are potential NAK senders.

For both cases an interval size T corresponding to R receivers is used to scale the timer distributions. The expected number $E(X)$ of NAKs is calculated for the worst case with R potential NAK senders. The **expected NAK latency** $E(M_p)$ is calculated for the average case, where R is substituted by $R_l = pR$ in $E(M)$ given by equation (7). The **expected NAK latency** $E(M_p)$ is higher than $E(M)$, since the feedback latency increases with a decreasing number R of receivers.

The pair of performance measures (expected NAK latency, expected number of NAKs) = $(E(M_p), E(X))$ is calculated with respect to the interval size T for each distribution. For the uniform distribution $(E(M_p), E(X))$ is uniquely determined by T. For the exponential distribution $E(X)$ (equation (13)) is minimized and the corresponding λ_o is used for the calculation of the expected NAK latency $E(M_p)$.

Figures 6 and figure 7 show the expected NAK latency $E(M_p)$ with respect to the number $E(X)$ of feedback messages in the worst case for $R = 10^2$ and $R = 10^6$ receivers and a packet loss probability of $p = 10^{-2}$.

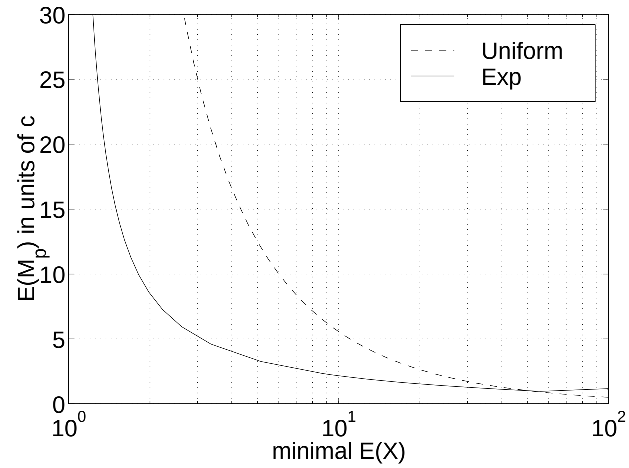

Figure 6: NAK latency $E(M_p)$ for optimal implosion avoidance with $R = 10^2$ receivers.

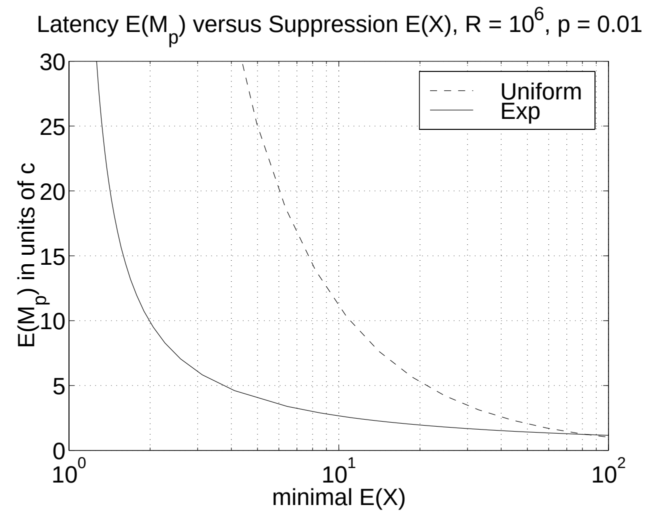

Figure 7: NAK latency $E(M_p)$ for optimal implosion avoidance with $R = 10^6$ receivers.

It can be observed that the **exponential distributed timer choice** outperforms the **uniform distributed timer choice** for reliable multicast with NAK feedback: For the same expected number $E(X)$ of NAKs in the worst case, uniform distributed timers result in higher NAK latencies in the average case.

Figure 7 shows that it is possible to adjust the exponential distribution for $R = 10^6$ receivers such that in the worst case at average 4 NAKs are returned and in the average case, the first NAK is delayed by only 5 one-way delays c.

We showed that the exponential distribution outperforms the uniform distribution for NAK feedback as encountered in reliable multicast and will henceforth just consider the exponentially distributed timer choice.

In the following section the feedback mechanism based on the exponentially distributed timer choice is examined for its robustness to loss and heterogeneous delays.

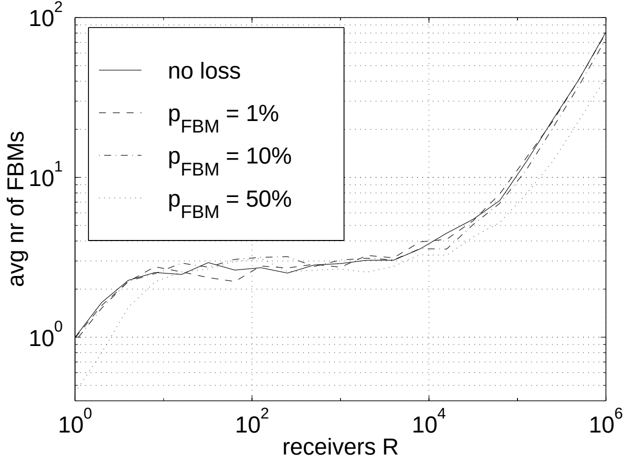

Figure 8: Average number of FBMs with loss p_{FBM}, $\lambda = 10$, $T = 10c$.

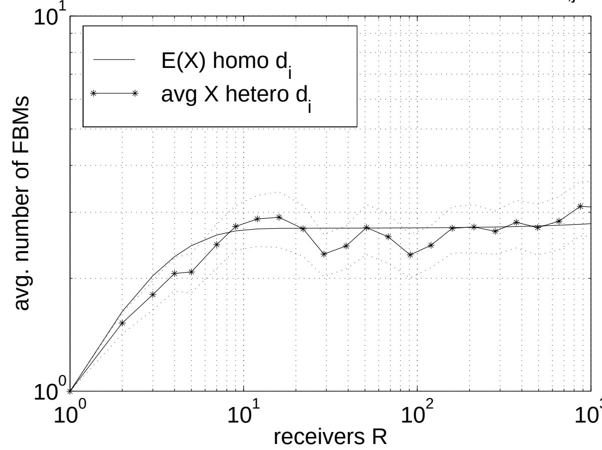

Figure 9: Expected number $E(X)$ of FBMs for heterogeneous sender-receiver delays $d_i \in (0, 2c), d_{i,j} = c$ and interval size $T = 10c, \lambda = 10$.

4 Robustness

The loss of responses can lead to a feedback implosion, since a lost FBM will not suppress other feedback sendings.

Again we consider the worst case, where a FBM is lost directly at the feedback sender and therefore lost for all other receivers. We simulated the feedback 100 times and used parameters $\lambda = 10$ and $T = 10c$ in order to achieve simulation results that correspond to the former analytical results (see figure 4). FBMs were lost with different probabilities $p_{FBM} = 1\%, 10\%, 50\%$ and compared to the case of loss free conditions. Figure 8 shows that the suppression performance of the timer mechanism is not sensitive to loss of FBMs for loss rates up to $p_{FBM} = 10\%$. We experienced a similar robustness also for the average feedback delay. For the very high loss rate of $p_{FBM} = 50\%$ the average number of FBMs is decreased compared to loss free conditions and the average feedback latency is slightly increased.

We can conclude that feedback suppression by exponentially distributed timers is very robust with respect to the loss of FBMs.

4.1 Impact of Heterogeneous Delays

In a network receivers have different delays to the sender and different delays between each other. In order to understand the influence of heterogeneous delays on the timer mechanism, we examine the following two cases:

- Heterogeneous sender-receiver delays d_i, but homogeneous delays $d_{i,j} = c$ between receivers.
- Homogeneous sender-receiver delays $d_i = c$, but heterogeneous delays $d_{i,j}$ between receivers.

Both cases are compared to the case, where the delays between sender and receivers and between receivers are homogeneous $d_{i,j} = d_i = c$.

Heterogeneous delays d_i, or $d_{i,j}$ were distributed in both cases on the interval $[0, 2c]$ via the beta distribution (see [9]) with parameters $a = 2$ and $b = 2$. The interval size for the timer choice was again chosen to be $T = 10c$. This means that the average heterogeneous delay (either $\bar{d}_i = c$, or $\bar{d}_{i,j} = c$) equals the homogeneous delays c.

We simulated the FBM suppression by exponentially distributed timers with $\lambda = 10$ for this heterogeneous case for $R = 1, \ldots, 10^3$ receivers and used 95% confidence intervals.

Heterogeneous delays to the sender

Let us consider the case where the delays between the sender and the receivers are heterogeneous and the delay between any pair of receivers i, j is homogeneous $d_{i,j} = c$.

Figure 9 illustrates that FBM suppression performs better for small groups $R < 10$ in the case of heterogeneous sender-receiver delays, than for homogeneous sender-receiver delays. This is caused by a wider spread of timer settings over $[0, 2c + T]$ due to the heterogeneous reception times d_i of the request for feedback, instead of a more narrow setting in $[c, c + T]$ with homogeneous sender-receiver delays $d_i = c$.

As the group size R increases FBM suppression does not benefit anymore from heterogeneous sender-receiver delays, since the impact of the number R of receivers on the density of the timer settings, and therefore on suppression, is higher than the small difference of the interval sizes.

Heterogeneous delays between receivers

Let us now consider a homogeneous sender-receiver delay $d_i = c$, but heterogeneous delays between receivers $d_{i,j} \in [0, 2c]$. This is, for instance, the case for a satellite distribution, where receivers are additionally connected among each other and to the sender via a terrestrial multicast feedback channel. The sender-receiver delay via the satellite is the same for all receivers (homogeneous $d_i = c$) and the delay between receivers via the terrestrial multicast channel is heterogeneous $d_{i,j}$.

Figure 10 shows that for all numbers R of receivers suppression benefits from heterogeneous delays between receivers. The reason

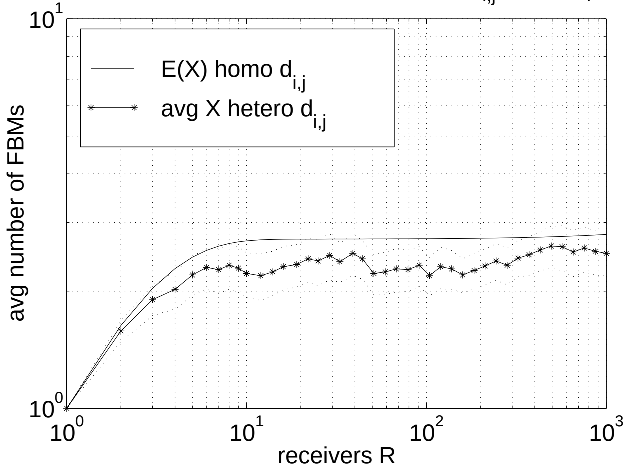

Figure 10: Expected number $E(X)$ of FBMs for heterogeneous inter-receiver delays $d_{i,j} \in (0, 2d)$ and interval size $T = 10c, \lambda = 10$.

is that also feedback sendings other than the one caused by the minimal timer perform suppression, so may e.g. the 3rd smallest be suppressed by the 2nd smallest, leading to a suppression of timers that would cause a FBM sending in the homogeneous case.

From this section we can conclude that feedback suppression by **exponentially distributed timers** is:

- **not** sensitive to loss of feedback messages
- **not** sensitive to heterogeneous delays between sender and receiver
- **not** sensitive to heterogeneous delays between receivers.

All cases are constructive for feedback suppression with probabilistic exponential timers and lead to a better suppression performance.

5 Discussion and Related Work

Ammar defined the feedback problem as response collection via several cost functions [10]. The most research on the the feedback implosion problem was driven by reliable multicast feedback.

There are two major classes of feedback mechanisms for multicast that are solutions to the *feedback implosion* problem:

- *Hierarchical approaches* [11, 12, 13, 14, 15]: Are an inherent solution to the *feedback implosion* and ensure a limited number of FBMs by accumulation/filtering in subgroups.
- *Approaches based on MAC protocols* [16, 17, 2, 3]: The feedback problem in multicast communication is related to the problem of Medium Access Control: The multicast channel constitutes the shared medium and messages sent on the multicast channel are seen by every connected group member. A token mechanism as in token ring is proposed in [16] and random timers with exponential back-off as in CSMA/CD [18] are used in XTP [17] or the SRM protocol [2, 3].

Hierarchical approaches require the expensive setup of the hierarchy of subgroups and can not be employed in a scenario like satellite distribution with unicast backward channels. Approaches based on MAC protocols suffer from scalability problems. Tokens lead to high feedback latencies and random timers in [2, 3] are based on a uniform distribution. SRM [2] exploits heterogeneous delays for a deterministic suppression, but needs a delay estimate $\hat{d}_i$ to the sender. This involves at least one packet sending from every receiver i, resulting for large groups of R receivers in a high amount of control traffic proportional to R. The optimal deterministic timers setting from Grossglauser [20] ensures only one NAK based on the knowledge of the delay and on network support for the timer setting.

Our mechanism does not suffer from any of these problems, since it is a pure end-to-end mechanism. It does not rely on delay estimates to other receivers and state and its complexity is independent of the number of receivers. It does not need any network support except for data delivery and it does not need topological information. It can be employed in any kind of multicast capable network.

Another end-to-end solution based on probabilistic feedback with exponential steps is the probing method of Bolot [19] that proceeds in discrete rounds. Using discrete rounds leads to very good performance for suppression, but a higher feedback latency than using just one round.

6 Conclusions

We investigated probabilistic feedback for multicast groups up to 10^6 receivers by analysis and simulation. Our main results are:

- Exponentially distributed timer settings lead to a lower feedback latency and better feedback suppression than existing schemes based on uniformly distributed timer settings.
- Probabilistic feedback with exponential timers is scalable with the number of receivers and avoids feedback implosion while assuring moderate feedback latency.

Based on these results we proposed a new timer-based feedback scheme that requires very few state, does not need any network support other than data delivery, and adapts to the number of receivers:

- It avoids feedback implosion and feedback arrives fast.
- It is robust under loss of feedback messages.
- It works with heterogeneous as with homogeneous delays between multicast group members and can therefore be employed on nearly any kind of network including satellite-based networks.
- It allows to adjust the parameters dependent on the trade-off between average number of feedback messages returned and the latency for the feedback.

The authors are grateful to Martin Lacher for simulating the mechanism with heterogeneous delays. We also thank Raymond Knopp and the anonymous reviewers for many helpful comments. Eurecom's research is partially supported by its industrial partners: Ascom, Cegetel, France Telecom, Hitachi, IBM France, Motorola, Swisscom, Texas Instruments, and Thomson CSF.

References

[1] J. Nonnenmacher, E. W. Biersack, and Don Towsley, "Parity-based loss recovery for reliable multicast transmission," in *SIGCOMM '97*, Cannes, France, Sept. 1997, pp. 289–300.

[2] S. Floyd, V. Jacobson, C. Liu, S. McCanne, and L. Zhang, "A reliable multicast framework for light-weight sessions and application level framing," *Submitted to IEEE/ACM Transactions on Networking*, 1996.

[3] P. Sharma, D. Estrin, S. Floyd, and V. Jacobson, "Scalable timers for soft state protocols," in *Proc. INFOCOMM 97*, Apr. 1997.

[4] Jörg Nonnenmacher and Ernst Biersack, "Scalable optimal multicast feedback," Tech. Rep., Institute EURECOM, BP 193, 06904 Sophia Antipolis, FRANCE, July 1997.

[5] Athanasios Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, Inc., 3rd edition, 1991.

[6] Enrique Castillo, *Extreme Value Theory in Engineering*, Academic Press, Inc., 1988.

[7] D. Towsley, J. Kurose, and S. Pingali, "A comparison of sender-initiated and receiver-initiated reliable multicast protocols," *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 3, pp. 398–406, 1997.

[8] Maya Yajnik, Jim Kurose, and Don Towsley, "Packet loss correlation in the mbone multicast network," in *Proceedings of IEEE Global Internet*, London, UK, November 1996, IEEE.

[9] Paul E. Pfeiffer, *Probability for Applications*, Springer Verlag, New York, Berlin, Heidelberg, 1990.

[10] Mostafa H. Ammar and George N. Rouskas, "On the performance of protocols for collecting responses over a multiple-access channel," in *Proceedings of INFOCOM'91*. IEEE, 1991, vol. 3.

[11] Rajendra Yavatkar, James Griffoen, and Madhu Sudan, "A reliable dissemination protocol for interactive collaborative applications," in *Proceedings of ACM Multimedia*, San Francisco, CA USA, 1995, ACM, pp. 333–344.

[12] S. Paul, K. K. Sabnani, J. C. Lin, and S. Bhattacharyya, "Reliable multicast transport protocol (rmtp)," *IEEE Journal on Selected Areas in Communications, special issue on Network Support for Multipoint Communication*, vol. 15, no. 3, pp. 407 – 421, April 1997.

cast," in *International Zürich Seminar*, Zürich, Switzerland, February 1996.

[14] J.M. Bonnin and J.J. Pansiot, "A scalable collect," in *HIPPARCH Workshop*, Sweden, June 1997.

[15] Dante DeLucia and Katia Obraczka, "Multicast feedback suppression using representatives," in *IEEE Infocom97*, 1997.

[16] J.M.Chang and N.F.Maxemchuk, "A broadcast protocol for broadcast networks," in *Proceedings of GLOBECOM*, Dec. 1993.

[17] Timothy W. Strayer, Bert J. Dempsey, and Alfred C. Weaver, *XTP - THE XPRESS TRANSFER PROTOCOL*, Addison-Wesley, 1992.

[18] Institute of Electrical and Electronics Engineers, "Carrier sense multiple access with collision detection (csma/cd) access method and physical layer specifications," *IEEE Std 802.3-1985*, 1985.

[19] J.C. Bolot, T. Turletti, and I. Wakeman, "Scalable Feedback Control for Multicast Video Distribution in the Internet," in *Proceedings of SigComm*. ACM, Sept. 1994.

[20] Matthias Grossglauser, "Optimal deterministic timeouts for reliable scalable multicast," *IEEE Journal on Selected Area in Communications*, vol. 15, no. 3, pp. 422–433, April 1997.