

Asymptotically Efficient Backlog Estimators for Frame Aloha

Luca Barletta and Flaminio Borgonovo

Dipartimento di Elettronica, Informazione e Bioingegneria
Politecnico di Milano, Italy

luca.barletta@polimi.it

Abstract—The dynamic frame Aloha protocol has been shown to reach efficiency e^{-1} when the number N of tags to be identified is known and approaches infinity. Results available in literature do not achieve efficiency e^{-1} for large N when N is unknown. In this paper we analytically show that the best tag estimation procedure let the protocol reach efficiency e^{-1} . However, the convergence to this result shows an oscillatory behavior that on average goes like $1/\ln N$. The oscillation can be reduced at will at the expenses of the convergence speed. A practical algorithm is proposed for achieving the claimed efficiency and numerical simulations are shown to validate the results.

I. INTRODUCTION

Dynamic Frame Aloha (DFA) is a collision resolution protocol proposed in 1983 [1] for satellite multiple access, and rediscovered some time ago for radio-frequency identification (RFID) [2], [3]. DFA is a popular topic among researchers, as it is testified by the very many publications over decades up to nowadays. Many papers refer to variations in the basic algorithm and protocol in order to improve the efficiency. Others study the performance taking into account new physical capabilities, such as receiver capture, or limitations, e.g., energy. However, referring to the basic protocol as it is standardized, in spite of the many works about its best configuration and performance, many basic issues remain to be investigated.

In RFID, DFA is used to identify an unknown number of tags N , that potentially can be very large, and in order to minimize the average identification time $L(N)$ some parameters must be optimized. Crucial to the performance, that varies with N , is the continuous estimation, during the entire process, of the number of tags yet to be identified. A number of estimators have been produced in the past, see for example, [1], [4]–[11]; interesting comparisons are given in [10], [11]. The attained results are rarely comparable because of different settings, and are somehow limited, most of them referring to a restricted range of N , or frame length, in some cases looking for some simple practical backlog estimators; their performance have been evaluated often by simulations, with the exception of [12], [13], and in few cases by a numerical evaluation of some model [14].

Although applications that relate to some large N are possible, only few works [12], [13], [15] have considered the asymptotic behavior of DFA. Asymptotic conditions provide a well defined case study, allowing the comparison of different

alternatives; it also gives away with the initial frame-setting problem, since their influence on $L(N)$ becomes negligible. The limit case in which a perfect estimate is used has been thoroughly and analytically investigated in [15], proving an asymptotic efficiency of e^{-1} . However, other important theoretical issues, such as the best estimation procedure and whether or not the DFA performance can reach the theoretical maximum, remain unanswered.

In this paper we continue the attempt to provide asymptotic answers, here investigating the best estimation process, showing that it can be split into two phases with two different estimates, for each of which we produce the best estimator. The former is the asymptotic Maximum Likelihood Estimator of the population size N , and the latter is proven to be the simple and long established Schoute's estimate [1]. Furthermore, we prove that such estimators are consistent and allow DFA to reach asymptotic efficiency e^{-1} , although with an overhead whose weight that decreases as $1/\ln N$. Finally, we show that our results can be extended to finite values of N , and that, at least for $N \geq 1000$, the efficiency of DFA, though less than its asymptotic value, is still the best possible.

The paper is organized as follows. Some basics on DFA, background, and problems, together with explanations on our approach, are given in Sec. II. In Sec. III we analyze the problem of estimating the initial population size N , while in Sec. IV we prove the results about the efficiency of the entire protocol. Conclusions are given in Sec. V

II. BACKGROUND

In brief, DFA operates as follows: an initial number N of users, also called *tags*, reply to a reader interrogation on a slotted time axis where slots are grouped into frames of length r_i , where i is the index of the frame; a tag is allowed to transmit only one packet per frame in a randomly chosen slot. In the first frame all tags transmit, but only a part of them avoid collisions with other transmissions and get through. The remaining number of tags n_i , often referred to as the *backlog*, re-transmit in the following frames until all of them succeed.

Although the standard allows stopping the exploration of the frame at any slot, and restart a new frame, in this paper we limit ourself to the basic version of DFA in which the Frame Restart is not allowed.

Outcomes of slots, i.e., successfully used, not used, or collided, are continuously observed to derive an estimate of the backlog, \hat{n}_i , which is used to set the length r_i of the next frame until all tags have been identified. The problem arises

to get at each frame a suitable estimate \hat{n}_i , and to determine the most favorable r_i .

From the protocol point of view, the goal is to minimize the average number of slots $L(N)$ needed to identify N tags, which depends on N and the way in which the length of the subsequent frames is assigned. In the following, we refer to the efficiency $N/L(N)$, a more convenient performance figure.

The optimal frame setting and performance of DSA when N is known have been thoroughly investigated in [15]. There we have shown that the optimum frame setting is $r_i = n_i$ at all frames, providing an efficiency that asymptotically reaches e^{-1} , an often recurring figure in S-Aloha.

When N is an unknown parameter, as usually assumed in DFA, three problems arise that affect the efficiency of the protocol, namely the backlog estimate \hat{n}_i , the optimal setting of frame lengths once \hat{n}_i is known, and the setting of the initial frame length r_0 , since at the beginning of the procedure no clues can be obtained on N . All the three issues are of paramount importance in determining the efficiency of Frame Aloha; however, up to now, none of them have been satisfactorily investigated together.

Many of works have appeared in the literature, [4]–[11], [14], deal with the first two problems. Results are rarely comparable because of different settings such as r_0 , or a restricted range of N , or limitations on the frame length r_i . No one has satisfactorily addressed the problem of conveniently setting r_0 , especially when N can vary from zero to a large unknown value, in which case r_0 can not be set to large values to avoid affecting the cases with low N . Again, no one has explicitly and satisfactorily addressed the case where in the first frames we observe all collisions, and how to produce the very first estimate of the population size. Furthermore, only few investigations have been produced on asymptotic performance [12], [13], and a systematic approach that could predict, for example, whether the protocol can reach e^{-1} efficiency, is missing.

Here we focus on a very large N and a small r_0 , that we set to $r_0 = 1$ since asymptotically its value has no impact on the performance. With such assumptions, in the first frame we observe all collided slots; since this observation brings no clues on the value of N , we must increase the length r_i of the subsequent frames with some predetermined law, until a frame appears where the slots are not all collided. At this point an estimate of N can be drawn, the frame length can be properly set, and the procedure enters the subsequent phase in the best possible conditions. Thus, the protocol can be seen as composed of two parts, the *convergence phase*, where an estimate of N is derived, and the *solution phase*, where tags are identified starting with a suitable frame length. If the estimate of N is perfect, as it happens with a consistent estimator when $N \rightarrow \infty$, in the solution phase we can set $r_i = n_i$, exactly as with known N , and tags are identified with efficiency e^{-1} . The overall efficiency is then dictated by how the lengths of the two phases compare.

With finite, but large N , the estimate \hat{N} derived in the convergence phase presents some approximation, and the

solution-phase efficiency can be improved by using an estimate drawn by subsequent observations. Then, we set $r_i = \hat{n}_i$, in accordance to the best setting when N is known [15], a good practice if \hat{n}_i presents a small estimation error and n_i is large, as it should be the case with an unbiased and consistent estimate when N is large. The two-phases modeling clearly shows that we are allowed two different estimators that can be optimized separately, an observation entirely new.

III. ANALYSIS OF THE CONVERGENCE PHASE

A. Asymptotic Behavior

In the following we assume that in the convergence phase the frame size is increased with a geometric law of parameter h , so that at step i the frame length is $r_i = \text{round}(h^i)$, for $i = 0, 1, \dots$. As r_i increases some non-collisions, i.e., empty or successful slots, begin to appear, say at frame U , and, at this point, an estimate \hat{N} can be provided. From here the *solution phase* begins, where, starting at frame $U + 1$, the estimate \hat{N} is continuously updated until all tags are identified.

Next we show that the observation drawn at frame U becomes more and more relevant as N increases, providing sufficient information for inferring the exact value N in the limit $N \rightarrow \infty$.

We start by evaluating the distribution of RV U , the first frame index at which a non-collision is observed. Let denote by X_i the number of non-collided slots in frame i . The probability that U coincides with frame i , $P_U(i) = \mathbb{P}(U = i)$, is given by

$$P_U(i) = \mathbb{P}(X_i > 0, X_{i-1} = 0, X_{i-2} = 0, \dots, X_0 = 0). \quad (1)$$

The frame length at frame i is $r_i = \text{round}(h^i) \approx h^i$ and the average traffic is $\gamma_i = N/\text{round}(h^i) \approx N/h^i$. In the approaching phase the traffic γ_i is very large, and successes and empties are rare events, allowing us to adopt Poisson as their distribution; therefore, (1) can be approximated to

$$P_U(i) \approx (1 - (1 - e^{-\gamma_i}(1 + \gamma_i))^{r_i}) \prod_{j=0}^{i-1} (1 - e^{-\gamma_j}(1 + \gamma_j))^{r_j}. \quad (2)$$

In the Appendix the following Lemma is proved:

Lemma 1. As $N \rightarrow \infty$ the function of γ

$$f(\gamma) = (1 - e^{-\gamma} - \gamma e^{-\gamma})^{N/\gamma} \quad (3)$$

tends to a unit step function centered in $\gamma = \ln N$, and such that $f(\ln N) = e^{-1}$.

Denoted by $\pi(\gamma)$ the function in the right-hand side (rhs) of (2) considered as function of $\gamma_i = \gamma$, we have

Lemma 2. As $N \rightarrow \infty$ function $\pi(\gamma)$ in (2) becomes

$$\pi(\gamma) = \begin{cases} 1 - e^{-1} & \gamma = \ln N/h \\ 1 & \ln N/h < \gamma < \ln N \\ e^{-1} & \gamma = \ln N \\ 0 & \text{elsewhere} \end{cases} \quad (4)$$

Proof. The proof is straightforward since, by Lemma 1, the first term in the rhs of (2) tends to a γ -reversed unit step centered in $\gamma = \ln N$, while terms of the remaining product can be written as $f(\gamma h^{i-j})$, each of them converging to a unit step centered in $\gamma = \ln N/h^{i-j}$, so that the product itself converges to a unit step centered in $\gamma = \ln N/h$. \square

We are now in the position to find the asymptotic behavior of P_U for large N . Starting from $r_0 = 1$ at frame $i = 0$ and increasing the frame size as h^i leads to a sequence $\{\gamma_i = N/h^i\}$, or $\{\gamma_i = h^{\nu-i}\}$, if we use the exponential notation $N = h^\nu$. Then we have the following theorem specifying the distribution of U , the number of the first not all collided frame:

Theorem 1. *As $h^\nu = N \rightarrow \infty$ the probability law P_U presents a unit mass in $U = u$ where u is the only index such that*

$$\nu - \ln(\beta\nu)/\beta < u < \nu - \ln(\beta\nu)/\beta + 1 \quad (5)$$

being $\beta = \ln h$.

Proof. The index frame u is such that γ_u lies within the support of $\pi(\gamma)$ in (4), that is,

$$\ln N/h < \gamma_u < \ln N, \quad (6)$$

or, using the exponential notation, and $\beta = \ln h$,

$$\nu\beta/h < h^{\nu-u} < \nu\beta.$$

Taking the logarithm we get

$$\ln(\beta\nu)/\beta - 1 < \nu - u < \ln(\beta\nu)/\beta, \quad (7)$$

which can be rewritten as in (5). Since the span of the range in (5) is less than one, while u must be an integer, there is only one u obeying (5). \square

A possible exception seems to be the case where γ_u coincides with the upper extreme $\ln N$. In this case N belongs to the set $\{N_u^*\}$ solutions of equation

$$N_u^* = r_u \ln N_u^*. \quad (8)$$

However, this scenario is not possible since r_u must be integer, while the solutions $\{N_u^*\}$ are irrational.

Figure 1 shows the values of u , as function of the population size exponent ν . This figure is purely qualitative, as it represents what happens in the limit $\nu \rightarrow \infty$. The value u increases as a staircase function, between the two envelopes that represent the bounds in (5): these bounds are parallel straight lines in the limit, thus showing that u increases at regular steps in ν .

Next we give an alternative interpretation of (5) by first fixing the value of N and then finding the boundaries on N that determine the change of frame index u .

Let us multiply the terms in (6) by r_u , thus getting the inequality

$$r_u \ln N/h < N < r_u \ln N.$$

Making N coincide with the lower and the upper extremes, and exploiting (8), shows that the boundaries of N can be expressed as in the following

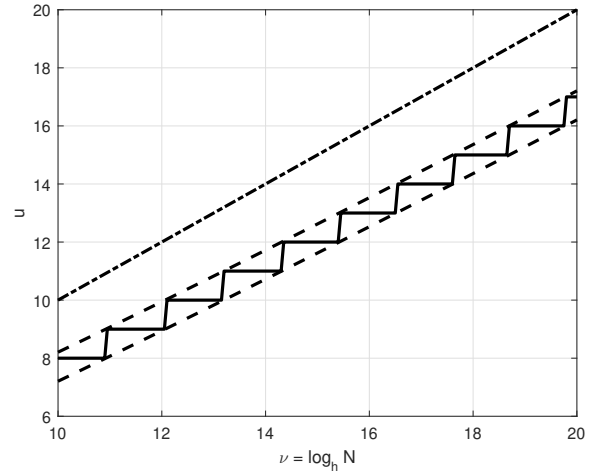


Figure 1. The staircase curve represents the values of u , assumed to be integer, as function of the population size exponent ν and $h = 2$, together with the dash curves representing their extremes and the straight line $\nu = u$.

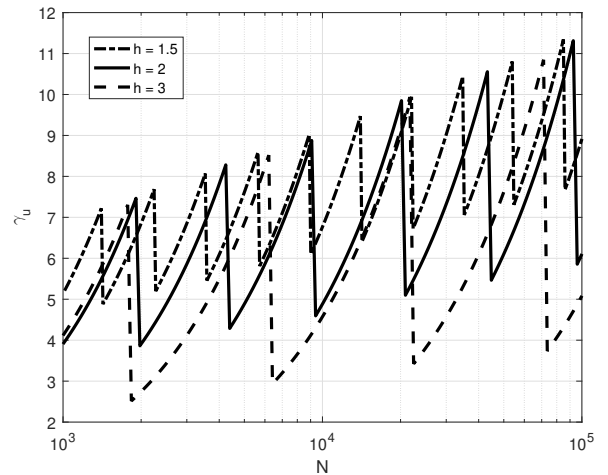


Figure 2. Behavior of the traffic γ versus the population size N , for different values of $h = \{1.5, 2, 3\}$.

Corollary 1. *The range of N that provides the same frame index u is given by*

$$N_{u-1}^* < N < N_u^*. \quad (9)$$

As a function of N , the steps of u in Fig. 1 occur at values $\{N_u^*\}$ defined in (9), or at values $\nu_u^* = \log_h N_u^*$. When N is within the interval $[N_{u-1}^*, N_u^*]$, the traffic γ_u increases linearly in N . When N is the closest integer to N_{u-1}^* the traffic is the closest value to $N_{u-1}^*/r_u = (\ln N_{u-1}^*)/h$; When N is the closest integer to N_u^* the traffic is the closest value to $N_u^*/r_u = \ln N_u^*$. When the upper extreme N_u^* is surpassed, the traffic suddenly drops to N_u^*/r_{u+1} , due to the increment in u , and then starts increasing again. Therefore we have

Corollary 2. *As N increases the behavior of traffic $\gamma_u(N)$ follows an increasing sawtooth behavior with extremes, in*

each period, given by

$$(\ln N_{u-1}^*)/h < \gamma_u < \ln N_u^*. \quad (10)$$

Figure 2 shows the behavior of traffic γ_u as function of the population size N for $h = \{1.5, 2, 3\}$. Discontinuities arise at values N_u^* . Again, the figure is purely qualitative, as it represents what happens in the limit $N \rightarrow \infty$. In any case, it shows that higher values of h cause larger oscillations in γ , denoting that low h are more precise in determining the range of γ , i.e., the range of N .

B. Finite N

When N is a finite number, the distribution of U changes with respect to what is described in Th. 1, and U can also assume values in the surroundings of u . Since we are dealing with large N , in the sequel we assume that the distribution of U takes values only in $\{u, u \pm 1\}$. Then, the step function $f(\gamma)$ becomes less sharp, and the function $\pi(\gamma)$ alters its form with respect to the rectangle (4) and becomes more similar to what is shown in Fig. 3; in particular, the breadth of $\pi(\gamma)$ exceeds one in logarithmic units, but it is still less than two. This means that, now, up to two elements of sequence $\{\gamma_i = N/h^i\}$, depending on the actual value of N , can lie in the support of $\pi(\gamma)$; one, the element of index u , lies within the rectangle, while the other, the element of index $u - 1$ (or $u + 1$), lies to the left (or to the right) of the rectangle. Their probabilities are, in fact, proportional to the height of the curve. When N is such that γ_u lies in the center of the rectangle, its probability is $P_U(u) \approx 1$; however, when N is close to the upper extreme N_u^* , and γ_u approaches the upper extreme of the rectangle, the probability of the term to the left with smaller traffic, which corresponds to $P_U(u + 1)$, i.e., a larger r , rises so that the distribution becomes bivariate. When N is close to the lower extreme N_{u-1}^* , and γ_u approaches the lower extreme of the rectangle, $P_U(u - 1)$ rises so that the distribution becomes bivariate again. Therefore, when N increases, the distribution of U changes periodically from a distribution with only one probability mass to a distribution with two probability masses.

The curve describing the average $E\{U\}$ versus N becomes a staircase function with smoothed steps, and the standard deviation of U oscillates periodically as shown in Fig. 4. Here, we observe that for $h = \{2, 3\}$, N is large enough to show values of N where U has zero variance. These values are those in the middle of range (9), $N \approx (N_u^* + N_{u-1}^*)/2$, which make γ_u lie at the center of rectangle (4) in Fig. 3, where the distribution of U has a unit mass; on the other side, when $N \approx N_u^*$ the variance is maximum because the distribution of U is bivariate. Focusing on $h = \{2, 3\}$, Fig. 3 also show that for values $N \approx N_u^*$, with traffic at the extremes of the rectangle, the bivariate distribution of N assumes values in u and $u \pm 1$ with probability ≈ 0.5 , which gives, in fact, standard deviation ≈ 0.5 . As N approaches infinite the bivariate distribution only appears in intervals around the N_u^* 's that shrink as N increases until they vanish in the limit. Therefore, the standard deviation is greater than zero in such shrinking intervals, but its peak is still ≈ 0.5 .

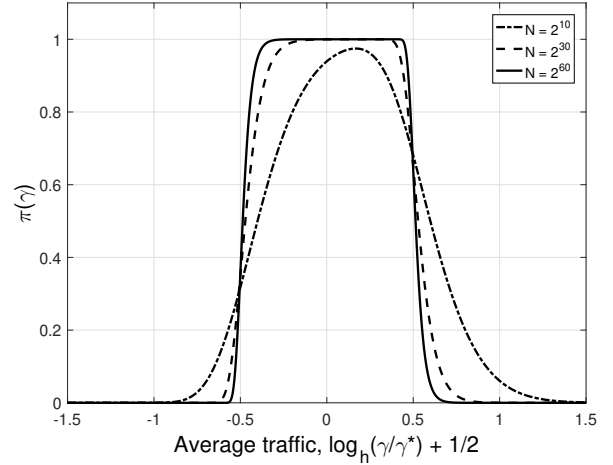


Figure 3. Shape of function $\pi(\gamma)$ of (2) for different values of the population size $N = h^\nu$ and $h = 2$.

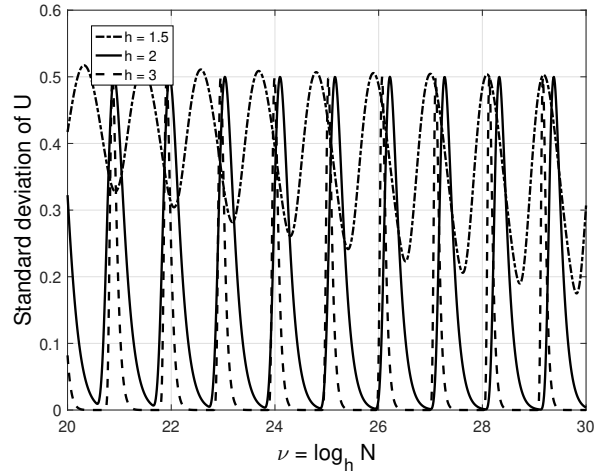


Figure 4. Standard deviation of index U for different values of the population size $N = h^\nu$ and $h = \{1.5, 2, 3\}$.

The traffic γ_u , shown in Fig. 2, at finite N becomes the RV γ_U whose average behavior changes to a smoother periodical shape, as shown in Fig. 5.

We now present some properties of the RVs represented by the number of successes S_i and empty slots E_i in frame $U = i$. In the Appendix we prove the following Lemma:

Lemma 3.

$$E\{S_i\} = \frac{N e^{-\gamma_i}}{1 - f(\gamma_i)}, \quad i \geq 0. \quad (11)$$

Furthermore, $E\{S_i\}$ is a decreasing function of N , and as $N \rightarrow \infty$ we have

$$1 < E\{S_u\} < (N_{u-1}^*)^{(1-1/h)} \quad (12)$$

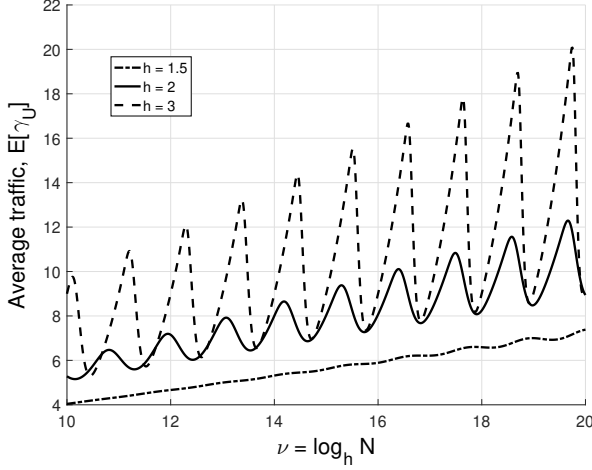


Figure 5. Behavior of the average traffic $E\{\Gamma\}$ versus the population size N , in the case $h = \{1.5, 2, 3\}$.

where the extremes are approached when N approaches the extremes N_u^* and N_{u-1}^* . The following also holds:

$$\text{Var}\{S_i\} \leq E\{S_i\}, \quad i \geq 0. \quad (13)$$

We note that, as N increases within the interval $[N_{i-1}^*, N_i^*]$ and γ_i increases, $E\{S_i\}$ decreases, presenting again a sawtooth behavior as N changes, the same as γ_i , though inverted.

Turning to the frequency of successes S_i/r_i , by Lemma 3, we have

$$E\left\{\frac{S_i}{r_i}\right\} = \frac{\gamma_i e^{-\gamma_i}}{1 - f(\gamma_i)}, \quad i \geq 0 \quad (14)$$

and:

Theorem 2. As $N \rightarrow \infty$ we have the following convergences in probability for $i \geq 0$

$$\frac{S_i}{r_i} \rightarrow \frac{\gamma_i e^{-\gamma_i}}{1 - f(\gamma_i)} \rightarrow 0, \quad \frac{E_i}{r_i} \rightarrow \frac{e^{-\gamma_i}}{1 - f(\gamma_i)} \rightarrow 0 \quad (15)$$

Proof: By Lemma 3 we have

$$\text{Var}\left\{\frac{S_i}{r_i}\right\} = \frac{\text{Var}\{S_i\}}{r_i^2} \leq \frac{E\{S_i\}}{r_i^2} \leq \frac{1}{r_i} \rightarrow 0 \quad (16)$$

which proves the convergence to (14). Convergence of E_i/r_i can be proved in a similar way. ■

C. Consistent Estimates

Here we derive an estimate \hat{N} of the population size N , based on the observations that refer to frame U , namely $U = i$, S_i and E_i . In order to simplify the notation, hereafter we omit the index i .

Theorem 3. For large N the ML estimate of the population size N , \hat{N}_{ML} , is the solution of the following equation

$$\frac{S + E}{r} = \frac{\hat{N}_{ML}}{r} e^{-\hat{N}_{ML}/r}, \quad (17)$$

and is consistent; that is, as $N \rightarrow \infty$ we have the following convergence in probability

$$\hat{N}_{ML} \rightarrow N \rightarrow \infty. \quad (18)$$

Furthermore, for finite N , \hat{N}_{ML} is biased, underestimating N .

Proof: The likelihood of N , or equivalently of the traffic $\gamma = N/r$, given the outcome in frame $U = i$ is

$$\mathcal{L}(\gamma|U=i, S=a, E=b) \propto \mathbb{P}(S=a, E=b, U=i|\gamma). \quad (19)$$

When N is very large S , E , and C , the latter representing the number of collided slots in frame i , follow a Poisson distribution of average respectively equal to rp_s , rp_e and rp_c , where $r = h^i$, $p_s = \gamma e^{-\gamma}$, $p_e = e^{-\gamma}$, and $p_c = 1 - e^{-\gamma} - \gamma e^{-\gamma}$. Using the shorthand notation

$$\Pi = \prod_{j=0}^{i-1} (1 - e^{-\gamma h^{i-j}} - \gamma h^{i-j} e^{-\gamma h^{i-j}})^{r/h^{i-j}}, \quad (20)$$

(19) reads as

$$\begin{aligned} \mathcal{L}(\gamma|i, a, b) &\propto \frac{(rp_s)^a}{a!} e^{-rp_s} \frac{(rp_e)^b}{b!} e^{-rp_e} \frac{(rp_c)^c}{c!} e^{-rp_c} \Pi \\ &= p_s^a p_e^b p_c^c \frac{r^r e^{-r}}{a! b! c!} \Pi \propto p_s^a p_e^b p_c^c \Pi \\ &= \gamma^a e^{-(a+b)\gamma} (1 - e^{-\gamma} - \gamma e^{-\gamma})^{r-a-b} \Pi, \end{aligned} \quad (21)$$

where $c = r - a - b$. Using the approximations

$$(1 - e^{-\gamma} - \gamma e^{-\gamma})^{r-a-b} \approx e^{-(r-a-b)(e^{-\gamma} + \gamma e^{-\gamma})}, \quad (22)$$

$$\begin{aligned} (1 - e^{-\gamma h^{i-j}} - \gamma h^{i-j} e^{-\gamma h^{i-j}})^{r/h^{i-j}} \\ \approx e^{-r/h^{i-j} (e^{-\gamma h^{i-j}} + \gamma h^{i-j} e^{-\gamma h^{i-j}})}, \end{aligned} \quad (23)$$

valid for $\gamma \gg 1$, the log-likelihood of γ is

$$\begin{aligned} \ln \mathcal{L}(\gamma|i, a, b) &\propto -\gamma(a+b) - (r-a-b)(e^{-\gamma} + \gamma e^{-\gamma}) \\ &+ a \ln \gamma - \sum_{j=0}^{i-1} \frac{r}{h^{i-j}} (e^{-\gamma h^{i-j}} + \gamma h^{i-j} e^{-\gamma h^{i-j}}). \end{aligned} \quad (24)$$

The derivative of the log-likelihood with respect to γ is

$$\begin{aligned} \frac{d}{d\gamma} \ln \mathcal{L}(\gamma) &\propto \frac{a}{\gamma} - (a+b) + (r-a-b)\gamma e^{-\gamma} \\ &+ \sum_{j=0}^{i-1} r h^{i-j} \gamma e^{-\gamma h^{i-j}}. \end{aligned}$$

The largest term of the summation is $rh\gamma e^{-\gamma h}$ which, for a large N that implies a large γ , becomes negligible for $h > 1$, so that we can disregard the whole summation. Furthermore, we can also disregard the first term a/γ and write

$$\frac{d}{d\gamma} \ln \mathcal{L}(\gamma) \propto - (a+b) + (r-a-b)\gamma e^{-\gamma}. \quad (25)$$

Equalling the above expression to zero yields the equation:

$$\frac{a+b}{r-a-b} = \frac{\hat{N}}{r} e^{-\hat{N}/r},$$

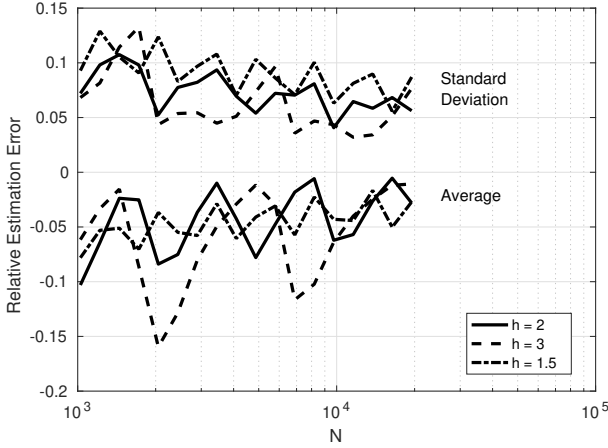


Figure 6. Behavior of the average and standard deviation of the fractional error of \hat{N} as N itself changes, for $h = \{1.5, 2, 3\}$.

which provides (17) because $r \gg a + b$.

Taking the average of (17) conditional to $U = i$, and disregarding E with respect to S , we get for all $i \geq 0$

$$\frac{N/r e^{-N/r}}{1 - f(N/r)} = \mathbb{E} \left\{ \frac{\hat{N}_{\text{ML}}}{r} e^{-\hat{N}_{\text{ML}}/r} \right\}. \quad (26)$$

The small value of the second derivative of the function $x e^{-x}$ for large x allows writing

$$\mathbb{E} \left\{ \hat{N}_{\text{ML}}/r e^{-\hat{N}_{\text{ML}}/r} \right\} \approx \mathbb{E} \left\{ \hat{N}_{\text{ML}}/r \right\} e^{\mathbb{E} \left\{ -\hat{N}_{\text{ML}}/r \right\}}, \quad (27)$$

so that from (26) we can derive the following inequality

$$N/r e^{-N/r} \leq \mathbb{E} \left\{ \hat{N}_{\text{ML}}/r \right\} e^{\mathbb{E} \left\{ -\hat{N}_{\text{ML}}/r \right\}}, \quad (28)$$

which implies $N \geq \hat{N}_{\text{ML}}$, i.e., that \hat{N}_{ML} underestimates N . However, if we let $N \rightarrow \infty$ in (17), by Th. 2 we can replace S/r with its asymptotic value $\gamma e^{-\gamma}$ and, still disregarding E/r , (17) converges (in probability) to

$$N/r e^{-N/r} = \hat{N}_{\text{ML}}/r e^{-\hat{N}_{\text{ML}}/r}, \quad (29)$$

that is, $\hat{N}_{\text{ML}} = N$, which proves (18). ■

D. Practical considerations

The results of Th. 3 show that for large, or infinite, values of N , \hat{N}_{ML} is virtually optimum; However, we must investigate its validity for moderately large values of N , such as $N = 1000$, more common in practice. Specifically, equation (17) might suggest that RV fluctuations in the LHS could cause too large fluctuations in the traffic estimate $\hat{\Gamma}_{\text{ML}}$ since the average values expected for $\gamma = N/r$ start from $\gamma \approx 4$ for $N \geq 1000$. This is not so, as exemplified in Fig. 6, that shows the average and the standard deviation of the fractional estimation error, i.e., $(\hat{N} - N)/N$, for different values of N , attained by simulating 10000 procedures for each N , and adopting $h = \{1.5, 2, 3\}$. The average value starts from about 10% at $N = 1000$ and decreases, as expected, as N increases. The

standard deviation presents a similar behavior. Figure 6 also confirms that \hat{N}_{ML} underestimates N .

The oscillations of average and standard deviations are similar to those already observed, e.g., in Fig. 5. The variance reaches the smallest values in correspondence of values $N \approx N_u^*$. In fact, as shown by Lemma 3, in those points S_u reaches its minimum variance, actually zero, so that the entire variance of \hat{N} is only due to r_U .

The effect of this error, and its bias, is negligible with respect to the performance of the whole protocol.

IV. FRAME ALOHA EFFICIENCY

A. Asymptotic Behavior

Given the estimate $\hat{N}_{U,\text{ML}}$ of the previous section, the estimate of the backlog at the beginning of frame $U + 1$ is given by $\hat{n}_{U+1} = \hat{N} - S_U$.

Therefore, disregarding S_U with respect to N , the average length of the solution phase is such that

$$\lim_{N \rightarrow \infty} \frac{L''}{N} = e. \quad (30)$$

As for the convergence phase, the number of slots up to frame i where the first non-collisions are observed, is given by

$$\begin{aligned} L' &= \sum_{i=0}^{\infty} P_U(i) \sum_{j=0}^i r_j = \sum_{i=0}^{\infty} P_U(i) \sum_{j=0}^i h^j \\ &= \sum_{i=0}^{\infty} P_U(i) \frac{h^{i+1} - 1}{h - 1} \approx \frac{h}{h - 1} \mathbb{E} \{ r_U \}. \end{aligned} \quad (31)$$

In the limit we have $\mathbb{E} \{ r_U \} \rightarrow r_u$, so that

$$\frac{L'}{N} \rightarrow \frac{h}{h - 1} \frac{r_u}{N} \rightarrow 0, \quad h > 1, \quad (32)$$

where the convergence to zero is proven by inequality (10). Therefore, in the limit, the overhead of the convergence phase length is zero and the overall efficiency,

$$\eta(N) = 1/(L'/N + L''/N), \quad (33)$$

becomes, by (30) and (32), e^{-1} .

B. Convergence Behavior

In Sec. III-D we have seen that, at the end of the convergence phase, the estimate presents an oscillating behavior, depending on the value of N , that affects the average length of phase itself and the overall efficiency.

As N increases, term L'/N (32) oscillates because r_u/N oscillates. This term decreases on the average as $1/\ln N$, and reaches zero asymptotically. To investigate term L''/N , we need more details.

After the $(U + 1)$ -th frame the backlog is

$$n_{U+2} = (N - S_U)(1 - e^{-(N - S_U)/r_{U+1}}). \quad (34)$$

Therefore, the average length of the solution phase is

$$\begin{aligned} L'' &= \mathbb{E} \{ r_{U+1} + e \cdot n_{U+2} \} \\ &= \mathbb{E} \left\{ \hat{N} - S_U + e(N - S_U)(1 - e^{-(N - S_U)/(\hat{N} - S_U)}) \right\}, \end{aligned}$$

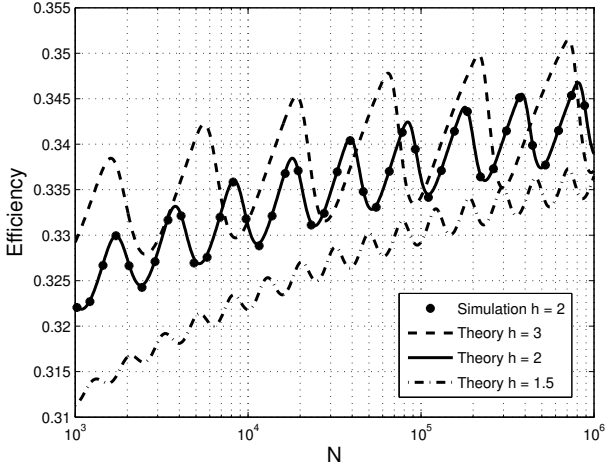


Figure 7. Protocol efficiency $\eta(N)$ versus the initial number of tags N for $h = \{1.5, 2, 3\}$. Simulated results are reported for $h = 2$.

which provides

$$\begin{aligned} \frac{L''}{N} &= \mathbb{E} \left\{ \frac{\hat{N}}{N} - \frac{S_U}{N} + e(1 - \frac{S_U}{N})(1 - e^{-\frac{N(1-S_U/N)}{\hat{N}(1-S_U/N)}}) \right\} \\ &\approx \mathbb{E} \left\{ \hat{N}/N - S_U/N + e(1 - S_U/N)(1 - e^{-1}) \right\}. \end{aligned} \quad (35)$$

Term (35) converges to e , as seen in (30), because of the consistency of \hat{N} and Theorem 2:

$$\text{plim}_{N \rightarrow \infty} \frac{S_U}{N} = 0. \quad (36)$$

In (35), both terms $\mathbb{E} \left\{ \frac{\hat{N}}{N} \right\}$ and $\mathbb{E} \{ S_U/N \}$ oscillates with the same period and phase as the first term, but with an amplitude by far smaller than the former, which decreases due to (18) and (36).

As an example, the plot of (33), derived using (31) and (35), for N in the range $10^3 < N < 10^6$ and $h = \{1.5, 2, 3\}$ is shown in Fig. 7. The oscillation of the efficiency is the one that affects almost all variables we have discussed, and must be ascribed to the limited resolution at which the convergence phase identifies N , the resolution decreasing as h increases. The oscillation can be reduced to any small value by taking h as close as possible to one, at the expenses of a slower convergence to e^{-1} , as shown by (32).

In practice we can not use the perfect backlog estimate in the solution phase. We could think of extending the MLE estimate in subsequent frames, taking into account the entire history. However, if N is not too small, such complex evaluation is not needed.

An interesting backlog estimate has been provided by Schoute in [1], being the first and the simplest ever appeared:

$$\hat{n} = \text{round}(2.39c), \quad (37)$$

where c is the number of collisions actually observed in the frame. It is derived assuming that in the observed frame the

traffic is in optimal conditions, i.e., $\gamma = 1$, i.e., $r_i = n_i$. The Schoute's proposal comes from approximating the number of tags transmitting in a slot by a Poisson variate. Hence, the average number of tags in a collided slot is $H = (1 - e^{-1})/(1 - 2e^{-1}) \approx 2.39$, and the estimate is then (37).

In [12], [13] we addressed the asymptotic efficiency of the Schoute's estimate starting with a mismatched frame length, actually $r_0 = 1$, and observed a convergence phase where the frame length is increased geometrically with $h = 2.39$. Once some non-collisions in frame U have occurred, Schoute's protocol proceeds again in the geometric increase phase, identifying some tags, until \hat{n} approaches n . We have analytically proved that this convergence indeed occurs, so that an efficient solution phase begins. However, this additional phase presents an average length that is not negligible with respect to the solution-phase length, and causes an overhead that reduces the asymptotic efficiency to 0.311.

In the same paper we have proved that the asymptotic efficiency of the Schoute's estimate increases to the theoretic maximum e^{-1} if the initial frame length r_0 happens to be equal to N . This shows the robustness of the estimate with respect to random fluctuations caused by collisions, and has suggested us to adopt Schoute's estimate in the solution phase, after the MLE estimate \hat{N} has been obtained. The simulation results for the case $h = 2$, averaged over 10^4 runs, are reported as dots in Fig. 7. The perfect match with the solid curve proves, in fact, the perfect suitability of this estimate. The pseudocode for one run of the Monte Carlo simulation is reported in the Algorithm.

Algorithm Dynamic Frame Aloha with efficient estimate

```

 $N \leftarrow h^{\nu_0}, L' \leftarrow 0, L'' \leftarrow 0$ 
 $i \leftarrow 0, r \leftarrow h^i, C \leftarrow r$ 
while  $C = r$  do
   $(C, S) \leftarrow \text{draw}(N, r)$ 
   $L' \leftarrow L' + r$ 
   $i \leftarrow i + 1$ 
   $r \leftarrow r \cdot h$ 
end while
 $n \leftarrow N - S$ 
 $\hat{\gamma} \leftarrow \text{solve}(\hat{\gamma}e^{-\hat{\gamma}} = (r - C)/r)$ 
 $r \leftarrow \hat{\gamma} \cdot r - S$ 
while  $C > 0$  do
   $(C, S) \leftarrow \text{draw}(n, r)$ 
   $L'' \leftarrow L'' + r$ 
   $n \leftarrow n - S$ 
   $r \leftarrow \text{round}(2.39 \cdot C)$ 
end while
 $L \leftarrow L' + L''$ 

```

V. CONCLUSIONS

In this paper we have analytically investigated the asymptotic behavior of dynamic frame Aloha for RFID when the population size N can range from small to very large and unknown values. Starting from an initial frame of length one,

we have shown that the entire procedure can be split into two parts. In the first, the *convergence phase*, the frame length increases geometrically, until some non-collisions are observed at frame U , where a consistent estimate of N is produced. In the second phase a simpler backlog estimate can be used. We have proved that, as $N \rightarrow \infty$ the overall efficiency reaches e^{-1} , the best possible. Also, we have provided a practical algorithm that achieves the best performance already for $N \approx 1000$. The convergence of the protocol efficiency to e^{-1} is rather slow, as goes like $1/\ln N$, and presents an oscillating behavior whose amplitude can be reduced at the expenses of the convergence speed.

APPENDIX

PROOF OF LEMMA 1

Since we are dealing with the case where $\gamma \gg 1$, we also have $e^{-\gamma} + \gamma e^{-\gamma} \ll 1$, so that we can use the approximation $1 - x \approx e^{-x}$ for $x \rightarrow 0$ and (3) becomes

$$f(\gamma) \approx e^{-(N/\gamma)e^{-\gamma} - Ne^{-\gamma}}. \quad (38)$$

Applying the change of variable $\gamma = t \ln N$, after some algebra we have

$$f(t \ln N) = g(t) \approx e^{\left(-\frac{1}{N^{t-1}} \left(\frac{1}{t \ln N} + 1\right)\right)}. \quad (39)$$

We now evaluate $g(t)$ in three points while letting $N \rightarrow \infty$, namely $t = 1$ and $t = 1 \pm \epsilon$, where $\epsilon > 0$. We get

$$g(1 - \epsilon) \rightarrow 0, \quad g(1) \approx e^{-1}, \quad g(1 + \epsilon) \rightarrow 1. \quad (40)$$

This shows that as $N \rightarrow \infty$ function $g(t)$ tends to an unit step centered in $t = t^* = 1$, therefore $f(\gamma)$ tends to an unit step centered in $\gamma = \ln N$.

PROOF OF LEMMA 3

We denote by \tilde{S} the binomial RV with parameters r and $p = \gamma e^{-\gamma}$, i.e., the approximate probability that in a slot there is a success. Also, \tilde{E} represents the binomial RV with parameters r and $p = e^{-\gamma}$, i.e., the approximate probability that the slot is empty. Since we are assuming large N , and $\gamma \gg 1$, so that $\gamma e^{-\gamma} \gg e^{-\gamma}$, in the following we neglect $E\{\tilde{E}\}$ with respect to $E\{\tilde{S}\}$.

For the expected value of RV S_i we have

$$\begin{aligned} E\{S_i\} &= E\left\{\tilde{S} \mid U = i\right\} = \frac{\sum_k k \mathbb{P}(\tilde{S} = k, U = i)}{\mathbb{P}(U = i)} \\ &= \frac{\sum_k k \mathbb{P}(\tilde{S} = k) \prod}{\mathbb{P}(\tilde{S} + \tilde{E} > 0) \prod} = \frac{E\{\tilde{S}\}}{\mathbb{P}[\tilde{S} + \tilde{E} > 0]}, \\ &= \frac{r\gamma_i e^{-\gamma_i}}{1 - f(\gamma_i)} = \frac{N e^{-\gamma_i}}{1 - f(\gamma_i)} \end{aligned} \quad (41)$$

where \prod is the product for $j = 0, \dots, i-1$ that appears in (2).

It is easy to show that the derivative of (41) with respect to N is always negative. For $i = u$ we asymptotically have

$1 - f(\gamma_u) \rightarrow 1$, and $E\{S_u\}$ in the two extremes of interval (10) becomes

$$\sup_N E\{S_u\} \rightarrow N_{u-1}^* e^{-(\ln N_{u-1}^*)/h} \rightarrow N_{u-1}^{*(1-1/h)} \quad (42)$$

$$\inf_N E\{S_u\} \rightarrow N_u^* e^{-\ln N_u^*} = 1. \quad (43)$$

In the same way we have derived (41) we can prove that

$$E\{S_i^2\} \approx \frac{E\{\tilde{S}^2\}}{\mathbb{P}[\tilde{S} > 0]}, \quad i \geq 0, \quad (44)$$

$$\begin{aligned} \text{Var}\{S_i\} &= \frac{E\{\tilde{S}^2\}}{\mathbb{P}[\tilde{S} + \tilde{E} > 0]} - \left(\frac{E\{\tilde{S}\}}{\mathbb{P}[\tilde{S} + \tilde{E} > 0]}\right)^2 \\ &= \frac{\text{Var}\{\tilde{S}\}}{\mathbb{P}[\tilde{S} + \tilde{E} > 0]} - \frac{E\{\tilde{S}^2\}}{\mathbb{P}[\tilde{S} + \tilde{E} > 0]} \left(\frac{1}{\mathbb{P}[\tilde{S} + \tilde{E} > 0]} - 1\right) \\ &\leq \frac{\text{Var}\{\tilde{S}\}}{\mathbb{P}[\tilde{S} + \tilde{E} > 0]} \leq \frac{E\{\tilde{S}\}}{\mathbb{P}[\tilde{S} + \tilde{E} > 0]} \approx E\{S_i\} \end{aligned} \quad (45)$$

for all $i \geq 0$, where we have exploited the fact that for the binomial distribution we have $\text{Var}\{\tilde{S}\} \leq E\{\tilde{S}\}$.

REFERENCES

- [1] F. Schoute, "Dynamic frame length Aloha," *IEEE Trans. Commun.*, vol. 31, no. 4, pp. 565 – 568, Apr. 1983.
- [2] *Information technology Radio frequency identification for item management Part 6: Parameters for air interface communications at 860 MHz to 960 MHz*, Intern. Organization for Standardization Std., 2004.
- [3] *Class 1 Generation 2 UHF Air Interface Protocol Standard Version 1.0.9*, EPCglobal Std., 2005.
- [4] H. Vogt, "Efficient object identification with passive RFID tags," in *Proc. First Intern. Conf. Pervasive Computing*, ser. Pervasive '02. London, UK: Springer-Verlag, 2002, pp. 98–113.
- [5] J.-R. Cha and J.-H. Kim, "Novel anti-collision algorithms for fast object identification in RFID system," in *11th Intern. Conf. on Parallel and Distributed Systems*, vol. 2, Jul. 2005, pp. 63 –67.
- [6] C. Floerkemeier, "Bayesian transmission strategy for framed Aloha based RFID protocols," in *IEEE Intern. Conf. on RFID*, Mar. 2007.
- [7] W.-T. Chen, "An accurate tag estimate method for improving the performance of an RFID anticollision algorithm based on dynamic frame length Aloha," *IEEE Trans. Autom. Sci. Eng.*, vol. 6, no. 1, Jan. 2009.
- [8] C.-F. Lin and F.-S. Lin, "Efficient estimation and collision-group-based anticollision algorithms for dynamic frame-slotted Aloha in RFID networks," *IEEE Trans. Autom. Sci. Eng.*, vol. 7, no. 4, Oct. 2010.
- [9] B. Knerr, M. Holzer, C. Angerer, and M. Rupp, "Slot-wise maximum likelihood estimation of the tag population size in FSA protocols," *IEEE Trans. Commun.*, vol. 58, no. 2, pp. 578–585, February 2010.
- [10] A. Zanella, "Estimating collision set size in framed slotted aloha wireless networks and rfid systems," *IEEE Communications Letters*, vol. 16, no. 3, pp. 300–303, March 2012.
- [11] D. Deng, C. Lin, T. Huang, and H. Yen, "On number of tags estimation in rfid systems," *IEEE Systems Journal*, vol. 11, no. 3, Sep. 2017.
- [12] L. Barletta, F. Borgonovo, and I. Filippini, "Asymptotic analysis of Schoute's estimate for dynamic frame Aloha," in *Int. Conf. Software, Telecom. and Computer Networks (SoftCOM)*, Sept 2015, pp. 113–118.
- [13] —, "Asymptotic analysis of backlog estimates for dynamic frame Aloha," *Journal of Commun. Software and Systems*, vol. 12, no. 1, pp. 83–90, 2016.
- [14] L. Zhu and T.-S. Yum, "Optimal framed Aloha based anti-collision algorithms for RFID systems," *IEEE Trans. Commun.*, vol. 58, no. 12, pp. 3583 –3592, Dec. 2010.
- [15] L. Barletta, F. Borgonovo, and M. Cesana, "A formal proof of the optimal frame setting for dynamic-frame Aloha with known population size," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 7221–7230, Nov 2014.