Optimal Retrial and Timeout Strategies for Accessing Network Resources

Lavy Libman and Ariel Orda
Department of Electrical Engineering
Technion – Israel Institute of Technology
Haifa, Israel 32000
{*libman@tx, ariel@ee*}*.technion.ac.il*

Abstract—**The notion of *timeout* (namely, the maximal time to wait before retrying an action) turns up in many networking contexts, such as packet transmission, connection establishment, etc. Usage of timeouts is encountered especially in large-scale networks, where negative acknowledgments (NACKs) on failures have significantly higher delays than positive acknowledgments (ACKs), and frequently are not employed at all. Selection of a proper timeout involves a tradeoff between waiting too long and loading the network needlessly by waiting too little. The common approach is to set the timeout to a large value, such that, unless the action fails, it is acknowledged within the timeout duration with a high probability. This approach is conservative and leads to overly long, far from optimal, timeouts.**

We take a quantitative approach with the purpose of computing and studying the optimal timeout strategy. The above tradeoff is modeled by introducing a 'cost' per unit time (until success) and a 'cost' per repeated attempt. The optimal timeout strategy is then defined as one that a selfish user would follow to minimize its expected cost. We discuss the various practical interpretations that these costs may have. We then derive the formulae for the optimal timeout values and study some of their fundamental properties. In particular, we identify the conditions for making parallel attempts from the outset to be worthwhile. In addition, we demonstrate a striking property of *positive feedback*. This motivates us to study the interaction resulting when many users selfishly apply the optimal timeout strategy; specifically, we use a noncooperative game model, and show that it suffers from an inherent instability problem. Some implications of these results on network design are discussed.

Keywords— **Timeout strategy, packet retransmission, connection establishment, noncooperative games, network pricing.**

I. Introduction

A generic issue that turns up in many networking contexts is that of *retrial*. In a typical scenario, an agent needs to accomplish an action of some kind (e.g. transmission of a packet, establishing a connection, etc), but every individual attempt is subject to failure because of noise or transient network conditions; therefore, the agent must repeat making attempts until the final success (or, possibly, until giving up). Examples are ubiquitous at all networking layers, including but not limited to:

- collision resolution in shared channels, e.g. IEEE 802.3 [1],[2],[3], where a frame is retransmitted after collision;
- reliable end-to-end transport protocols, such as TCP [4], which use packet retransmission to overcome losses at the network layer;
- connection establishment in ATM [5], which employs *crankback* to find a route with the necessary resources;
- application-specific retrials, e.g. in SMTP [6] and HTTP [7].

Furthermore, examples of retrials after failures extend well beyond the realm of networking, and include submission of computation jobs to time-shared computers, phonecall-operated service centers, and many more.

This research was supported by the Israeli Ministry of Science.

To recognize a failure from success, the agent must receive an *acknowledgment* from the network. The acknowledgment can be either negative (NACK) or positive (ACK), where a constructive reply from the remote party also constitutes an implicit positive acknowledgment. At times, particularly if the cause of failure is easily detectable (e.g. a full queue in a server, or collision in a shared channel), a NACK can be generated quickly. However, such cases are exceptional, and it is more typical for the delay of NACKs to be significantly higher than that of ACKs. This asymmetry between ACKs and NACKs is especially characteristic of large-scale networks and is a consequence of their structure, particularly the high variability of delay. In fact, for this reason, often explicit NACKs are not implemented at all. A natural consequence of this asymmetry is that the decision between success and failure is based on a *timeout*, defined as the duration of time after which the action is attempted again.

Selection of a proper timeout involves a tradeoff between waiting too long, on the one hand, and loading the network needlessly by waiting too little, on the other. The common approach, employed in the vast majority of cases where timeouts are used, is to set the timeout to a large value, such that, unless the action fails, it is acknowledged within the timeout duration with a high probability (usually very close to 1). This approach is conservative; it puts an absolute priority to the "common good" (the desire not to load the network) at the expense of the "private good" (the agent's desire for fast completion), and leads to overly long timeouts, sometimes even up to a point that requires human intervention. For example, the TCP algorithm [4],[8] makes a strong emphasis on avoiding so-called 'spurious' retransmissions.

While the conservative approach works reasonably well when the (*a priori*) failure probability is very low, it is far from optimal in many cases. To see the inefficiency, consider a lightly-loaded network consisting of high-capacity but noisy links, with high propagation delays. It may then be reasonable to make a repeated attempt well before the typical delay of an acknowledgment, since the network can accommodate the extra attempt easily. Taking this to the extreme, it may even be worthwhile to make several attempts in parallel at the outset, before starting to wait for any acknowledgment at all. This reduces the expected time until success, not only because every individual attempt may fail by itself, but also because one of the subsequent attempts may succeed earlier than the original (e.g. by taking a different route).

Motivated by the above observations, this paper embarks on a generic study of timeout strategies, looking to describe ways to compute optimal ones and investigate their general properties. By "generic" we mean that no specific problem or context is assumed for the analysis; both the probability of failure and the distribution of delay from the attempt to its acknowledgment (in case of success) can be arbitrary. The only important assumption, made mainly for mathematical tractability, is that the delay and failure probability are independent among attempts. While this assumption limits the applicability of the analysis, it is justified in many practical situations — in particular, when repeated attempts try a different or modified action every time (rather than repeating the same action), in terms of, e.g., route to the destination, a lower amount of requested resources, etc. Possible examples include, but are not limited to, the following:

- a client retrieving information from one of a large pool of mirror servers (which, if properly deployed, should represent equal and independent alternatives);
- fault-tolerant connections, which preallocate parallel (independent) paths to provide fast restoration;
- connection establishment, which attempts different routes to the destination (crankback), different amounts of requested resources, or different priorities or service classes.

The assumption of independence between attempts is less justified when the retrial attempts simply repeat the same action, in the hope that the failure was caused by some transient condition which is no longer there (e.g. packet retransmission in TCP over the same path), since in those cases the failure probability of an attempt typically depends on the time elapsed from a previous (failed) one. For such contexts, the paper's analysis can only be applied in special cases, e.g., for TCP, if the failures are mostly caused by "white noise" or if the underlying routing layer employs path splitting and thus sends packets over different paths.

In order to be able to compare between timeout strategies and compute optimal ones, the above-mentioned tradeoff between time and load must be quantified. Therefore, we model the tradeoff by associating a 'cost' per unit of time till success and a 'cost' per attempt, and defining the optimal timeout strategy as one that a selfish agent would follow to minimize its expected cost. These costs can have various practical interpretations, and it should not be taken literally that the agent is indeed 'charged' for them, or that it is necessarily selfish. For example, the time cost may be associated with the disutility incurred by the application due to increased delay. Similarly, a 'social' agent that refrains from retransmitting to avoid loading the network for others behaves as if it had a high (quickly increasing) attempt cost; in particular, if the computation of timeouts takes place in equipment provided by the network operator and cannot be controlled by the end user, the cost per attempt can simply be the operator's assessment of the user's "contribution" to the network load. Nonetheless, the literal interpretation of the costs also has merit; it suggests using *pricing* as a management tool, through which the network can affect the timeout strategies of its users, and, to some degree, control their flows. In this sense, our work fits into the broad framework of flow and congestion control by pricing [9],[10].

After defining the model, we derive the formulae for the optimal timeout values and study some of their properties. In particular, we identify the conditions for making parallel attempts from the outset to be worthwhile. In addition, we demonstrate a striking property of *positive feedback*: a higher network load can, in some cases, lead to a shorter optimal timeout. This motivates us to study the interaction resulting when many users selfishly apply the optimal timeout strategy, using a noncooperative game model [11], and show that it suffers from an inherent instability problem. The practical implications of these results are discussed.

The rest of this paper is structured as follows. Our model and assumptions are described in section II. The considerations in selecting the first timeout are described in section III. Section IV discusses the dependency on a deterministic propagation delay. Section V establishes the framework for computing the subsequent timeouts and proves several general properties of the asymptotic dependence of the timeout strategy when the price per retry tends to zero. Section VI discusses the game resulting from the interaction of multiple selfish users sharing a network, and, in particular, demonstrates an inherent instability problem in this game. Finally, section VII summarizes the results and their practical implications, and outlines related problems to be studied in future research. To maintain presentation continuity and clarity, some of the technical calculations and proofs are omitted from the main text and gathered in the Appendix.

II. Model Definitions and Assumptions

For the sake of concreteness, we present the analysis in terms of the optimal timeout strategy for reliable access of parallel servers in a lossy network. Nevertheless, as explained in the Introduction, our study applies to a variety of problems in different contexts involving retrials and timeouts. Therefore, the following terms should be understood generically.

Consider a client with a request for some kind of service that can be fulfilled by any of a large number of mirror servers. The network connecting the client and the servers is unreliable, and can lose either the service request or reply packets; in addition, a service request can be dropped by a server, e.g. due to congestion. The clients's policy will be to transmit the request and wait for an acknowledgment[1]; if a certain time (a *timeout*) elapses without the request having been acknowledged, it will be retransmitted, and so on until the acknowledgment finally arrives. Formally, we define a *strategy* of the client as an infinite vector of non-negative timeouts, which the client waits between retransmissions as long as the reply does not arrive. The timeout may change for every subsequent transmission attempt, and can be a function of the number of transmission attempts so far and of the time differences between them.

The client's goal is to find the optimal timeout strategy, i.e. one that brings the total expected cost to a minimum. We assume that the client's cost is composed of a price of a per unit of time and b per request; to put it differently, if it takes a total of T time units and P request transmissions to get a reply, then the total cost is $J(T, P) = a \cdot T + b \cdot P$. As explained in the Introduction, these prices quantify the tradeoff between waiting too long and loading the network too much; properly selected, they can embed the delay or congestion cost (disutil-

[1] We use the terms *reply* and *acknowledgment* interchangeably.

ity) of the client's application or the network itself, and need not necessarily be taken literally. We note that other forms of cost can also be considered, as long as the cost function $J(T,P)$ increases in both its arguments. For example, an application with a rigid delay bound requirement can be modeled by a step-like cost function, with a sharp increase for T that is at or around the value of the bound. However, such cases are specific and, in our view, deserve to be studied separately (using, perhaps, the same methodology as in this paper). In contrast, the linear form of cost function, i.e. with constant time and retransmission prices, is rather generic and occurs naturally and frequently. For instance, the charge that networks levy for usage time is typically linear, at any rate for time scales of a typical application lifetime (e.g. 'ether time' in cellular networks). Likewise, the cost per retransmission, whether interpreted literally (e.g. in terms of energy for a cellular device) or as the user's projection on the network congestion, can also be fairly considered constant in typical cases (at least for 'small' clients which do not have the ability to congest the network completely on their own). The linear cost structure was advocated, albeit for a different purpose, in [12], from where we borrow the notation of a for cost per time and b for cost per unit of transmitted information.

The optimal strategy depends, of course, on the client's estimation of the loss probability and delay distribution of its requests in the network and in the servers, and on the interactions it may have with other network traffic, including its own requests transmitted earlier. We make the following assumptions in our analysis.

- Independence: Each retransmitted request follows a new, independent path in the network to an independent server. Either of the requests can be acknowledged (only one reply is required), and the acknowledgments may arrive to the client in any order, regardless of the order of request transmissions.[2]
- Homogeneity of loss: The probability of a request to be lost (either by the network or the server) is L, and is the same for all servers. If a request is lost, its acknowledgment never arrives.
- Homogeneity of delay: If a request is not lost, then the *a priori* probability density function of the delay between its transmission and its acknowledgment is given by $d(t)$, and is the same for all servers. In particular, it is not affected by the rate that previous requests were transmitted recently.[3]

To illustrate the timeout tradeoff within this framework, consider first what happens if $b = 0$, i.e. the client is not charged anything for retransmissions but only for the total connection time. Then, it is easy to see that the best scheme is simply to retransmit the request over and over indefinitely. Since there is no charge for a retransmission, there is no incentive for the client to wait at all: by retransmitting as early as possible, it increases the chance of getting an early reply (whether previous requests are lost or not) and therefore reduces the expected waiting time. On the other hand, for a fixed b, as $a \to 0$, the client is willing to wait indefinitely for an acknowledgment (assuming that the delay probability density does not decay actually to zero for any finite delay). The client pays nothing for the extra wait, but increases the chance that the missing acknowledgment will arrive after all and that the request will therefore not have to be retransmitted. Having considered the extreme cases, the subsequent analysis assumes finite and nonzero values for a and b.

Before continuing, we note one case, potentially of practical importance, in which the optimal timeout strategy is easily found. Suppose that the client is unable to hold more than one connection attempt at a time, either due to network limitations or lack of local resources. We shall call such a client *simple*. Then, once the client decides to retransmit, it needs to abandon the previous attempt and will not live to see a reply even if the original request was not lost. In this case, it is obvious that the optimal strategy consists of identical timeouts, since there is no difference between the computation of the first timeout and that of any subsequent one. To compute the optimal timeout value, we note that with a timeout of τ, the expected cost is $(a\tau + b)$ times the number of unsuccessful requests, plus a times the expected reply time of the last (successful) request. Formally,

$$E(\tau) \triangleq \left[\frac{1}{(1-L)\int_0^\tau d(t)\,dt} - 1\right](a\tau + b) + a\frac{\int_0^\tau t \cdot d(t)\,dt}{\int_0^\tau d(t)\,dt}, \tag{1}$$

and the optimal timeout, i.e. the one that brings (1) to a minimum, can be found either by numerical minimization or, in some cases, analytically by equating the derivative to zero:

$$\begin{aligned}\frac{\partial E(\tau)}{\partial \tau} &= \left[\frac{1}{(1-L)\int_0^\tau d(t)\,dt} - 1\right] \cdot a - \\ &- \left[\frac{d(\tau)}{(1-L)\left[\int_0^\tau d(t)\,dt\right]^2}\right] \cdot (a\tau + b) + \\ &+ a \cdot \frac{d(\tau)\left[\int_0^\tau d(t)\,dt \cdot \tau - \int_0^\tau t \cdot d(t)\,dt\right]}{\left[\int_0^\tau d(t)\,dt\right]^2} = 0. \end{aligned} \tag{2}$$

Having considered simple clients, we now focus on the more interesting case where the client can hold an unlimited number of connections at a time; hence, there is no reason for all timeouts to be identical. Moreover, except for special cases of extremely simple delay distributions (such as deterministic), it is impossible — neither for us here nor for any client implementation in practice — to find the optimal strategy directly, i.e. by calculating the expected cost of every strategy and solving the resulting minimization problem in an infinite number of variables. However, even if it were feasible, computing the entire timeout strategy in advance would probably be overkill in most cases: in any system with 'reasonable' delays, losses, and costs, it is most unlikely that the exact value of, say, the 100-th timeout is going to be needed, so it is futile to spend the required computational effort to find it in advance. Rather, it is logical for the client to employ an online approach of computing each timeout as it is needed, taking into account the times of the previous transmissions, but ignoring any future retransmissions that may be required. To put it differently, as long as no request is acknowledged, we suggest that the client continuously decides at every point in time whether it is more profitable to make another transmission now or to wait some more, assuming that this

[2] The implications of the independence assumption, and the contexts it can be applied in, have been considered in the Introduction.

[3] We note the homogeneity assumptions are introduced mainly for simplicity of presentation, and point out that to alleviate them (i.e. extend the model to let the loss probability and delay distribution vary among the servers) is rather straightforward.

transmission will be the last one and will not be lost; however, having made it, the client immediately gets back to deciding continuously when to retransmit next. In the rest of this paper, we elaborate and study some general properties of this online (in a sense, 'greedy') approach. Incidentally, it should be intuitively obvious that acting as described above will approximate the 'truly' optimal strategy well if the loss probability is small.

III. The First Timeout

The client obviously starts by transmitting the request once. Suppose that, at a certain time later, the request is still unacknowledged; then, it is either lost or simply suffers a long delay. Denote by $p(t)$ the *a posteriori* probability density of the additional delay (from this moment) to be t; thus, if the time elapsed so far since the request transmission is τ,[4]

$$p(t)_{|\tau} = \frac{(1-L)\cdot d(t+\tau)}{(1-L)\int_\tau^\infty d(x)\,dx + L}. \tag{3}$$

Note that $p(t)$ is not a legitimate probability density in the usual sense, since its integral is less than 1; the missing probability mass, i.e.,

$$1 - \int_0^\infty p(t)\,dt = \frac{L}{(1-L)\int_\tau^\infty d(x)dx + L}, \tag{4}$$

is the probability that the additional delay is not any finite number, corresponding to the *a posteriori* probability that the request was lost. This probability is increasing in τ, except during periods of time for which the *a priori* delay probability density is zero.

Having computed $p(t)$, should the client retransmit the request immediately, or should it wait some more before deciding to retransmit? An immediate retransmission is worthwhile if it gains at least as much as it loses compared with any $\Delta t > 0$ of additional waiting, and the optimal timeout is the smallest value of τ after which this condition is satisfied.

Suppose that the original request is not lost. Retransmitting immediately can still be profitable if the retransmitted request's acknowledgment arrives before the original one. Assume that τ time units have elapsed, the original request is about to arrive within t more time units, and the retransmitted request arrives within θ time units. Then, if $\theta < t$, a time charge of $a(t-\theta)$ is saved. Retransmitting immediately will, therefore, save on the average

$$a\int_0^\infty \left[\int_0^t (t-\theta)\,d(\theta)\,d\theta\right] p(t)\,dt,$$

so the difference in favor of immediate retransmission due to this factor only is

$$a\int_0^\infty \left[\int_0^t (t-\theta)\,d(\theta)\,d\theta\right] [p(t) - p(t+\Delta t)]\,dt.$$

[4]The a posteriori probability density will be denoted by $p(t)_{|\tau}$, or, if there is no ambiguity as to the elapsed time since the transmission, simply by $p(t)$.

In addition, if the original request is lost, retransmitting immediately gains $a\Delta t$ relative to waiting an extra Δt. The downside of retransmitting immediately is that it costs an extra b if the original request is not lost and the reply is about to arrive within the next Δt. Adding up all the terms, we get the condition to retransmit immediately, namely, that the expected saving in time charge is at least as high as the expected extra retransmission cost for any $\Delta t > 0$:

$$\boxed{\begin{aligned} &a\int_0^\infty \left[\int_0^t (t-\theta)\,d(\theta)\,d\theta\right] [p(t) - p(t+\Delta t)]\,dt + \\ &\quad + a\Delta t\left(1 - \int_0^\infty p(t)\,dt\right) \ge b\int_0^{\Delta t} p(t)\,dt. \end{aligned}} \tag{5}$$

A. Example: M/M/1-induced Delay and Loss

We now wish to demonstrate the results that an application of formula (5) leads to with a specific delay distrubtion and loss probability. Unfortunately, it is a well-established truism that an accurate description of delay and loss in real networks is very complex if at all possible; therefore, for practical reasons, clients must settle for reasonably simple approximations. One commonly used model of network delay, and one that we shall use (solely for the purpose of this example), is that of the M/M/1 queue [13], in which the customers' arrival is Poisson with rate λ, and the service time is exponentially distributed with rate μ. To capture the dependence of loss probability on congestion, suppose that the request is lost if it enters the queue at a time that it has more than a certain threshold m of customers. The *a priori* request loss probability is, therefore, the probability of the queue to be with m customers or more,

$$L = \rho^m, \tag{6}$$

where $\rho \triangleq \frac{\lambda}{\mu}$. The probability density of a request delay is

$$d(t) = (\mu - \lambda)\,e^{-(\mu-\lambda)t}. \tag{7}$$

Although formulae (6)–(7) are derived from the simplest of models, they do capture the essentials of the exponential dependence of delay and power-type dependence of loss on the network load, which can be considered typical in many cases; furthermore, even when a more detailed model can describe the exact delay and loss expressions governing a specific instance, clients are still likely to base their timeout calculations on a simplified model such as the above, in order to maintain computational simplicity (recall that the timeout strategy is computed online). Therefore, we reckon the discussion of the properties of the resulting timeout strategy in this and the following sections is insightful in general. We also point out, without elaborating on details, that minor modifications to the model, such as taking the request delay to be just the queue waiting time, excluding the service time (i.e. $d(t) = (1-\rho)\left[\lambda e^{-(\mu-\lambda)t} + \delta(t)\right]$), or considering an M/M/1/m queue limited to a total of m customers altogether (which leads to $L = \frac{(1-\rho)\rho^m}{1-\rho^{m+1}}$), result in timeout expressions with essentially similar properties.

Computing the *a posteriori* delay density function after τ time units, according to formula (3),

$$p(t)_{|\tau} = \frac{(1-\rho^m)(\mu-\lambda)\,e^{-(\mu-\lambda)(t+\tau)}}{(1-\rho^m)\,e^{-(\mu-\lambda)\tau}+\rho^m}, \qquad (8)$$

and substituting into the timeout condition (5), we get (skipping the details of the integral calculations, which can be found in the Appendix as identities (27)–(30)) that the condition to retransmit is

$$(1-\rho^m)\,e^{-(\mu-\lambda)\tau}\frac{1-e^{-(\mu-\lambda)\Delta t}}{\Delta t}\left[\frac{a}{2(\mu-\lambda)}-b\right]+ + a\rho^m \geq 0. \qquad (9)$$

Obviously, this condition is satisfied for any τ and Δt if the expression in brackets is nonnegative. Even if it is negative, since $e^{-(\mu-\lambda)\tau}\cdot\frac{1-e^{-(\mu-\lambda)\Delta t}}{\Delta t}$ is decreasing both in τ and in Δt and is bounded by $\mu-\lambda$, the condition is still fulfilled for any τ and Δt if

$$\left[\frac{a}{2}-b(\mu-\lambda)\right](1-\rho^m)+a\rho^m \geq 0,$$

or, by a different grouping of the terms, the timeout is zero if

$$\boxed{\frac{a}{b} \geq \mu\cdot\frac{2(1-\rho)(1-\rho^m)}{1+\rho^m}.} \qquad (10)$$

For any $\frac{a}{b}$ below this threshold, the optimal timeout (that attains equality in (9)) is

$$\boxed{\tau = -\frac{1}{\mu-\lambda}\ln\left[\frac{\frac{a}{b}}{\mu-\lambda-\frac{a}{2b}}\cdot\frac{\rho^m}{1-\rho^m}\right].} \qquad (11)$$

We can make the following remarks on the obtained expressions:

- For a given price per unit time, there exists a threshold below which the price per request does not give the client enough incentive to wait before transmitting again.
- For a large m, and if ρ is not too close to 1, the threshold (in terms of $\frac{a}{b}$) is virtually equal to $2\mu(1-\rho)$. In particular, setting $\frac{a}{b}\geq 2\mu$ causes the client's first timeout to be 0 for any load.
- For a large m and a given $\frac{a}{b}$, the timeout τ grows approximately linearly with m.
- As $\frac{a}{b}$ tends to zero, τ grows without bound, as explained in the discussion of the extreme cases above (section II).
- For $\frac{a}{b}$ close to zero, τ behaves approximately as $\ln\frac{b}{a}$. Thus, to increase the retransmission timeout additively, the retransmission price must grow multiplicatively relative to the time price.

To gain some further insight on the behavior of the optimal timeout, we plot the expression (11) as a function of the load λ, keeping the price ratio $\frac{a}{b}$ constant. Figure 1 shows the plots for a few selected values of $\frac{a}{b}$; the buffer size was set to $m=30$ throughout. The plots are normalized for $\mu=1$; thus, the horizontal axis shows $\rho=\lambda/\mu$, the vertical axis shows τ in units of $1/\mu$, and $\frac{a}{b}$ is given in units of μ.

From these plots, one can deduce the following conclusions:

- For a given m and $\frac{a}{b}$, the retransmission timeout is mostly a decreasing function of the load λ, especially when the network is lightly loaded (λ is close to 0). This introduces an effect of *positive feedback* — as λ (the arrival rate of other requests) grows, the client retransmits more frequently.

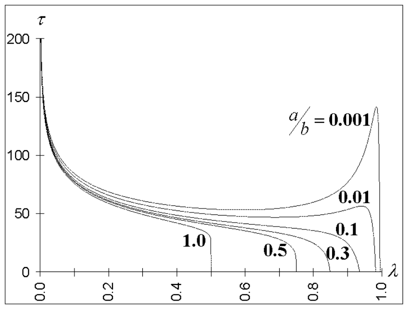

Fig. 1. The timeout vs λ

- For sufficiently low values of $\frac{a}{b}$, there is a range of λ for which the timeout is an increasing function of λ; for higher values of $\frac{a}{b}$, it becomes a monotonously decreasing function.
- For a given m and $\frac{a}{b}$, the timeout is a continuous function of λ; as soon as λ passes the value that causes $\frac{a}{b}$ to be equal to the threshold, the retransmission timeout drops to 0.

IV. Adding a Propagation Delay

Suppose that the connection delay is composed of a deterministic part (e.g. propagation delay) and a random part (e.g. queuing delay). We keep the stochastic part notations as before ($d(t)$ and $p(t)_{|\tau}$), and denote the deterministic delay by T.

This extension can be incorporated into our model in a straightforward way. We compute the "overall" *a priori* delay distribution function

$$\boxed{d'(t) = \begin{cases} 0 & t<T, \\ d(t-T) & t\geq T, \end{cases}} \qquad (12)$$

and the "overall" *a posteriori* probability distribution, after τ time units, for the request to arrive within t more time units

$$\boxed{p'(t)_{|\tau} = \begin{cases} 0 & \tau+t<T, \\ p(t-(T-\tau))_{|0} & \tau<T,\ \tau+t\geq T, \\ p(t)_{|\tau-T} & \tau\geq T. \end{cases}} \qquad (13)$$

Once $d'(t)$ and $p'(t)$ are computed, the optimal timeout can be found from the condition (5) with $d'(t)$ and $p'(t)$ substituted for $d(t)$ and $p(t)$, respectively.

The following series of propositions and theorems present a few general properties of the optimal timeout, that hold regardless of the specific form of the delay distribution.

Proposition 1. *The optimal timeout is either $\tau=0$ or $\tau\geq T$.*

Proposition 2. *If the optimal timeout is not zero, then it is not less than τ^*+T, where τ^* is the optimal timeout for the same delay distribution but without a propagation delay.*

Theorem 1. *The threshold value of $\frac{a}{b}$, beyond which the optimal timeout is zero, behaves asymptotically as $\Theta\left(\frac{1}{T}\right)$ for $T\to\infty$.*

The above propositions and theorem are proved in the Appendix.

A. Example: M/M/1-induced Delay and Loss

We now apply the above ideas to the M/M/1-induced example, described by (6)–(8).

Substituting $p'(t)_{|0}$ into (5), we obtain the condition for $\tau = 0$ and $\Delta t \geq T$ (the condition is satisfied trivially for $0 < \Delta t < T$, since the right-hand side of (5) is then 0):

$$(1-\rho^m) \cdot \frac{1-e^{-(\mu-\lambda)(\Delta t - T)}}{\Delta t} \left[\frac{a}{2(\mu-\lambda)} - b\right] + a\rho^m \geq 0. \tag{14}$$

Now, it can be easily verified (e.g. by taking the derivative with respect to Δt and equating it to zero) that

$$\sup_{T \leq \Delta t < \infty} \frac{1-e^{-(\mu-\lambda)(\Delta t-T)}}{\Delta t} = \frac{\mu-\lambda}{-\operatorname{plog}\left(-e^{-[(\mu-\lambda)T+1]}\right)}, \tag{15}$$

where $\operatorname{plog}(\cdot)$ (the product-log function) denotes the inverse function of $f(x) = xe^x$, such that $x = -\operatorname{plog}(-y)$ (for $0 < y \leq \frac{1}{e}$) is the largest positive solution to the equation $y = xe^{-x}$. Condition (14) is therefore satisfied for every Δt (and, consequently, the optimal timeout is zero) if

$$\frac{a}{b} \geq \mu \frac{2(1-\rho)(1-\rho^m)}{1+\rho^m\left[-2\operatorname{plog}\left(-e^{-[(\mu-\lambda)T+1]}\right)-1\right]}. \tag{16}$$

Note, incidentally, that the threshold is indeed $\Theta\left(\frac{1}{T}\right)$, because $-\operatorname{plog}(-e^{-x}) \approx x + \ln x$ for large values of x ($x \gg 1$).

For $\frac{a}{b}$ below this threshold, the timeout can be computed by substituting $p'(t)_{|\tau}$ for $\tau \geq T$ in (5):

$$(1-\rho^m)\, e^{-(\mu-\lambda)(\tau-T)} \frac{1-e^{-(\mu-\lambda)\Delta t}}{\Delta t} \cdot \cdot \left[\frac{ae^{-(\mu-\lambda)T}}{2(\mu-\lambda)} - b\right] + a\rho^m \geq 0, \tag{17}$$

therefore (using the same technique as in section III-A),

$$\tau = -\frac{1}{\mu-\lambda} \ln\left[\frac{\frac{a}{b}}{\mu-\lambda-\frac{a}{2b}e^{-(\mu-\lambda)T}} \cdot \frac{\rho^m}{1-\rho^m}\right] + T, \tag{18}$$

and, indeed, it is seen to be shifted by more than T relative to the optimal timeout without a propagation delay, given by (11).

V. Subsequent Timeouts

The above analysis led to some results concerning the timeout for the first retransmission. The results for subsequent retransmissions need to be extended, in a rather straightforward way.

First, we define some relevant notation. Suppose that the (still unacknowledged) request has been transmitted k times so far, and denote the vector of elapsed times since the transmissions by $\vec{\tau} = \langle \tau_1, \tau_2, \ldots, \tau_k \rangle$. As before, denote by $p(t)_{|\tau_i}$ the *a posteriori* probability density of the i-th transmission to arrive after an additional delay of t (from this moment), i.e.

$$p(t)_{|\tau_i} = \frac{(1-L)\cdot d(t+\tau_i)}{(1-L)\int_{\tau_i}^{\infty} d(x)\,dx + L}, \tag{19}$$

and denote by $p(t)_{|\vec{\tau}}$ the probability density of the random variable that is the minimum (in a stochastic sense) of the random variables defined by $p(t)_{|\tau_i}$, $1 \leq i \leq k$; hence, $p(t)_{|\vec{\tau}}$ is the probability density of the earliest arrival from this moment to be after an additional delay of t.

The expression for $p(t)_{|\vec{\tau}}$ is

$$\boxed{p(t)_{|\vec{\tau}} = \sum_{i=1}^{k}\left\{p(t)_{|\tau_i} \prod_{j\neq i}\left[1-\int_0^t p(\theta)_{|\tau_j}\,d\theta\right]\right\}.} \tag{20}$$

This expression sums the probability densities that one of the transmissions will be acknowledged at time t from now, multiplied by the probabilities of all the other transmissions not to be acknowledged until then. Note that the probability of transmission j not to be acknowledged by time t is taken as $1-\int_0^t p(\theta)_{|\tau_j}\,d\theta$ and not as $\int_t^{\infty} p(\theta)_{|\tau_j}\,d\theta$, because of the missing probability mass at infinity corresponding to the probability of the request to be lost altogether. Once $p(t)_{|\vec{\tau}}$ is computed, the condition under which it is desirable to make another retransmission immediately is given by (5), with $p(t)_{|\vec{\tau}}$ substituted for $p(t)$.

An alternative approach to handle the subsequent timeouts, which can be more convenient in some cases, is via the *a posteriori* cumulative delay distribution function $P(t)_{|\vec{\tau}} = \int_0^t p(\theta)_{|\vec{\tau}}\,d\theta$. (Again, we note that this is not a cumulative probability distribution in the traditional sense, because it does not tend to 1 as $t \to \infty$.) It is given by the expression

$$P(t)_{|\vec{\tau}} = 1 - \prod_{i=1}^{k}\left[1-P(t)_{|\tau_i}\right], \tag{20'}$$

which reflects the fact that the probability of the earliest request *not* to arrive by time t is simply the probability of *no* request to arrive during that time. In addition, it is easily seen that $p(t)_{|\vec{\tau}}$, given by formula (20), is the same as $\frac{dP(t)_{|\vec{\tau}}}{dt}$, using the rule of the derivative of a product and the chain rule.

We note a general property of expression (20).

Theorem 2. *Consider two vectors $\vec{\tau}_1$, $\vec{\tau}_2$ such that $\vec{\tau}_1$ is a subsequence of $\vec{\tau}_2$, i.e. $\vec{\tau}_2$ represents a situation which includes all the transmissions described by $\vec{\tau}_1$ plus some others. Then, if condition* (5) *is satisfied for any Δt with $p(t)_{|\vec{\tau}_2}$, then it is also satisfied for any Δt with $p(t)_{|\vec{\tau}_1}$.*

Proof. See Appendix. □

Corollary. *In the resulting timeout strategy, all the timeouts are at least as long as the first one.*

Proof. Denote the first timeout by τ_1. Then condition (5) is not satisfied (for every Δt) for any $\tau < \tau_1$. Consequently, in light of the theorem, it will not be satisfied (for every Δt) after a vector $\vec{\tau}_2$ that contains τ as one of its components. Therefore, τ cannot be the optimal timeout after any subsequent retransmission. □

We now focus our analysis on a specific yet interesting issue of immediate retransmissions at the outset (i.e. making parallel attempts initially, before starting to wait for any acknowledgment). It was already noted that, when $b = 0$, the client retransmits the request immediately again and again until receiving the acknowledgment. Therefore, we expect that as $\frac{a}{b}$ increases, the number of immediate retransmissions (alternatively, the number of zero timeouts in the beginning of the optimal strategy) increases as well, and we seek to describe the asymptotic dependence of this number on the various parameters. For this purpose, we introduce a couple of notations:

$D(t)$ - the cumulative delay distribution of a single request, $D(t) = \int_0^t d(t)\,dt$. In case that $D(t)$ has discontinuities, we assume that it is upper-semicontinuous.

$t_0 = \inf\{t|D(t) > 0\}$ - the smallest possible delay that can be attained by a single request.

Theorem 3. *As $\frac{a}{b} \to \infty$,*

a) If $D(t_0) > 0$, then the number of immediate retransmissions is $O\left(\log \frac{a}{b}\right)$.

b) If $D(t_0) = 0$ and can be expanded around t_0 as $D(t) = \sum_i C_i \cdot (t - t_0)^{\alpha_i}$, for certain constants C_i such that $C_1 > 0$ and $0 < \alpha_1 < \alpha_2 < \ldots$, then the number of immediate retransmissions is $O\left(\left(\frac{a}{b}\right)^{\frac{\alpha_1}{\alpha_1+1}}\right)$.

Theorem 4. *For any fixed $\frac{a}{b}$, if a propagation delay of T is added to the delay distribution (as in section IV), then the number of immediate retransmissions is $O(\log T)$ for $T \to \infty$.*

The proofs of these theorems pass through a rather long series of intermediate lemmas. The details can be found in the Appendix.

As a common special case, if the delay distribution is continuous and its density at t_0 is finite and strictly positive, then $\alpha_1 = 1$, and, consequently, the number of immediate retransmissions is $O\left(\sqrt{\frac{a}{b}}\right)$. Such are, for example, the exponential distribution, the uniform distribution, and many more.

A. Example: Deterministic delay

One case in which the optimal timeout strategy can be found directly is that of a deterministic link delay T. By intuitive reasoning, it is not hard to see that the optimal timeout strategy consists of timeouts of either 0 or T; furthermore, the number of adjacent zero timeouts (i.e. immediate retransmissions) before every wait of T is constant.

If the loss probability of an individual request is L, then the optimal number of times to transmit it immediately is the number k^* that minimizes the expression $b \cdot k + a \cdot T \cdot L^k$. By taking the derivative with respect to k (ignoring the need of k to be a whole number) and equating it to zero, one easily obtains the solution,

$$k^* = \frac{\log\left(\frac{a}{b} T \log \frac{1}{L}\right)}{\log \frac{1}{L}}, \tag{21}$$

which clearly exhibits the logarithmic dependence on $\frac{a}{b}$ and on T predicted by the first case of theorem 3 and theorem 4 (for $t_0 = T$ and $D(t_0) = 1$).

B. Example: Exponentially-distributed delay without loss

Suppose that the delay distribution is exponential, $d(t) = \mu e^{-\mu t}$, and that requests are not lost, $L = 0$. This is another case in which the optimal timeout strategy can be found directly, by intuitive reasoning. Because of the memoryless property of the exponential distribution, after any elapsed time τ, the *a posteriori* distribution of the remaining time is the same as the *a priori* delay distribution. Consequently, there cannot be any positive timeouts in the optimal strategy: if it is profitable to resend a request after a time $\tau > 0$, then it must be at least as profitable to resend it immediately. Therefore, the optimal strategy must have the form of sending k requests immediately, followed by unlimited waiting until one of them is acknowledged. Since the average delay of the earliest among k independent exponentially-distributed variables is $\frac{1}{k\mu}$, the optimal number k^* is the one that minimizes $a \cdot \frac{1}{k\mu} + b \cdot k$. By taking the derivative and equating to zero, it is easily found to be

$$k^* = \sqrt{\frac{a}{\mu b}}. \tag{22}$$

This is consistent with the second case of theorem 3, as $D(t) = 1 - e^{-\mu t} = \sum_{i=1}^{\infty} \frac{(-1)^{i-1}(\mu t)^i}{i!}$, therefore $\alpha_1 = 1$.

VI. Stability Considerations

In this section, our aim is to consider the interaction among multiple clients that selfishly compete over the same servers, and estimate the effect that frequent retransmissions, caused by small values of the timeout, can have on the link load and, therefore, on the behavior of other clients. To begin, we again turn to the example of the M/M/1 queue, analyzed in some detail in section III-A. Recall that, in that example, an effect of positive feedback has been noted, namely, that as the queue load gets higher, the timeout (generally, except in some regions) gets smaller. As is generally the case with positive feeback, one can therefore expect that interaction among multiple clients selfishly acting according to each own's optimal strategy may lead to some kind of instability, and we shall demonstrate that (unfortunately) this is indeed the case.

Assume that the total load arriving to the network is a memoryless process with an average rate γ, and that it is composed of an infinite number of clients, each with an infinitesimal impact. Each client stays in the system and retransmits its request, according to its optimal policy, until it is acknowledged. The number of times that a client transmits its request is, obviously, a random value, depending on the number of times that a timeout elapses before the request is eventually acknowledged. We denote this random variable by N, and its mean by $\overline{N}$. Therefore, we have the following relation between the arrival rate of new clients and the rate of their requests:

$$\lambda = \gamma \cdot \overline{N}. \tag{23}$$

Since, unfortunately, attempting to compute $\overline{N}$ for the general model requires computing the optimal strategy and quickly leads to unmanageable expressions for subsequent timeouts, we shall limit ourselves to the model of simple clients for the rest of this section. With simple clients, the optimal timeouts are equal to each other and given by the solution to equation (2), which, upon substituting the M/M/1-induced loss and delay distribution ((6)–(7)) and some straightforward simplification, becomes

$$\{a\rho^m [(\mu - \lambda)\tau + 1] + (\mu - \lambda) b\} e^{-(\mu-\lambda)\tau} - a\rho^m = 0. \tag{24}$$

The solution to this equation is

$$\tau^* = \frac{1}{\mu - \lambda}\left[-\operatorname{plog}\left(-e^{-K}\right) - K\right], \tag{25}$$

where $K \triangleq 1+\frac{\mu b}{a}\cdot\frac{1-\rho}{\rho^m}$, and $\operatorname{plog}(\cdot)$ is the product-log function (recall equation (15)).

Assuming that all clients employ the optimal timeout policy, the probability of a request not to be acknowledged within the timeout period, and hence to be retransmitted, is

$$P_r \triangleq 1-(1-L)\int_0^{\tau^*} d(t)\,dt = \rho^m + (1-\rho^m)\,e^{-(\mu-\lambda)\tau^*}.$$

The average number of times for a request to be transmitted is

$$\overline{N} = \frac{1}{1-P_r} = \frac{1}{(1-\rho^m)\left(1-e^{-(\mu-\lambda)\tau^*}\right)},$$

and, therefore,[5]

$$\gamma = \frac{\lambda}{\overline{N}} = \mu\rho\,(1-\rho^m)\left(1-e^{-(\mu-\lambda)\tau^*}\right). \qquad (26)$$

Equation (26) describes the stable behavior of the client game for given prices and arrival rate. In other words, for a given arrival rate γ, the steady-state load $\lambda = \mu\rho$ will be one that solves (26). To gain some more insight, figure 2 provides a typical plot of the relation $\gamma(\lambda)$; this plot is obtained with $\frac{a}{b} = 0.2$, $m = 30$, and, as before, normalized for $\mu = 1$ (i.e. both γ and λ are given in units of μ).

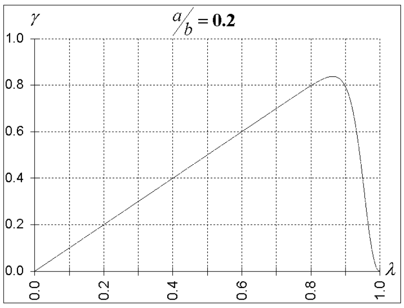

Fig. 2. Plot of a typical $\gamma(\lambda)$

The plot in figure 2 clearly shows that, for small values of γ, there are two solutions to equation (26). One is $\lambda \approx \gamma$, corresponding to 'reasonable' values of the timeout and a very small probability of retransmission. This solution can be seen to be stable, in the sense that a small perturbation in the arrival rate will cause the clients to converge to a nearby point. There is, however, another solution, with λ much closer to 1, which also describes an equilibrium and corresponds to a very small value of the timeout coupled with a relatively large loss probability. This is an unstable equilibrium; a small decrease in the load perceived by the clients will increase the timeouts and, eventually, bring the system back to the first equilibrium. Unfortunately, a small *increase* in the perceived load can have a devastating effect: it decreases the timeouts even further, and eventually brings about an equilibrium not described by (26), in which all the clients retransmit their requests with no timeout at all, the servers are overloaded, and requests are lost with probability 1. We emphasize that the qualitative relation between the client arrival rate and the total load and, in particular, the existence of this overloaded state as an equilibrium of the client game, does not strongly depend on the precise features of the chosen M/M/1 model; rather, it occurs in any queue-based model in which the delay and loss are both increasing functions of the load, and the load is taken to be the client arrival rate times the average number of requests per client. It is an inherent problem of the linear cost structure, and is a direct consequence of the fact that a and b do not depend on the actual load.

[5]In the derivation process of (26), it is implicitly assumed that each request's delay and loss are described by the M/M/1-induced formulae (6)–(7), independently of the behavior of the other clients; in particular, it is assumed that no effect of 'synchronization' occurs in spite of all clients employing the same strategy. In the context of clients accessing a 'pool' of infinitely many independent servers, this requires that the server for each retry is chosen at random, rather than, say, having all clients follow the same sequence of servers when retransmitting.

VII. CONCLUSION

In this paper we presented an analytic study of the problem of finding an optimal timeout strategy. Instances of this generic problem appear in various networking contexts, notably in large-scale networks where failures are not accompanied by immediate negative acknowledgments. Selection of a timeout always involves a tradeoff between the desire for fast completion and the desire not to overload the network. In contrast with most common approaches, which put an absolute priority on avoiding 'spurious' retransmissions, we quantified the tradeoff by a cost per unit time (a) and a cost per retransmission (b), and studied the dependency of the optimal timeout strategy (i.e. the one that minimizes the expected cost until completion) on the price ratio ($\frac{a}{b}$), as well as the loss probability and delay distribution.

We have derived the formulae for computing the timeout strategy within the framework of independent and identically-distributed attempts. We then proceeded to establish some general properties that hold for any loss probability and delay distribution. In particular, we studied how the number of immediate attempts depends on the price ratio $\frac{a}{b}$. We indicated the existence of a threshold value for $\frac{a}{b}$, below which the first timeout is strictly positive; in addition, we revealed the asymptotic dependency of the number of initial attempts for $\frac{a}{b} \to \infty$, and proved that this number always grows at a slower rate than $\frac{a}{b}$ (depending on the delay distribution, the growth can be either logarithmic or polynomial with an order of less than 1). The effect of adding a deterministic propagation delay to a fixed stochastic delay distribution was also studied; in particular, we have proved that the number of initial attempts increases with the propagation delay, at a logarithmic rate. It is noteworthy that, even though these asymptotic results were formally proved with the aid of a 'greedy' behavior assumption (as explained at the end of section II), they were shown to hold in a couple of examples where the optimal strategy could be computed directly (i.e. without the above assumption).

We have also demonstrated, through an example of loss probability and delay distribution induced by an M/M/1 queue model, that the optimal timeout strategy may exhibit a phenomenon of positive feedback, such that a higher 'ambient' network load causes the end user to decrease the timeouts between retransmissions and thus to increase the load even further. Using a model of a noncooperative game, this phenomenon was shown to cause instability and a potential congestion collapse if the end users selfishly employ their optimal timeout strategies.

The prime importance of the theory and results of this paper – in particular, the general properties of timeout strategies proved – is in providing a fundamental framework for the analysis of further problems in the area of optimal timeouts. Nevertheless, even though the contribution of this paper is primarily theoretic, the results do have a generic significance for certain issues in network design. The property presented by the corollary of theorem 2 provides an upper (worst-case) bound on the client's retransmission rate. The threshold value of $\frac{a}{b}$ is an indicator of the design quality; thus, if $\frac{a}{b}$ is above the threshold value, meaning that the cost of retransmission is low and every request is attempted at least twice, the network is probably over-provisioned. Finally, if a and b are interpreted literally, as usage prices, then our game-theoretic analysis shows that a simple pricing scheme, with constant time and transmission prices independent of the actual network load, cannot avoid the possibility of an overloaded equilibrium.

While the analysis in this paper is made for general loss and delay distributions, its applicability is still limited by the assumption of independent attempts. It does not apply to the timeout strategy when the retrials cannot modify the attempted action without limit (or, in the extreme case, when the retried actions are all identical), since, in that case, the probability of success of any individual attempt depends on the outcome of previous attempts (and the time elapsed). In terms of the client-server model of this paper, for example, the analysis of the timeout strategy for connection establishment with a single server, connected to the client by a single path handling requests in a FIFO order, would require taking the temporal dependency between attempts into account. If the client has a selection of a small number of partially overlapping paths, the *spatial* dependency (even for requests transmitted at the same time) must be considered as well. The extension of this paper's model, so as to account for both types of dependency, is the subject of ongoing research.

REFERENCES

[1] J. Hastad, T. Leighton, and B. Rogoff. Analysis of backoff protocols for multiple access channels. *SIAM Journal on Computing*, 25(4):740–774, August 1996.

[2] L.A. Goldberg and P.D. MacKenzie. Analysis of practical backoff protocols for contention resolution with multiple servers. *Journal of Computer and System Sciences*, 58(1):232–258, February 1999.

[3] IEEE LAN/MAN Standards Committee. IEEE standard 802.3, October 2000. http://standards.ieee.org/getieee802/802.3.html.

[4] V. Paxson and M. Allman. RFC 2988: Computing TCP's retransmission timer, November 2000. Internet RFC.

[5] ATM Forum. ATM PNNI v1.0 specification, March 1996.

[6] editor J. Klensin. RFC 2821: Simple mail transfer protocol, April 2001. Internet RFC.

[7] R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, and T. Berners-Lee. Hypertext transfer protocol – HTTP 1.1, June 1999. Internet RFC.

[8] R. Ludwig and K. Sklower. The Eifel retransmission timer. *ACM Computer Communications Review*, 30(3), July 2000.

[9] R.J. Gibbens and F.P. Kelly. Resource pricing and the evolution of congestion control. *Automatica*, 35(12):1969–1985, December 1999.

[10] R.J. La and V. Anantharam. Charge-sensitive TCP and rate control in the Internet. In *Proc. IEEE INFOCOM*, pages 1166–1175, Tel-Aviv, Israel, March 2000.

[11] R. B. Myerson. *Game Theory: Analysis of Conflict*. Harvard University Press, Cambridge, MA, 1991.

[12] F.P. Kelly. Charging and accounting for bursty connections. In L.W. McKnight and J.P. Bailey, editors, *Internet Economics*, pages 253–278. MIT Press, Cambridge, MA, 1997.

[13] L. Kleinrock. *Queueing Systems*, volume 1. John Wiley & sons, 1975.

APPENDIX

This Appendix contains the detailed proofs and calculations omitted from the main text.

I. CALCULATIONS IN SECTION III-A

The following integrals appear in the derivation of the timeout formulae for the M/M/1 example in section III-A. Here, $d(t)$ and $p(t)$ are the delay distributions of the M/M/1 queue, given by (7)-(8).

$$\int_0^{\Delta t} p(t)\,dt = \frac{(1-\rho^m)\,e^{-(\mu-\lambda)\tau}\left[1-e^{-(\mu-\lambda)\Delta t}\right]}{(1-\rho^m)\,e^{-(\mu-\lambda)\tau}+\rho^m}; \tag{27}$$

$$1-\int_0^{\infty} p(t)\,dt = \frac{\rho^m}{(1-\rho^m)\,e^{-(\mu-\lambda)\tau}+\rho^m}; \tag{28}$$

$$\int_0^{t}(t-\theta)\,d(\theta)\,d\theta = t-\frac{1-e^{-(\mu-\lambda)t}}{\mu-\lambda}; \tag{29}$$

$$\int_0^{\infty}\left[\int_0^{t}(t-\theta)\,d(\theta)\,d\theta\right]\left[p(t)-p(t+\Delta t)\right]dt =$$
$$= \frac{(1-\rho^m)\,e^{-(\mu-\lambda)\tau}\left(1-e^{-(\mu-\lambda)\Delta t}\right)}{(1-\rho^m)\,e^{-(\mu-\lambda)\tau}+\rho^m}\cdot\frac{1}{2(\mu-\lambda)}. \tag{30}$$

II. PROOFS IN SECTION IV

Proof of proposition 1. The optimal timeout is the smallest value of τ that satisfies condition (5) for every $\Delta t > 0$. Suppose, to the contrary, that the optimal timeout is τ^* such that $0 < \tau^* < T$; thus (5) is satisfied for any Δt when $p'(t)_{|\tau^*}$ is substituted for $p(t)$. Note that, due to the structure of $p'(t)$ (formula (13)),

$$p'(t)_{|0} = \begin{cases} p'(t-\tau^*)_{|\tau^*} & t \geq \tau^*, \\ 0 & t < \tau^*, \end{cases} \tag{31}$$

consequently, for any $\Delta t > \tau^*$, the following identities hold:

$$\int_0^{\Delta t} p'(t)_{|0}\,dt = \int_0^{\Delta t-\tau^*} p'(t)_{|\tau^*}\,dt, \tag{32}$$

$$\int_0^{\infty} p'(t)_{|0}\,dt = \int_0^{\infty} p'(t)_{|\tau^*}\,dt, \tag{33}$$

and, for any monotonously non-decreasing function $\uparrow(t)$,

$$\int_0^{\infty} \uparrow(t)\left[p'(t)_{|0}-p'(t+\Delta t)_{|0}\right]dt \geq$$

$$\geq \int_0^\infty \uparrow (t) \left[p'(t)_{|\tau^*} - p'(t+\Delta t - \tau^*)_{|\tau^*}\right] dt. \quad (34)$$

Substituting relations (32)–(34) into (5), we obtain that the fulfillment of the condition for any $\Delta t - \tau^*$ with $p'(t)_{|\tau^*}$ implies its fulfillment for Δt with $p'(t)_{|0}$. It follows that a timeout of 0 satisfies the respective condition, contradicting the assumption that the smallest value of τ satisfying it is τ^*. □

Proof of proposition 2. Suppose that the optimal timeout for the case when there is no propagation delay is τ^*; thus, condition (5) is not satisfied for any $\tau < \tau^*$. Now consider the same condition in the presence of a propagation delay, for a timeout of $\tau + T$. Since $p'(t)_{|\tau+T} = p(t)_{|\tau}$ by (13), the values of $p(t)$ in the condition remain unchanged.

As for the *a priori* function $d(t)$, it can be seen from (12) that, for any t, $\int_0^t (t-\theta)\, d'(\theta)\, d\theta < \int_0^t (t-\theta)\, d(\theta)\, d\theta$; substituting this inequality into (5), it is easily seen that the condition continues not to be satisfied. Consequently, the optimal timeout is at least as high as $\tau^* + T$. □

Proof of theorem 1. We begin by noting that the term

$$\int_T^\infty \left[\int_T^t (t-\theta)\, d'(\theta)\, d\theta\right] \left[p'(t)_{|0} - p'(t+\Delta t)_{|0}\right] dt,$$

can be shown, by swapping the order of integration followed by the variable substitutions of $t' = t - \theta$ and $\theta' = \theta - T$, to be equal to

$$\int_0^\infty \left[\int_0^\infty t' \left[p(t'+\theta')_{|0} - p(t'+\theta'+\Delta t)_{|0}\right] dt'\right] d(\theta')\, d\theta',$$

that is, to be independent of T. Hence, for convenience, it can be denoted by $\Phi(\Delta t)$. It can further be shown that

$$\begin{aligned}\Phi(\Delta t) \leq & \int_0^\infty \left[\int_0^\infty t' p(t'+\theta')_{|0}\, dt' - \right.\\ & \left. - \int_0^\infty (t'-\Delta t)\, p(t'+\theta')_{|0}\, dt'\right] d(\theta')\, d\theta' = \\ & = \int_0^\infty \left[\Delta t \cdot \int_0^\infty p(t'+\theta')_{|0}\, dt'\right] d(\theta')\, d\theta' \leq \\ \leq & \int_0^\infty \left[\Delta t \cdot \int_0^\infty p(t')_{|0}\, dt'\right] d(\theta')\, d\theta' = \Delta t \cdot \int_0^\infty p(t')_{|0}\, dt'.\end{aligned} \quad (35)$$

Substituting $p'(t)_{|0}$ into (5), we get that, in order for the timeout to be 0, the price ratio must satisfy

$$\frac{a}{b} \geq \sup_{\Delta t} \frac{\int_T^{\Delta t} p'(t)_{|0}\, dt}{\left(1 - \int_T^\infty p'(t)_{|0}\, dt\right) \Delta t + \Phi(\Delta t)} =$$

$$\sup_{\Delta t'} \frac{\int_0^{\Delta t'} p(t')_{|0}\, dt'}{\left(1 - \int_0^\infty p(t)_{|0}\, dt\right)(\Delta t' + T) + \Phi(\Delta t' + T)}, \quad (36)$$

where we have used the auxiliary notation $\Delta t' \triangleq \Delta t - T$. Now, the right-hand side of (36) can be bounded from above by $\frac{1}{\left(1-\int_0^\infty p(t)_{|0} dt\right)\cdot T}$ and from below (using (35)) by the value of $\frac{\int_0^{\Delta t'} p(t)_{|0} dt}{\Delta t' + T}$ with any specific $\Delta t'$ for which $\int_0^{\Delta t'} p(t)_{|0}\, dt > 0$. Hence, the threshold is indeed $\Theta\left(\frac{1}{T}\right)$ for $T \to \infty$. □

III. PROOFS IN SECTION V

A. Proof of theorem 2

To prove theorem 2, we show that when $p(t)_{|\vec{\tau}_2}$ is replaced by $p(t)_{\vec{\tau}_1}$, for any specific Δt, the right-hand side of condition (5) becomes smaller whereas the left-hand side becomes larger; consequently, for any Δt for which the condition is satisfied with $p(t)_{|\vec{\tau}_2}$, it is satisfied for the same Δt with $p(t)_{|\vec{\tau}_1}$ as well. The two following lemmas establish the corresponding inequalities.

Lemma A-1. *If $\vec{\tau}_1$ is a subsequence of $\vec{\tau}_2$, then $P(t)_{|\vec{\tau}_1} \leq P(t)_{|\vec{\tau}_2}$ for any t.*

Proof. Consider the product at the right-hand side of (20'). If $\vec{\tau}_1$ is a subsequence of $\vec{\tau}_2$, then, for any t, the product for $\vec{\tau}_1$ contains only a part of the terms contained in the product for $\vec{\tau}_2$, and consequently, is not smaller (all the product terms are less than or equal to 1). Therefore, $P(t)_{|\vec{\tau}_1} \leq P(t)_{|\vec{\tau}_2}$. □

Corollary. *For any Δt, the right-hand side of condition* (5) *is smaller for $\vec{\tau}_1$ than for $\vec{\tau}_2$.*

Lemma A-2. *For any Δt, the left-hand side of* (5) *is larger for $\vec{\tau}_1$ than for $\vec{\tau}_2$.*

Proof. For notational convenience, define $d(t) = 0$ for all $t < 0$; this definition allows to write the following transformation of the first term of the left-hand side of (5):

$$\int_0^\infty \left[\int_0^t (t-\theta)\, d(\theta)\, d\theta\right] \left[p(t)_{|\vec{\tau}} - p(t+\Delta t)_{|\vec{\tau}}\right] dt =$$

$$\int_0^\infty \left[\int_0^t (t-\theta)\, d(\theta)\, d\theta - \int_0^{t-\Delta t} (t - \Delta t - \theta)\, d(\theta)\, d\theta\right] p(t)_{|\vec{\tau}}\, dt.$$

Now use the rule of integration in parts, which states $\int u(t)\, v'(t)\, dt = u(t)\, v(t) - \int u'(t)\, v(t)\, dt$, for the functions $u(t) = \int_0^t (t-\theta)\, d(\theta)\, d\theta - \int_0^{t-\Delta t} (t-\Delta t-\theta)\, d(\theta)\, d\theta$ and $v(t) = P(t)_{|\vec{\tau}}$. Before applying the rule, we precompute, using the rule of the derivative of an integral and the chain rule,

$$u'(t) = \int_0^t d(\theta)\, d\theta - \int_0^{t-\Delta t} d(\theta)\, d\theta = \int_{t-\Delta t}^t d(\theta)\, d\theta;$$

$$\lim_{t\to\infty} u(t) = \int_0^\infty \Delta t \cdot d(\theta)\, d\theta = \Delta t.$$

Therefore,

$$\int_0^\infty \left[\int_0^t (t-\theta)\, d(\theta)\, d\theta - \right.$$
$$\left. - \int_0^{t-\Delta t} (t-\Delta t-\theta)\, d(\theta)\, d\theta \right] p(t)_{|\vec{\tau}}\, dt =$$
$$= P(\infty)_{|\vec{\tau}} \cdot \Delta t - \int_0^\infty \left[\int_{t-\Delta t}^t d(\theta)\, d\theta \right] P(t)_{|\vec{\tau}}\, dt.$$

The entire left-hand side of (5) can now be written as follows:

$$a \int_0^\infty \left[\int_0^t (t-\theta)\, d(\theta)\, d\theta \right] \left[p(t)_{|\vec{\tau}} - p(t+\Delta t)_{|\vec{\tau}} \right] dt +$$
$$+ a\Delta t \left(1 - \int_0^\infty p(t)_{\vec{\tau}_1}\, dt \right) =$$
$$= a \left\{ P(\infty)_{|\vec{\tau}} \cdot \Delta t - \int_0^\infty \left[\int_{t-\Delta t}^t d(\theta)\, d\theta \right] P(t)_{|\vec{\tau}}\, dt \right\} +$$
$$+ a\Delta t \left(1 - P(\infty)_{|\vec{\tau}} \right) =$$
$$= a \left\{ \Delta t - \int_0^\infty \left[\int_{t-\Delta t}^t d(\theta)\, d\theta \right] P(t)_{|\vec{\tau}}\, dt \right\}, \quad (37)$$

and it becomes clear that, since $P(t)_{|\vec{\tau}_1} \leq P(t)_{|\vec{\tau}_2}$ for any t (lemma A-1), (37) is larger for $\vec{\tau}_1$ than for $\vec{\tau}_2$. □

B. Proof of theorems 3–4

The proof of theorems 3–4 passes through a series of intermediate lemmas, which we proceed to establish. In fact, some of these lemmas are not specifically related to the topic of this text, but to properties of random sequences in general. Nevertheless, we prove them here for completeness.

For convenience, we recall a couple of notations from the main text and introduce some new ones:

$D(t)$ - the cumulative delay distribution function of a single request, $D(t) = \int_0^t d(t)\, dt$. In case that $D(t)$ has discontinuities, we assume that it is upper-semicontinuous.

t_0 - the infimum of the values that the delay of a single request can take, i.e. $t_0 = \inf\{t | D(t) > 0\}$.

$D^{-1}(x), 0 \leq x < 1$ - the inverse cumulative distribution function, defined as $D^{-1}(x) = \sup\{t | D(t) \leq x\}$. Thus, $D^{-1}(x)$ is monotonously non-decreasing and upper-semicontinuous. In particular, note that $D^{-1}(0) = \sup\{t | D(t) = 0\} = \inf\{t | D(t) > 0\} = t_0$.

M_k - the random variable which is the minimum of k independent identically distributed (i.i.d.) random variables, each distributed according to $d(t)$.

$D_k(t)$ - the cumulative distribution function of M_k; obviously, $D_k(t) = 1 - [1 - D(t)]^k$.

Lemma A-3. *If* $\mathrm{E}[M_k]$ *is finite, then*

$$\int_0^\infty \left[\int_0^t (t-\theta)\, d(\theta)\, d\theta \right] p(t)_{|\vec{0}^k}\, dt =$$
$$= \left(1 - L^k\right) \left\{ \mathrm{E}[M_k] - \mathrm{E}[M_{k+1}] \right\},$$

where L *is the loss probability of a single request,* $L = 1 - \int_0^\infty p(t)_{|0}\, dt$.

Proof. Since $D_{k+1}(t) = 1 - [1 - D(t)][1 - D_k(t)]$, it follows that

$$\mathrm{E}[M_{k+1}] = \int_0^\infty t \cdot dD_{k+1}(t) =$$
$$= \int_0^\infty t \cdot [1 - D(t)]\, dD_k(t) + \int_0^\infty t \cdot [1 - D_k(t)]\, dD(t),$$

and, consequently,

$$\mathrm{E}[M_k] - \mathrm{E}[M_{k+1}] =$$
$$= \int_0^\infty t \cdot D(t)\, dD_k(t) - \int_0^\infty t \cdot [1 - D_k(t)]\, dD(t).$$

Replacing $D(t)$ by $\int_0^t dD(\theta)$ and $[1 - D_k(t)]$ by $\int_t^\infty dD_k(\theta)$, followed by swapping the order of integration and the names of the integration variables, this expression is shown to be equal to

$$\mathrm{E}[M_k] - \mathrm{E}[M_{k+1}] =$$
$$= \int_0^\infty \int_0^t t \cdot dD(\theta)\, dD_k(t) - \int_0^\infty \int_t^\infty t \cdot dD_k(\theta)\, dD(t) =$$
$$= \int_0^\infty \int_0^t t \cdot dD(\theta)\, dD_k(t) - \int_0^\infty \int_0^t \theta \cdot dD(\theta)\, dD_k(t) =$$
$$= \int_0^\infty \left[\int_0^t (t-\theta)\, dD(\theta) \right] dD_k(t).$$

It remains to observe that $d(\theta)\, d\theta = dD(\theta)$ and $p(t)_{|\vec{0}^k}\, dt = dP(t)_{|\vec{0}^k} = \left(1 - L^k\right) dD_k(t)$, and the lemma follows. □

Lemma A-4. *If* $\mathrm{E}[M_k]$ *is finite, then*

$$\mathrm{E}[M_k] = \int_0^1 D^{-1}\left(1 - x^{1/k}\right) dx.$$

Proof. As noted above, $D_k(t) = 1 - [1 - D(t)]^k$, hence $t = D^{-1}\left(1 - [1 - D_k(t)]^{1/k}\right)$. Denote $x = 1 - D_k(t)$; then, adjusting the integral bounds accordingly, we get $\mathrm{E}[M_k] = \int_0^\infty t \cdot dD_k(t) = \int_0^1 D^{-1}\left(1 - x^{1/k}\right) dx$. □

Lemma A-5. *If $D(t_0) > 0$ and $D(t)$ has a finite mean, then $\mathrm{E}[M_k] - \mathrm{E}[M_{k+1}]$ behaves as $O\left(k \cdot (1 - D(t_0))^k\right)$ for $k \to \infty$.*

Proof. As stated above, $D_k(t) = 1 - [1 - D(t)]^k$. Therefore,

$$\mathrm{E}[M_k] = \int_0^{\infty} t \cdot dD_k(t) = \int_{t_0^-}^{\infty} t \cdot k [1 - D(t)]^{k-1} dD(t)$$

and it follows that

$$\begin{aligned}
&\mathrm{E}[M_k] - \mathrm{E}[M_{k+1}] = \\
&= \int_{t_0^-}^{\infty} t \cdot \left\{ k [1 - D(t)]^{k-1} - (k+1)[1 - D(t)]^k \right\} dD(t) \le \\
&\le \int_{t_0^-}^{\infty} t \cdot (k+1)[1 - D(t)]^{k-1} D(t)\, dD(t) \le \\
&\le (k+1)[1 - D(t_0)]^{k-1} \int_{t_0^-}^{\infty} t \cdot dD(t) = \\
&= (k+1)[1 - D(t_0)]^{k-1} \cdot \mathrm{E}[M_1].
\end{aligned}$$

Since $\mathrm{E}[M_1]$ (the mean of $D(t)$) is finite, this means that $\mathrm{E}[M_k] - \mathrm{E}[M_{k+1}] = O\left(k \cdot (1 - D(t_0))^k\right)$. □

Lemma A-6. *If $D(t)$ has a finite mean, $D(t_0) = 0$, and can be expanded around t_0 as $D(t) = \sum_i C_i \cdot (t - t_0)^{\alpha_i}$, for certain constants C_i such that $C_1 > 0$ and $0 < \alpha_1 < \alpha_2 < \ldots$, then $\mathrm{E}[M_k] - \mathrm{E}[M_{k+1}]$ behaves as $O\left(k^{-\frac{\alpha_1+1}{\alpha_1}}\right)$ for $k \to \infty$.*

Proof. If $D(t)$ is expandable around t_0 as given, then its inverse $D^{-1}(x)$ is can be expanded around 0 to

$$D^{-1}(x) = D^{-1}(0) + \sum_i \Lambda_i x^{\beta_i} = t_0 + \sum_i \Lambda_i x^{\beta_i}, \qquad (38)$$

where $\frac{1}{\alpha_1} = \beta_1 < \beta_2 < \ldots$ and $\Lambda_1 = \frac{1}{C_1^{1/\alpha_1}}$. Therefore, there is a value of $\epsilon > 0$ such that the series in (38) converges uniformly for any $0 \le x \le \epsilon$.

Applying lemma A-4, we have

$$\begin{aligned}
&\mathrm{E}[M_k] - \mathrm{E}[M_{k+1}] = \\
&= \int_0^1 \left[D^{-1}\left(1 - x^{1/k}\right) - D^{-1}\left(1 - x^{1/(k+1)}\right) \right] dx = \\
&= \left.\begin{aligned} &\left[\int_0^{(1-\epsilon)^k} D^{-1}\left(1 - x^{1/k}\right) dx - \int_0^{(1-\epsilon)^{k+1}} D^{-1}\left(1 - x^{1/(k+1)}\right) dx \right] + \end{aligned}\right\} \mathbf{T1} \\
&\left.\begin{aligned} &+ \left[\int_{(1-\epsilon)^k}^1 \sum_i \Lambda_i \left(1 - x^{1/k}\right)^{\beta_i} dx - \int_{(1-\epsilon)^{k+1}}^1 \sum_i \Lambda_i \left(1 - x^{1/(k+1)}\right)^{\beta_i} dx \right] \end{aligned}\right\} \mathbf{T2} \qquad (39)
\end{aligned}$$

Consider the first term of the rightmost side of (39) (denoted by $\mathbf{T1}$). Using a variable substitution of $x' = 1 - x^{1/k}$, it can be bounded as follows:

$$\begin{aligned}
\mathbf{T1} &\le \int_0^{(1-\epsilon)^k} \left[D^{-1}\left(1 - x^{1/k}\right) \right] dx = \\
&= \int_\epsilon^1 D^{-1}(x') \cdot k(1 - x')^{k-1} dx' \le \\
&\le k(1-\epsilon)^{k-1} \cdot \int_\epsilon^1 D^{-1}(x')\, dx' \le k(1-\epsilon)^{k-1} \cdot \mathrm{E}[M_1].
\end{aligned}$$

Since $\mathrm{E}[M_1]$ is finite, we conclude that the first term of the rightmost side of (39) decays as $O\left(k(1-\epsilon)^k\right)$ for $k \to \infty$.

Now consider the second term of the rightmost side of (39) (denoted by $\mathbf{T2}$). Our objective will be to prove that it is $O\left(k^{-(\beta_1+1)}\right)$; accordingly, we consider the limit of it divided by $\left(k^{-(\beta_1+1)}\right)$, for $k \to \infty$. Swapping the order of integration and summation (permissible since the series converge uniformly), and using the variable substitution of $x' = 1 - x^{1/k}$ (or, correspondingly, $x' = 1 - x^{1/(k+1)}$) again, we get

$$\begin{aligned}
&\lim_{k\to\infty} k^{\beta_1+1} \cdot \mathbf{T2} = \\
&= \lim_{k\to\infty} k^{\beta_1+1} \left[\sum_i \Lambda_i \int_0^\epsilon x'^{\beta_i} \cdot k(1-x')^{k-1} dx' - \right. \\
&\left. - \sum_i \Lambda_i \int_0^\epsilon x'^{\beta_i} \cdot (k+1)(1-x')^k dx' \right] = \\
&= \sum_i \Lambda_i \lim_{k\to\infty} k^{\beta_1+1} \int_0^\epsilon x'^{\beta_i} \left[k(1-x')^{k-1} - \right. \\
&\left. - (k+1)(1-x')^k \right] dx \le \\
&\le \sum_i |\Lambda_i| \lim_{k\to\infty} k^{\beta_1+1} \int_0^1 x'^{\beta_i} \left[k(1-x')^{k-1} - \right. \\
&\left. - (k+1)(1-x')^k \right] dx = \\
&= \sum_i |\Lambda_i| \lim_{k\to\infty} k^{\beta_1+1} \left[k \frac{\Gamma(\beta_i+1)\Gamma(k)}{\Gamma(\beta_i+k+1)} - \right. \\
&\left. - (k+1) \frac{\Gamma(\beta_i+1)\Gamma(k+1)}{\Gamma(\beta_i+k+2)} \right] = \\
&= \sum_i |\Lambda_i| \lim_{k\to\infty} k^{\beta_1+1} \left[\Gamma(\beta_i+1) \frac{\beta_i \Gamma(k+1)}{\Gamma(\beta_i+k+2)} \right] =
\end{aligned}$$

$$= \sum_i |\Lambda_i| \lim_{k\to\infty} k^{\beta_1+1} \left[\Gamma(\beta_i+1)\frac{\beta_i}{k^{\beta_i+1}}\right] = \\ = |\Lambda_1|\,\beta_1\Gamma(\beta_1+1),$$

where we have used the well-known formulae $\frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)} = \int_0^1 x^{m-1}(1-x)^{n-1}dx$ and $n\Gamma(n) = \Gamma(n+1)$, $\Gamma(\cdot)$ being the gamma-function. We conclude that the aforementioned limit is finite, and, consequently, the second term of the rightmost side of (39) is indeed $O\left(k^{-(\beta_1+1)}\right)$. Altogether, therefore, $\mathrm{E}[M_k] - \mathrm{E}[M_{k+1}]$ behaves as $O\left(k(1-\epsilon)^k + k^{-(\beta_1+1)}\right) = O\left(k^{-(\beta_1+1)}\right) = O\left(k^{-\frac{\alpha_1+1}{\alpha_1}}\right)$, which proves the lemma. □

Lemma A-7. *For any Δt, define the distribution $\hat{D}(t) \triangleq \begin{cases} D(t) & t<\Delta t \\ 1 & t\ge\Delta t\end{cases}$ (the delay distribution truncated to Δt), and the functions $\hat{d}(t)$, $\hat{p}(t)_{|\vec{\tau}}$ accordingly. Then*

$$\int_0^\infty \left[\int_0^t (t-\theta)\,d(\theta)\,d\theta\right]\left[p(t)_{|\vec{0}^k} - p(t+\Delta t)_{\vec{0}^k}\right]dt- \\ -\int_0^\infty \left[\int_0^t (t-\theta)\,\hat{d}(\theta)\,d\theta\right]\left[\hat{p}(t)_{|\vec{0}^k} - \hat{p}(t+\Delta t)_{\vec{0}^k}\right]dt \le \\ \le \Delta t\cdot[1-D(\Delta t)]^k .$$

Proof. Note that $\hat{p}(t+\Delta t)_{|\vec{0}^k} = 0$ for any $t>0$, while $p(t)_{|\vec{0}^k} = \hat{p}(t)_{|\vec{0}^k}$ for any $t<\Delta t$, and $d(\theta) = \hat{d}(\theta)$ for any $\theta<\Delta t$. Consequently,

$$\int_0^\infty \left[\int_0^t (t-\theta)\,d(\theta)\,d\theta\right]\left[p(t)_{|\vec{0}^k} - p(t+\Delta t)_{\vec{0}^k}\right]dt- \\ -\int_0^\infty \left[\int_0^t (t-\theta)\,\hat{d}(\theta)\,d\theta\right]\left[\hat{p}(t)_{|\vec{0}^k} - \hat{p}(t+\Delta t)_{\vec{0}^k}\right]dt = \\ = \int_{\Delta t}^\infty \left[\int_{\Delta t}^t (t-\theta)\,d(\theta)\,d\theta\right] p(t)_{|\vec{0}^k}\,dt- \\ -\int_0^\infty \left[\int_0^t (t-\theta)\,d(\theta)\,d\theta\right] p(t+\Delta t)_{\vec{0}^k}\,dt \le \\ \le \int_{\Delta t}^\infty \left[\int_{\Delta t}^t (t-\theta)\,d(\theta)\,d\theta\right]\left[p(t)_{|\vec{0}^k} - p(t+\Delta t)_{\vec{0}^k}\right]dt = \\ = \int_{\Delta t}^\infty \left[\int_\theta^\infty (t-\theta)\left[p(t)_{|\vec{0}^k} - p(t+\Delta t)_{\vec{0}^k}\right]dt\right] d(\theta)\,d\theta = \\ = \int_{\Delta t}^\infty \left[\int_\theta^\infty (t-\theta)\,p(t)_{|\vec{0}^k}\,dt- \\ -\int_\theta^\infty (t-\theta)\,p(t+\Delta t)_{\vec{0}^k}\,dt\right] d(\theta)\,d\theta \le \\ \le \int_{\Delta t}^\infty \left[\int_\theta^\infty (t-\theta)\,p(t)_{|\vec{0}^k}\,dt- \\ -\int_\theta^\infty (t-\theta-\Delta t)\,p(t)_{\vec{0}^k}\,dt\right] d(\theta)\,d\theta = \\ = \int_{\Delta t}^\infty \left[\Delta t\cdot\int_\theta^\infty p(t)_{|\vec{0}^k}\,dt\right] d(\theta)\,d\theta \le \\ \le \int_{\Delta t}^\infty \left[\Delta t\cdot\int_{\Delta t}^\infty p(t)_{|\vec{0}^k}\,dt\right] d(\theta)\,d\theta = \\ = \Delta t\left(1-L^k\right)[1-D(\Delta t)]^k\cdot\int_{\Delta t}^\infty d(\theta)\,d\theta \le \\ \le \Delta t\,[1-D(\Delta t)]^k .$$

□

Lemma A-8. *For any k and any $\Delta t>0$,*

$$\int_0^{\Delta t} p(t)_{|\vec{0}^k}\,dt \ge \int_0^{\Delta t} p(t)_{|0}\,dt.$$

Proof. By definition of the *a posteriori* cumulative distribution function (formula (20')),

$$\int_0^{\Delta t} p(t)_{|\vec{0}^k}\,dt = P(\Delta t)_{|\vec{0}^k} = 1-\left[1-P(t)_{|0}\right]^k \ge \\ \ge 1-\left[1-P(t)_{|0}\right] = P(t)_{|0} = \int_0^{\Delta t} p(t)_{|0}\,dt.$$

□

Proof of theorem 3. Choose any specific Δt such that $D(\Delta t)>0$ (in other words, any $\Delta t>t_0$), define the distribution $\hat{D}(t)$ as in lemma A-7, and define $\hat{M}_k$ to be the random variable which is the minimum of k independent identically distributed (i.i.d.) random variables, each distributed according to $\hat{D}(t)$. Obviously, $\hat{D}(t)$ has a finite mean, and lemmas A-3–A-6 are therefore applicable for $\hat{D}(t)$ and $\hat{M}_k$.

Suppose that the request is retransmitted immediately k times (excluding the first transmission). Then condition (5) must be satisfied with $p(t)_{\vec{0}^k}$ for every Δt, and, in particular, for the specific Δt chosen in the previous paragraph. Applying all the previous lemmas, we finally get

$$a\left(1-L^k\right)\left\{\mathrm{E}\left[\hat{M}_k\right] - \mathrm{E}\left[\hat{M}_{k+1}\right]\right\} + \\ + a\Delta t\,[1-D(\Delta t)]^k + a\Delta t\cdot L^k = \\ = a\int_0^\infty \left[\int_0^t (t-\theta)\,\hat{d}(\theta)\,d\theta\right]\hat{p}(t)_{|\vec{0}^k}\,dt+ \\ + a\Delta t\,[1-D(\Delta t)]^k + a\Delta t\cdot L^k \ge$$

$$\geq a \int_0^\infty \left[\int_0^t (t-\theta)\,\hat{d}(\theta)\,d\theta \right] \left[\hat{p}(t)_{|\vec{0}^k} - \hat{p}(t+\Delta t)_{|\vec{0}^k} \right] dt +$$
$$+ a\Delta t\,[1-D(\Delta t)]^k + a\Delta t \cdot L^k \geq$$
$$\geq a \int_0^\infty \left[\int_0^t (t-\theta)\,d(\theta)\,d\theta \right] \left[p(t)_{|\vec{0}^k} - p(t+\Delta t)_{|\vec{0}^k} \right] dt +$$
$$+ a\Delta t \cdot L^k \geq b \int_0^{\Delta t} p(t)_{|\vec{0}^k}\,dt \geq b \int_0^{\Delta t} p(t)_{|0}\,dt. \quad (40)$$

Now consider the theorem cases.

a) If lemma A-5 applies, then the leftmost side of (40) is $a \cdot O\left(k \cdot (1-D(t_0))^k + [1-D(\Delta t)]^k + L^k\right)$. This implies $\frac{a}{b} = \Omega\left(\min\left\{\frac{1}{k}\left[\frac{1}{1-D(t_0)}\right]^k, \left(\frac{1}{1-D(\Delta t)}\right)^k, \left(\frac{1}{L}\right)^k\right\}\right)$, hence $k = O\left(\log \frac{a}{b}\right)$.

b) If lemma A-6 applies, then the leftmost side of (40) is $a \cdot O\left(k^{-\frac{\alpha_1+1}{\alpha_1}} + [1-D(\Delta t)]^k + L^k\right) = a \cdot O\left(k^{-\frac{\alpha_1+1}{\alpha_1}}\right)$. Consequently, $\frac{a}{b} = \Omega\left(k^{\frac{\alpha_1+1}{\alpha_1}}\right)$, hence $k = O\left(\left(\frac{a}{b}\right)^{\frac{\alpha_1}{\alpha_1+1}}\right)$. $\square$

Proof of theorem 4. As in the proof of theorem 3, choose an arbitrary Δt such that $D(\Delta t) > 0$. If the request is retransmitted immediately k times (excluding the first transmission), then condition (5) must be satisfied with $p'(t)_{\vec{0}^k}$, in particular, for $\Delta t + T$.[6] Therefore

$$a\left(1-L^k\right)\left\{\mathrm{E}\left[\hat{M}_k\right] - \mathrm{E}\left[\hat{M}_{k+1}\right]\right\} +$$
$$+ a(\Delta t + T)[1-D(\Delta t)]^k + a(\Delta t + T)\cdot L^k =$$
$$= a \int_0^\infty \left[\int_0^t (t-\theta)\,\hat{d}(\theta)\,d\theta \right] \hat{p}(t)_{|\vec{0}^k}\,dt +$$
$$+ a(\Delta t + T)[1-D(\Delta t)]^k + a(\Delta t + T)\cdot L^k \geq$$
$$\geq a \int_0^\infty \left[\int_0^t (t-\theta)\,\hat{d}(\theta)\,d\theta \right] \left[\hat{p}(t)_{|\vec{0}^k} - \hat{p}(t+\Delta t+T)_{|\vec{0}^k} \right] dt +$$
$$+ a(\Delta t + T)[1-D(\Delta t + T)]^k + a(\Delta t + T)\cdot L^k \geq$$
$$\geq a \int_0^\infty \left[\int_0^t (t-\theta)\,d(\theta)\,d\theta \right] \left[p(t)_{|\vec{0}^k} - p(t+\Delta t+T)_{|\vec{0}^k} \right] dt +$$
$$+ a(\Delta t + T)\cdot L^k =$$
$$= a \int_0^\infty \left[\int_0^t (t-\theta)\,d'(\theta)\,d\theta \right] \left[p'(t)_{|\vec{0}^k} - p'(t+\Delta t+T)_{|\vec{0}^k} \right] dt$$
$$+ a(\Delta t + T)\cdot L^k \geq b \cdot \int_0^{\Delta t + T} p'(t)_{|\vec{0}^k}\,dt \geq$$

[6]In this proof, we use the notation of section IV; recall definitions (12)–(13).

$$\geq b \cdot \int_0^{\Delta t + T} p'(t)_{|0}\,dt = b \cdot \int_0^{\Delta t} p(t)_{|0}\,dt, \quad (41)$$

or

$$aT\left\{[1-D(\Delta t)]^k + L^k\right\} \geq b \cdot \int_0^{\Delta t} p(t)_{|0}\,dt -$$
$$- a\left(1-L^k\right)\left\{\mathrm{E}\left[\hat{M}_k\right] - \mathrm{E}\left[\hat{M}_{k+1}\right]\right\} -$$
$$- a\Delta t\left\{[1-D(\Delta t)]^k + L^k\right\}. \quad (42)$$

As $T \to \infty$ and $k \to \infty$, all but the first term at the right-hand side of (42) vanish, and it tends to $b \cdot \int_0^{\Delta t} p(t)_{|0}\,dt$, which is strictly positive by the selection of Δt. Consequently, $T = \Omega\left(\min\left\{\left(\frac{1}{1-D(\Delta t)}\right)^k, \left(\frac{1}{L}\right)^k\right\}\right)$, hence $k = O(\log T)$. $\square$

Lavy Libman received his B.Sc. degrees (summa cum laude) in Electrical Engineering and in Computer Engineering and his M.Sc. degree in Electrical Engineering from the Technion – Israel Institute of Technology, Haifa, Israel, in 1992 and 1997, respectively. Since 1999, he is a Ph.D. candidate in the Department of Electrical Engineering at the Technion. In 1992–1999 he served as an engineer in the Israel Defense Forces (with a research rank since 1996).

His current research interests are in game-theoretic and pricing aspects of network architecture and protocols, in particular flow and congestion control, routing, and provision of QoS.

Mr. Libman received the Merlin Award for distinguished students in Electrical Engineering, the Knesset (Israeli Parliament) Award for distinguished students, and the Wolff prize for distiguished Ph.D. students.

Mr. Libman is a member of the Executive Committee of IEEE Infocom'2002.

His e-mail address is: libman@tx.technion.ac.il

Ariel Orda (S'94-M'92-SM'97) received the B.Sc. (summa cum laude), M.Sc., and D.Sc. degrees in Electrical Engineering from the Technion – Israel Institute of Technology, Haifa, Israel, in 1983, 1985, and 1991, respectively.

Since 1994, he has been with the Department of Electrical Engineering at the Technion, where he is currently an Associate Professor and the Academic Head of the Computer Networking Laboratory. In 1993–1994, he was with the Center for Telecommunication Research, Columbia University, New York, NY, as a Visiting Scientist, and with AT&T Bell Laboratories, Murray Hill, NJ, as a Scientific Consultant. During the summers of 1992 and 1995–1999, he held visiting positions at AT&T Bell Laboratories, Murray Hill, NJ, IBM Watson, Hawthorne, NY, and Bell Laboratories - Lucent Technologies, Holmdel, NJ. Since 1991 he has held several consulting positions with the Israeli industry.

His current research interests include QoS routing, the application of game theory to computer networking, network pricing, and distributed network algorithms.

Dr. Orda received the Award of the Chief Scientist in the Ministry of Communication in Israel, a Gutwirth Award for outstanding distinction, the Research Award of the Association of Computer and Electronic Industries in Israel, and the Jacknow Award for Excellence in Teaching.

Dr. Orda is an Associate Editor of the IEEE/ACM Transactions on Networking and a Program Co-Chair of IEEE Infocom'2002.

His e-mail address is: ariel@ee.technion.ac.il