Optimization of communication protocols by stochastic delay mechanisms

J. Levendovszky, I. Koncz, P. Boros

Abstract—The paper is concerned with developing stochastic delay mechanisms for efficient multicast protocols and for smooth mobile handover processes which are capable of preserving a given Quality of Service (QoS). In both applications the participating entities (receiver nodes or subscribers) sample a stochastic timer and generate load after a random delay. In this way, the load on the networking resources is evenly distributed which helps to maintain QoS communication. The optimal timer distributions have been sought in different p.d.f. families (e.g. exponential, power law and radial basis function) and the optimal parameter have been found in a recursive manner. Detailed simulations have demonstrated the improvement in performance both in the case of multicast and mobile handover applications.

Keywords—Multicast communication, Stochactic delay mechanisms

I. Introduction

EFFICIENT multicast communication is one of the toughest challenges in packet switched networking and optimal protocols are still under development [5]. One of the difficulties arises from the large amount of signaling information (ACK/NACK) exchanged between the server and users which can overwhelm network capacities [7]. The objective is to develop stochastic timer distributions for generating NACK signals which avoid the misuse of bandwidth, i.e. the NACK signals do not flood the network with overwhelming signaling information. Therefore, the timer distribution should guarantee that the tail distribution of the number of NACKs is under a given threshold.

On the other hand, in mobile communication a large number of simultaneous reactive handovers has a negative impact on the access network performance, i.e. they can cause serious QoS degradation [2], [3], [4]. To circumvent this effect, a stochastic delay mechanism can spread the handover requests in time, resulting in a more balanced network load. To investigate the effect of the random delays on the handover mechanisms we have created a queueing model where the queues are organised into a tree topology. In this model single queues represent the performance of various system components, while the load which is caused by a handover request is modeled by a single packet [7].

The paper provides novel approaches to these problems by introducing a stochastic timer which can smooth out either the load presented by signaling information (ACK/NACK) in the case of multicast protocol or the handover requests presented by mobile users in b3G networks. The optimal timer distribution has been found n the families of exponential and power law distributions based on adaptive optimization. The solution can be applied to the most general networking scenario (heterogeneous networks with arbitrary random link delays).

Budapest University of Technology and Economics, Department of Telecommunications, Magyar tudosok korutja 2, H-1117 Budapest, Hungary, {borosp, levendov}@hit.bme.hu

II. Optimal stochastic delay mechanism for multicast communication

A. The model

Let us assume that there is a graph $G(V, E, d)$, with a node $s \in V$ referred to as "sender" and a subset of other nodes $\{r_j, j = 1, ..., J\} \subset V$ referred to as "receivers". There is a vector $\mathbf{c} = (c_1, ..., c_J)$ characterizing the distances of each receiver from the sender, and a matrix $\mathbf{C}$, the C_{ij} element of which describes the distance between receiver node i and receiver node j.

Upon sending a message, the sender also sends a timer p.d.f. to each receiver denoted by $f(t)$ (or specifies a parameter of a certain density family). When sending feedbacks, the receivers sample this timer p.d.f. and wait accordingly. If no feedback from other nodes arrives during the waiting period, then a feedback is generated on the corresponding node. Otherwise the feedback is suppressed.

In order to formulate the problem, let $X_i \in \{0, 1\}$ denote the random variable expressing whether a feedback is generated on node i ($X_i = 1$) or not ($X_i = 0$). We are concerned with evaluating the distribution of the aggregated number of NACKs $Y := \sum_{j=1}^{J} X_j$. Our endeavour is to develop some optimal timer distributions $f^{(opt)}(t)$ in order to achieve some desired properties of the distribution of Y denoted by P_Y.

One of such desired properties can be given as follows:

$$f^{(opt)}(t) : \max_{f(t)} P(A < Y < B) \tag{1}$$

for a given A and B. Before we delve into solving this problem, we list the possible density families we seek the optimum timer p.d.f. within.

B. The timer distributions

We considered the following timer distributions $f(t, \mathbf{w})$, where $\mathbf{w}$ denotes the free parameter(s) subject to optimization:

- the timer density is selected from the exponential distribution family:

$$f_{exp}(t, \lambda) =$$
$$= \begin{cases} \frac{1}{1-\exp(-\lambda)} \left(\frac{\lambda}{T}\right) \exp\left(-\frac{\lambda}{T}t\right) & 0 \le t \le T \\ 0 & \text{otherwise} \end{cases}$$

where $\mathbf{w} = \lambda$.

- the timer density is selected from the "power-law" distribution family:

$$f_{pow}(t, a) = \begin{cases} \frac{a}{T} \left(\frac{t}{T}\right)^{a-1} & 0 \le t \le T \\ 0 & \text{otherwise} \end{cases}$$

where $\mathbf{w} = a$.

- the timer density is selected from the Radial Basis Function distribution family:

$$f_{RBF}(t, \mathbf{x}, \mathbf{m}, \sigma) =$$

$$= \begin{cases} \dfrac{\sum_{i=1}^{K} x_i e^{-\frac{1}{2}\left(\frac{t-m_i}{\sigma}\right)^2}}{\int_0^T \sum_{i=1}^{K} x_i e^{-\frac{1}{2}\left(\frac{t-m_i}{\sigma}\right)^2} dt} & 0 \le t \le T \\ 0 & \text{otherwise} \end{cases}$$

where $\mathbf{w} = (\mathbf{x}, \mathbf{m}, \sigma)$, K denotes the number of components, x_i $(i = 1, \dots, K)$ is the weight of component i, m_i $(i = 1, \dots, K)$ is the "mean" value of component i, and σ is the common variance.

The RBF family has the advantage that it is a universal approximator in L^2, thus it can capture almost any L^2 measurable density function. The corresponding distribution functions can be expressed as follows:

- Exponential timer:

$$F_{exp} = \begin{cases} 0 & t < 0 \\ \dfrac{1 - \exp\left(-\frac{\lambda}{T} t\right)}{1 - \exp(-\lambda)} & 0 \le t \le T \\ 1 & T < t \end{cases}$$

- Power-law timer:

$$F_{pow} = \begin{cases} 0 & t < 0 \\ \left(\frac{t}{T}\right)^a & 0 \le t \le T \\ 1 & T < t \end{cases}$$

- RBF timer:

$$F_{RBF} =$$

$$= \begin{cases} 0 & t < 0 \\ \dfrac{\int_0^t \sum_{i=1}^{K} x_i e^{-\frac{1}{2}\left(\frac{l-m_i}{\sigma}\right)^2} dl}{\int_0^T \sum_{i=1}^{K} x_i e^{-\frac{1}{2}\left(\frac{l-m_i}{\sigma}\right)^2} dl} & 0 \le t \le T \\ 1 & T < t \end{cases}$$

C. Timer optimization assuming deterministic homogeneous delays

A homogeneous network is characterized by a graph with uniform link delays denoted by c. Now, with the distributions at hand, $P(X_i = 1)$ can be calculated based on the following expression:

$$p(\mathbf{w}) := P(X_i = 1) = \qquad (2)$$

$$= \int_c^{c+T} f(t - c, \mathbf{w}) \prod_{j=1, j \ne i}^{J} (1 - F(t - 2c, \mathbf{w})) dt.$$

It is clear that $Y = \sum_{j=1}^{J} X_j$ follows a binomial distribution, thus

$$P(A \le Y \le B) = \sum_{l=A}^{B} \begin{pmatrix} J \\ l \end{pmatrix} p(\mathbf{w})^l (1 - p(\mathbf{w}))^{J-l}.$$

Since p is a function of vector $\mathbf{w}$, therefore our objective is to find

$$\mathbf{w}_{opt} : \max_{\mathbf{w}} \Psi(\mathbf{w}), \qquad (3)$$

where $\Psi(\mathbf{w}) = P(A \le Y \le B)$. This means that we have to solve the following optimization problem for free parameters of the timer distribution functions:

$$\mathbf{w}_{opt} : \max_{\mathbf{w}} \Psi(\mathbf{w}) = \max_{\mathbf{w}} P\left(A \le \sum_{i=1}^{J} X_i \le B\right) =$$

$$\max_{\mathbf{w}} \sum_{l=A}^{B} \begin{pmatrix} J \\ l \end{pmatrix} p(\mathbf{w})^l (1 - p(\mathbf{w}))^{J-l}.$$

Due to the fact that the binomial distribution is differentiable with respect to $p(\mathbf{w})$ and $p(\mathbf{w})$ is differentiable with respect to $\mathbf{w}$, the optimization problem set forth by (3) can be solved by the gradient search method:

$$\mathbf{w_{k+1}} = \mathbf{w_k} + \delta_k grad\Psi(\mathbf{w_k}),$$

where δ_k is referred to as a relaxation parameter.

D. Timer optimization in the case of random homogeneous delays

As in a real communication network the link delays are random, parameter c is assumed to be a random variable with density $g(c)$. This case is referred to as "random and homogeneous" as each delay in the graph changes subject to the same random variable. This random variable is supposed to follow either a Gaussian or a uniform distribution. We characterized both of these distributions by their mean and variances. Therefore our objective function can be expressed as follows:

$$\Psi(\mathbf{w}|c) = P(A \le Y \le B|c)$$

$$\Psi(\mathbf{w}) = E_c \Psi(\mathbf{w}|c) = \int \Psi(\mathbf{w}, c) g(c) dc,$$

where $g(c) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(c-m)^2}{2\sigma^2}}$ in the case of a normal distribution and $g(c) = \frac{1}{U-L}$ in the case of a uniform distribution, where $L = m - \frac{\sqrt{12}\sigma}{2}$ and $U = m + \frac{\sqrt{12}\sigma}{2}$.

Now we have to perform the following optimization problem based on (3):

- Normal distribution:

$$\Psi(\mathbf{w}) = K \int_0^T \Psi(\mathbf{w}|c) \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(c-m)^2}{2\sigma^2}} dc,$$

where $K = \frac{1}{\int_0^T \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(c-m)^2}{2\sigma^2}} dc}$.

- Uniform distribution:

$$\Psi(\mathbf{w}) = Z \int_{\max(0,L)}^{\min(T,U)} \Psi(\mathbf{w}|c) \frac{1}{U-L} dc$$

where $Z = \frac{1}{\int_{\max(0,L)}^{\min(T,U)} \frac{1}{U-L}} = \frac{U-L}{\min(T,U) - \max(0,L)}$.

As a result, it can be proven that the obtained $\Psi(\mathbf{w})$ is differentiable again with respect to $\mathbf{w}$ which means that $\mathbf{w}_{opt} : \max_{\mathbf{w}} \Psi(\mathbf{w})$ can be solved by gradient search:

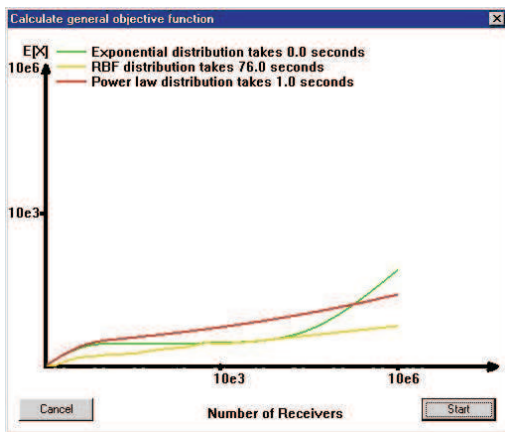$$\mathbf{w_{k+1}} = \mathbf{w_k} + \delta_k grad\Psi(\mathbf{w_k}).$$

Fig. 1. Expected number of feedback messages dependent on the number of receivers.

Fig. 2. Probability of aggregated NACK message numbers ($P(0 \leq Y \leq 7)$) for several timer distributions in the case of 15 receivers with deterministic and homogenous(a) , and non homogenous link delay

E. Numerical results

We have investigated the average number of NACKs E(Y) as a function of the number of recievers in the case of homogenous deterministic delays. The length of the delay interval was $T = 10c$, while the parameter of the stochastic timer distribution was set to $\lambda = 10$ in the case of exponential; and $a = 10$ in the case of power law distribution. The obtained results are depicted by Figure (1). The simulations were carried out for exponential, power law and RBF distribution.

It can be inferred that the number of NACKS can further be decreased by using RBF. Furthermore, one can investigate the probability of the number of NACKs falling into a pre-determined interval. The corresponding results are indicated by Figure 2. One can see that in the case of exponential delay distribution the optimal λ parameter (which maximizes the probability that the number of NACKs falls into the region $0 \leq Y \leq 7$) is $\lambda_{opt} = 2.21$, while in the case of power-law distribution the optimal value is $a_{opt} = 1.55$. The maximal probability that the number of NACKs falls into the region $0 \leq Y \leq 7$ is 0.661, while in the case of power-law distribution the maximal probability is 0.416. The RBF function gives the best result, the optimal probability is 0.981.

Secondly, numerical results have been calculated for the case of inhomogeneous and deterministic link delays by using exponential, power-law and RBF timer distributions. The corresponding numerical results are shown in figure 2.

One can see that in the case of exponential delay distribution the optimal λ parameter (which maximizes the probability that the number of NACKs falls into the region $0 \leq Y \leq 7$) is $\lambda_{opt} = 1.64$, while in the case of power-law distribution the optimal value is $a_{opt} = 1.45$. The maximal probability that the number of NACKs falls into the region $0 \leq Y \leq 7$ is 0.683, while in the case of power-law distribution the maximal probability is 0.611. The RBF function gives the best result, the optimal probability is 0.823.

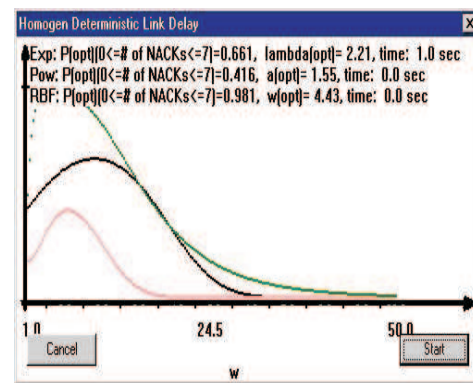In the case of homogeneous but random link delays, we tested a network which included 15 receivers too, and the maximal timer delay value was set $T = 10$. Furthermore, we assumed that the link delays are subject to a uniform p.d.f. Based on the optimized **w** parameters, the obtained performance is indicated in Figure 3, where the probability $P(0 \leq Y \leq 7)$ is plotted against $E(Y)$.
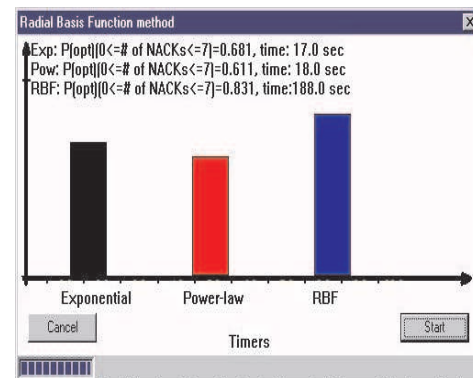

Fig. 3. Probability of aggregated NACK message numbers ($P(0 \leq Y \leq 7)$) for several timer distributions in the function of 'mean' parameter, and in the case of RBF method

From this figure, one can see that for lower mean delay values the timer distributions yield nearly the same performance, while in the case of higher mean delay values the RBF distribution gives a much better performance. This comes down to the extraordinary approximation capabilities of RBF. In the case of inhomogeneous random link delays we tested the methods with exponential, power-law and RBF timer distributions on the same network topology. The possible link delays were $c_{-1} = 0.3$ msec, $c_0 = 1.1$ msec and $c_1 = 1.9$ while the the average delay was $E(c_0) = 0.99$). The maximal timer delay value was set $T = 10$. Based on the optimized **w** parameters, by using the RBF method borrowed from reliability analysis the obtained performance is shown by Figure 3.

III. STOCHASTIC DELAY MECHANISM FOR SMOOTH HANDOVER

The rapidly growing bandwidth demand in mobile applications accelerated the development of novel mobile communication technologies leading to the concept of Beyond-3G Networks (B3G). This concept is fully based on the Internet Protocol (IP) and is not restricted to any specific access technology or mobility protocol. Therefore, the most important properties of the B3G networks are mobility support, high throughput, and Quality of Service (QoS) provisions.

To ensure QoS communication, upon a handover request the mobile nodes sample a stochastic timer and handover is initiated only after the sample expires. In this way, the handover requests are spread in time so that each one can be processed accordingly and the radio link will not be congested at all. As a consequence, QoS can be maintained and more effective resource utilization becomes possible.

A. The model

Based on the general properties of the B3G networks a queueing network model is constructed to find an optimal handover strategy. This model takes the additional processing load which is generated by a handover which appears at different levels of network hierarchy into account. As a result, the stochastic timer described by a p.d.f. will determine the input distribution of a hierarchical queueing network. Our analysis aims at deriving the QoS parameters (Packet Loss Probability and Mean Packet Delay) based on the stationary distribution of this queueing network. In this way, the analytical relationship between the timer p.d.f. and QoS measures can be revealed and the parameters of the timer p.d.f. can be subject to optimization for guaranteeing smooth handover in terms of optimal QoS parameters.

The p.d.f. of the stochastic timer introduced to avoid large load caused by simultaneous handovers is denoted by $f(t)$. The investigated p.d.f is the truncated exponential described in Section 1. Users initiating handover sample this stochastic timer and the handover will commence if the random time provided by the stochastic timer has elapsed. [7], [5], [6]

B. Modeling the input process generated by the handover

Let $p_k = \int_{t_k}^{t_{k+1}} f(t)\,dt$ denote the probability that the sampled timer expires in the time interval in the $[t_k; t_{k+1}]$. The input process generated by the handover is denoted by $Y(k)$ has the following distribution

$$P(Y(k) = n | N(k) = m) = \qquad (4)$$

$$= \binom{m}{n} p_k^n (1 - p_k)^{m-n}, \qquad (5)$$

where

$$P(N(k) = m) = \qquad (6)$$

$$= \binom{N(0)}{m} \left(\int_0^{k\Delta} f(t)\,dt \right)^m * \qquad (7)$$

$$\left(1 - \int_0^{k\Delta} f(t)\,dt \right)^{N(0)-m}. \qquad (8)$$

C. Handover processing in queueing scenarios - the single queue approach

A G/D/1 queueing model can be used to represent the handover process, where the queue length fulfils the following stochastic differential equation:

$$q(k+1) = \lceil q(k) - 1 \rceil^+ + Y(k+1). \qquad (9)$$

Here $q(k)$ stands for the number of waiting requests in the queue with a length of L packets and $Y(k)$ denotes the number of arriving requests at time k. $Y(k)$ is a stochastic variable.

The stationary distribution of the queue length $\overline{\pi}$ can be calculated by solving the equation $\overline{\pi} = \mathbf{P}\overline{\pi}$. Based on $\overline{\pi}$ the QoS parameters (cell loss probability and the mean cell delay) can be calculated as

$$P_{cell\,loss} = \frac{\sum_{k=1}^{\infty} k \sum_{l=0}^{L} \pi_l P_{L-l+k+1}}{\sum_{i=1}^{M} n_i m_i} \qquad (10)$$

and

$$\mathbf{E}(q) = \sum_{k=1}^{L} k\pi_k. \qquad (11)$$

We assume that no packets are waiting at time 0: $\pi(k) = [\pi_0(k), \pi_1(k), \dots, \pi_L(k)] = [1, 0, \dots, 0]$. [1] Thus, $\pi(1)$ can be written as

$$\pi(1) = \overline{\pi}(0)\mathbf{P}(0), \qquad (12)$$

and consequently,

$$\pi(k) = \overline{\pi}(0) \prod_{j=0}^{k-1} P(k). \qquad (13)$$

Our goal is to find an optimal stochastic timer distribution $f_{opt}^{(W)}(t)$ for which

$$f_{opt}^{(W)}(t) : \frac{min}{W} \left(P_{cell\,loss}(k), \mathbf{E}^{(q)}(k) \right), \qquad (14)$$

$$\forall k \in [0, \dots, J]$$

We seek the optimal timer distribution in the family of truncated exponential and a Radial Basis functions, which yield a one- or multi-dimensional optimization problem, respectively.

D. Handling handovers in a hierarchical network structure

Queues are connected according to a tree topology. In the following example a two-level queueing system is described (see figure 4) where two G/D/1/L queues are connected to a third one. A handover request can enter the system on the lower level and it is considered to be successful if it leaves the upper and lower buffer with minimal delay.

$$q_1(k+1) = \lceil q_1(k) - 1 \rceil^+ + Y_1(k+1) \qquad (15)$$

$$q_2(k+1) = \lceil q_2(k) - 1 \rceil^+ + Y_2(k+1) \qquad (16)$$

$$Q(k+1) = \lceil Q(k) - 1 \rceil^+ + U(k+1) \qquad (17)$$

$$U(k) =$$
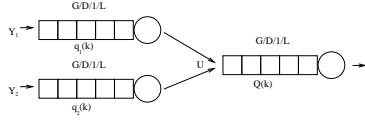
Fig. 4. A two-level queueing system

$$labeleq : U - k = \begin{cases} 2 \ if \ q_1(k) \neq 0 \cup q_2(k) \neq 0 \\ 1 \ if \ q_1(k) = 0 \cup q_2(k) \neq 0 \\ ;; |q_1(k) \neq 0 \cup q_2(k) = 0 \\ 0 \ if \ q_1(k) = 0 \cup q_2(k) = 0 \end{cases} \quad (18)$$

Considering that Y_1 and Y_2 are independent, it can be written:

$$P\left(U\left(k\right) = 2\right) \quad = \quad (19)$$

$$= P(q_1(k) > 0)P(q_2(k) > 0) = \quad (20)$$

$$= (1 - \pi_0^{(1)})(1 - \pi_0^{(2)}) \quad (21)$$

$$P(U(k) \quad = \quad 1) = (1 - \pi_0^{(1)})\pi_0^{(2)} + \quad (22)$$

$$= (1 - \pi_0^{(2)})\pi_0^{(1)} \quad (23)$$

$$P(U(k) = 0) = \pi_0^{(1)}\pi_0^{(2)} \quad (24)$$

Our goal is to find the optimal $f_{opt}^{(\overline{W})}$ stochastic timer distributions so that

$$f_{opt}^{(\overline{W})}(t) : \frac{min}{\overline{W}} \quad (P_{cell \ loss}(k), \mathbf{E}(k)), \quad (25)$$

$$\forall k \in [0, \dots, J]$$

where based on equation (10) and (11):

$$P_{cell\_loss}(k) \quad = \quad (26)$$

$$= \sum_{q=1}^{3} P_{cell\_loss}^{(q)}(k) + \quad (27)$$

$$+ P_{cell\_loss}^{(1)}(k)P_{cell\_loss}^{(3)}(k) \quad + \quad (28)$$

$$+ P_{cell\_loss}^{(2)}(k)P_{cell\_loss}^{(3)}(k), \quad (29)$$

$$\mathbf{E}(k) = max(\mathbf{E}^{(1)}(k)+ \quad (30)$$

$$\mathbf{E}^{(3)}(k), \mathbf{E}^{(2)}(k) + \mathbf{E}^{(3)}(k))$$

$$\forall k \in [0, \dots, J]$$

Since equation (14) and (25) has transformed the problem of finding the proper timer distribution into a multidimensional optimization problem, here we focus on the numerical approaches of optimization.

To find an optimal $f^{(\overline{W})}(t)$ we used exhaustive search on the discretised timer distribution. The maximum allowable loss probability was 0.05.
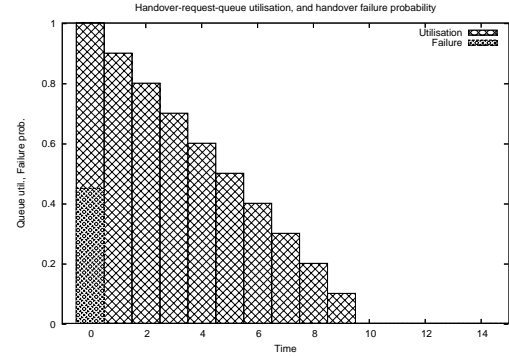

Fig. 5. Queue utilization and handover failure probability when no stochastic timer is used
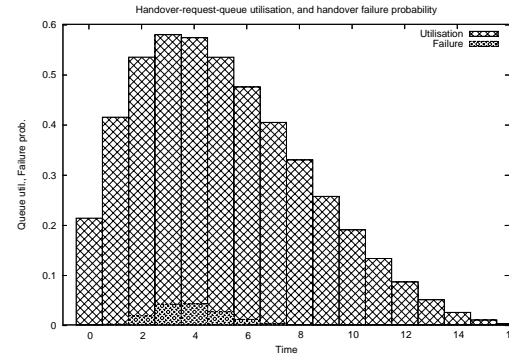

Fig. 6. Queue utilization and handover failure probability when using an exponential timer distribution

E. Numerical results

For performance analysis, the following parameters were used: queue length is 10 requests, service time is 1 request per unit time, the number of mobile nodes is 20. In a working network configuration the service time is in the range of 0.1-1 ms, depending on the performance of the components. Based on this if we transform our values for a real configuration then we can say that a handover is successful only if its serving latency is below 2-20 ms. If the delay is more than this (the queue runs over) the handover is considered to be unsuccessful, and handover request retransmission occurs.
In figure 5 the handover-request-queue utilization and handover failure probability is shown when no stochastic timer is used.

9 out of 20 requests are dropped, this indicates that a delay mechanism should be used to lower the load on the network. In figure 6 the effect of an exponential timer distribution with parameter $\lambda = 0.5$ is shown. It can be seen, that the overall utilization is much better now and the failure ratio stays on a moderate level as well.

In figure 7 the queue utilization and cell loss probability is shown when using near-optimal timer distribution for a single queue model. The utilization in the early phase is high, and is decreasing with time. The cell loss probability is acceptable, as well.

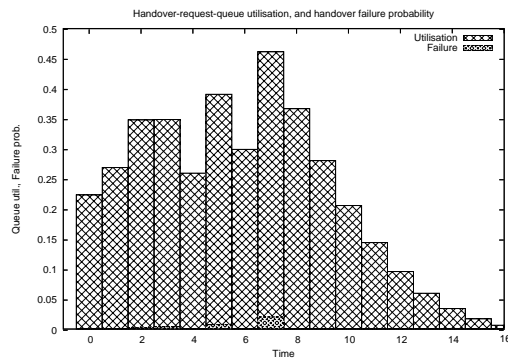In figure 8 the used stochastic timer distributions are shown.

Fig. 7. Queue utilization and handover failure probability when using near-optimal timer distribution
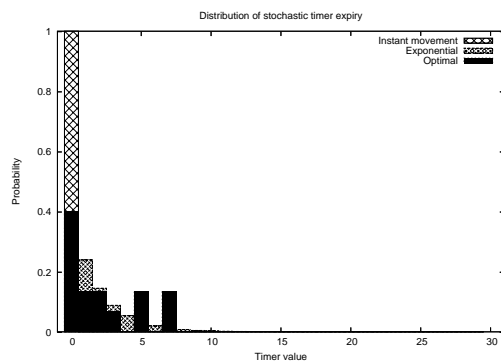

Fig. 8. Stochastic timer distributions

IV. CONCLUSIONS

In this paper novel stochastic delay mechanisms have been developed for optimal multicast and handover performance. In both applications the participating entities sample a stochastic timer and generate load after a random delay. In this way, the load on the networking resources is evenly distributed which helps to maintain QoS communication. The optimal timer distributions have been sought in different p.d.f. families (e.g. exponential, power law and radial basis function) and the optimal parameter have been found in a recursive manner. Detailed simulations have demonstrated the improvement in performance both in the case of multicast and mobile handover applications.

V. ACKNOWLEDGEMENT

REFERENCES

[1] L. Kleinrock, *"Queueing Systems Vol. I.: Theory"*
[2] D. Johnson, C. Perkins, J. Arkko, *"Mobility Support in IPv6", draft-ietf-mobileip-ipv6-24.txt*
[3] H. Soliman, C. Castelluccia, K. El-Malki, L. Bellier, *"Hierarchical Mobile IPv6 mobility management (HMIPv6)", draft-ietf-mobileip-hmipv6-08.txt*
[4] IST-1999-10050 BRAIN project, *"BRAIN architecture specifications and models, BRAIN functionality and protocol specification D2.2"*, www.ist-brain.org
[5] J. Nonmacher, E.W. Biersack, *Scalable Feedback for Large Groups, IEEE/ACM Trans. Networking*, vol. 7., pp 375-386, June 1999.
[6] P. Sharma, D. Estrin, S. Floyd, and V. Jacobson, *Scalable timers for soft state protocols*, presented at INFOCOM , Kobe, Japan.
[7] Levendovszky J., David, T., van der Meulen, E.C., *Stochastic feedback mechanisms in multicast communications*, 23th Symposium on Information and Communication Theory in the Benelux, Lovain-la-Neuve, the Netherlands, May, 2002.