Analysis of Dynamic Frame Aloha with Frame Restart

Luca Barletta
Institute for Advanced Study
Technische Universität München
D-85748 Garching, Germany
luca.barletta@tum.de

Flaminio Borgonovo, Ilario Filippini
Dipartimento di Elettronica, Informatica e Bioingegneria
Politecnico di Milano
I-20133 Milano, Italy
borgonov@elet.polimi.it

Abstract—This paper provides an analysis of the Dynamic frame Aloha (DFA) protocol with frame Restart. Although a previous work has numerically provided the best restarting strategy when the number of tags N is known, that strategy is represented by a table whose entries are so many to be impractical for high values of N. Here we provide a simpler set of equations that, besides the optimal strategy, lead to a very simple sub-optimal strategy whose performance is practically indistinguishable from the optimal one. Furthermore, we prove that asymptotically the Frame Restart property is useless, being the corresponding efficiency equal to e^{-1}, the same as the plain DFA. We investigate the strategies and performance when N is unknown, furthermore providing a procedure, called AE^2, that is able to asymptotically reach the theoretical efficiency e^{-1}.

Index Terms—RFID, Collision Resolution, Frame Aloha, Frame Restart, Tag Estimation, Asymptotic Efficiency.

I. Introduction

Dynamic frame Aloha (DFA) is a collision resolution protocol originated in [1] and rediscovered some time ago for radio-frequency identification (RFID) [2], [3]. In this paper we will use terminology borrowed from RFID parlance, though the derived results are valid in general.

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; 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, often referred to as the *backlog*, re-transmit in the following frames until all of them succeed. Outcomes of slots, i.e., successfully used, not used, or collided, are continuously observed to derive an estimate of the backlog, $\hat{n}$, which is used to set the length r of the next frame till all tags have been identified. The problem arises to get at each frame a suitable estimate $\hat{n}$, and to determine the most favorable frame length r.

The protocol performance is measured by $L(N)$, the average number of slots needed to identify all tags. Since in the RFID environment the number N of tags to be identified is a constant, the performance can also be expressed by the protocol efficiency $N/L(N)$.

In the original version of DFA, that we refer to as *Frame Complete* (FC), a new frame is started only when the previous is finished. Many analyses have appeared on this protocol version, and its behavior is well understood (see for example [1],

[4]–[9]). Here we are concerned with the version of the cited RFID standards that allow for *Frame Restart* (FR), i.e., the capability to restart a new frame at any slot even though the present frame is not finished.

Although some papers have dealt with the performance of DFA-FR [7], [10], some basic issues are still unanswered. None of them has attempted an analysis, in order to determine the best frame setting and strategy, i.e., the rules that dictate when the frame exploration should be interrupted and the frame restarted, with the notable exception of [11]. A complete comparison with FC is lacking, in particular when N goes to infinity. Finally, none has thoroughly addressed the critical issue of the initial frame size, and the strategy when N is unlimited, a case not included in the current standard, but that could be considered in future applications.

In this paper we address the above cited issues about DFA-FR. On the analysis point of view we show that the performance $L(N)$, with N known, can be attained recursively by the performance of a system with $N - 1$ tags, solving a set of equations. In this way we are able to derive the optimal strategy and performance up to some finite value of N. Then we derive a very simple suboptimal strategy, which provides a performance practically coincident with the optimal, and that can be used for values of N where the exact analysis is intractable. We also show that as $N \to \infty$ the optimal performance of FR coincides with that of FC, i.e., $N/L(N) \to e^{-1}$. Then we turn to the case were N is unknown and analyze by simulation some estimation techniques, which shows a remarkable efficiency loss with respect the perfect estimation case. Finally, we address the asymptotic issue, proposing a procedure that is proved to asymptotically reach the theoretical efficiency e^{-1}. On the whole, our results seems to show that, with unknown N, the FR property does not offer appreciable advantages over FC.

The paper is organized as follows. In Sec. II, we discuss the performance of the most relevant backlog estimates appeared in literature. In Sec. III, we present an analysis of DFA assuming that both the backlog and the number of colliding packets are known. In Sec. IV, the analysis is extended to the case where the number of colliding packets is unknown. In Sec. V, we discuss and analyze some backlog estimation proposals, while in Sec. VI we present a procedure able to asymptotically reach the theoretical efficiency e^{-1}. Concluding remarks are given in Sec. VII.

II. PREVIOUS WORKS

To understand identification with FR some basic issues about FC are needed. Although many proposals have appeared in order to estimate the number of tags at the beginning of each frame with FC, the most appealing in our understanding is still the original one suggested in [1] by Schoute. Here, it has been shown that in optimal conditions, i.e., when the average number of transmissions per slot is one, the average number of collisions in a collided slot is 2.39. Therefore after a complete frame exploration, a good estimate of the backlog is provided by

$$\hat{n} = \text{round}(2.39c), \tag{1}$$

where c is the number of collided slots observed. This method is appealing because of its simplicity; it converges to the final identification for any initial frame length r_0, and it is also quite efficient. The performance of this protocol is illustrated in Fig. 1. Here, efficiencies are reported for different initial-frame lengths r_0, namely for the case $r_0 = N$ and for the case with known N, which represents a benchmark. In [8] we have proved that, with known N, the best strategy is to set the frame length equal to the backlog, i.e., $r = n$, and that the asymptotic efficiency reaches $e^{-1} \approx 0.367$.

When starting with $r_0 = N$, the figure shows that the algorithm provides a performance which is practically equal to the benchmark, reaching the theoretical asymptotic efficiency e^{-1}, as we have also analytically proved in [9]. However, in practical cases r_0 is set ignoring N. From the example we see a phenomenon, common to all identification procedures, caused by the mismatch between r_0 and N, which causes the efficiency to drop when N is not close to r_0. From the figure we also see that Schoute's algorithm can not reach the asymptotic efficiency e^{-1}. Actually, in [9] we have proved that Schoute's asymptotic efficiency is 0.311, due to the overhead of the convergence of the estimate from $\hat{N} = r_0$ to N.

With FR, after a frame restart, the observation of the outcomes in the frame is incomplete and, therefore, an estimation method such (1) can not guarantee the convergence. An estimation and restart strategy is represented by the Q-algorithm defined in ISO 18000-6 C [2]. It keeps a backlog estimate $\hat{n}$ at the beginning of the frame which is multiplied by a constant β whenever a collision occurs and is divided by β whenever an empty slot is detected. A successful slot leaves the estimate unchanged. Although the Q-Algorithm requires only modest computational resources, it does not specify a method to compute the crucial control parameter β, which is critical for the convergence.

The results we show in this section have been derived by simulating the standard algorithm, where the adopted value of $\beta = 2^\alpha$ is indicated in [7] as the one that provides a reasonable performance across different frame sizes. Since in the standard the frame size must be a power of two, the current frame length is calculated as $r = 2^Q$ where $Q = \text{round}(Q_{\text{fp}})$. Q_{fp} is updated at each slot according to the slot outcome 0, and > 1. If the outcome is a success the current frame size is reduced
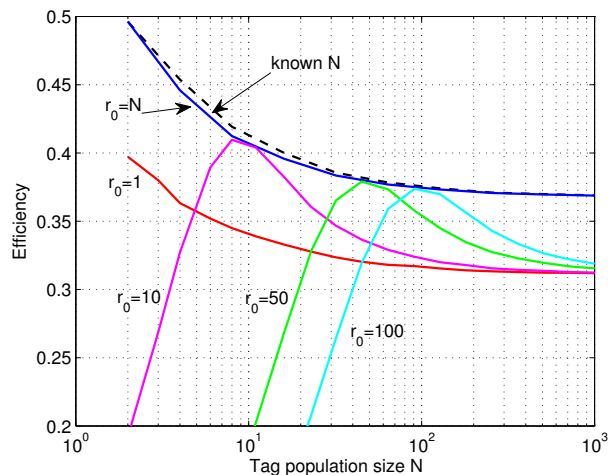

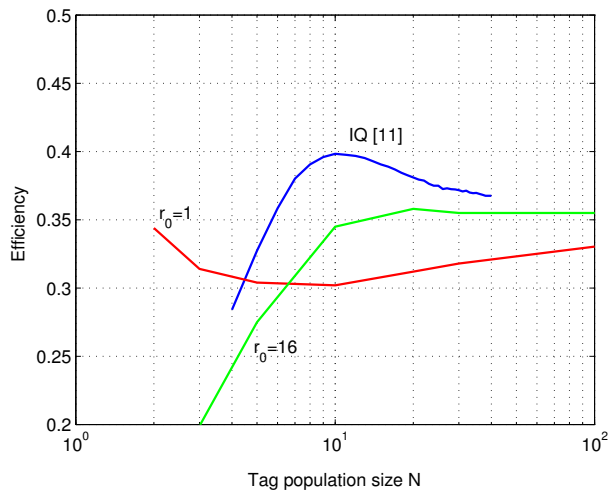Fig. 1. Efficiency of Schoute's protocol with different initial frame length r_0.

Fig. 2. Efficiency of the Q-algorithm with different initial frame length $r = 0$.

by one, otherwise the new size is computed using

$$\begin{aligned}
Q_{\text{fp}}^{(0)} &= \max\{0, Q_{\text{fp}} - \alpha\} \\
Q_{\text{fp}}^{(>1)} &= \min\{15, Q_{\text{fp}} + \alpha\} \\
\alpha &= \max\{0.1, \min\{0.5, \frac{0.8}{\log_2(r)}\}\}.
\end{aligned} \tag{2}$$

The frame is re-started each time the remaining length of the current frame does not match the calculated one. Efficiency results are shown in Fig. 2 for two initial frame sizes, namely $r_0 = 1$ and $r_0 = 16$. The performance is not encouraging, since it appears to be lower than in the FC case in Fig. 1, although the loss of about 0.02 can be ascribed to the constraint on the frame length that must be a power of two, as in the standard. Other losses are due to the error in the estimate of the backlog, and the initial frame size r_0, which causes an efficiency loss when N is not matched with r_0.
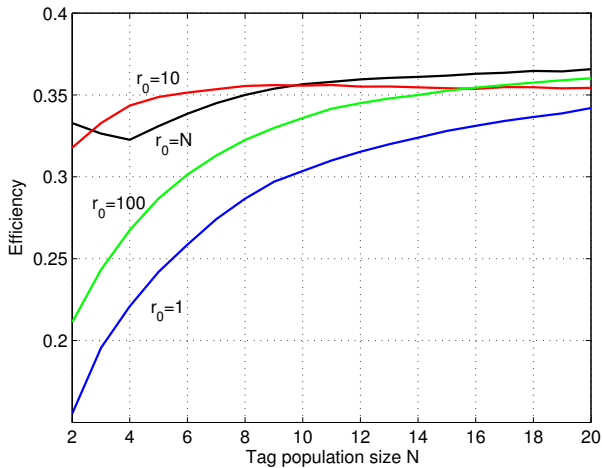
In [7] Floerkemeier uses a Bayesian method to derive

Fig. 3. Efficiency of Floerkemeier's protocol with different initial frame length r_0.

the distribution of the backlog n at each slot, based on the distribution at previous slot and on the observation of the outcome, i.e., E empty, S successful, and C collided. The distribution update equation is

$$p_n = P(n|z_{1:t}, o_{1:j}) = \alpha P(n|z_{1:t}, o_{1:j-1}) P(o_j|n, z_{1:t}, o_{1:j-1}) \tag{3}$$

where α denotes a normalizing constant, $o_{1:j}$ represents the outcome from the first j slots in the current frame, $z_{1:t}$ the outcomes of previous frames. Since consecutive frames are independent given the number of transmitting terminals, the following holds

$$P(o_j|n, z_{1:t}, o_{1:j-1}) = P(o_j|n, o_{1:j-1}). \tag{4}$$

Relation (3) provides a recurrent expression by which the "a posteriori" probability distribution of the original population size n, at slot j of frame t can be derived by the same expression at slot $j-1$ of frame t, starting from the "a priori" distribution of the number of transmitting terminals. The "a posteriori" probability p_n is then used to find the frame length r that maximizes the throughput of the next slot

$$E[S(r)] = \sum_n S(n,r)p_n = \sum_n \frac{n}{r}\left(1 - \frac{1}{r}\right)^{n-1} p_n. \tag{5}$$

This maximization is carried out slot by slot and a new frame is restarted whenever the frame length so determined changes with respect to the old one. The evaluation of (4) is rather complex and the method suggested in the original paper is not practical; therefore we have devised a computational method whose details have been given in [12].

In Fig. 3 we report the calculated performance of this procedure in three cases corresponding to three different initial frame size, namely $r_0 = 1, r_0 = 10,$ and $r_0 = 100,$ assuming an uniform population distribution in $[0; 1000]$. The case $r_0 = N$, the ideal one, is reported as a benchmark since N is not known in practice. The figure shows the effect of the

mismatch due to the wrong setting of the first frame, which is even increased when the range of the initial distribution is increased, say to 10000. In fact, with small r_0 and large N we very likely observe all collisions in the first slots of the frame, which causes the "a posteriori" distribution to concentrate on the maximum range values, probably far from optimal conditions. This behavior is of great concern when N has no limited value. The fact that the curve $r = 10$ is above the $r0 = N$ for $N < 10$ is due to the inaccuracy of the latter in estimating low values of N. Even in this case, where a powerful estimate procedure has been used, the FR procedure seems not to perform definitely better than in the FC case.

Finally, as we have already mentioned in the introduction, the analysis with known N proposed in [11] numerically provides the best strategy and the best frame setting up to $N = 40$. Here, the authors show that the optimal FR strategy only depends on the three variables n, c, k, where n is the number of unidentified tags (the backlog) at the present slot, c is the number of slots collided up to the present slot, and k is the number of slots remaining to the end of the frame. The analysis is carried out showing that the process defined over states (n, c, k) is a Markov Chain and the solution is attained by evaluating the first passage time to the final state where all tags are identified, which implies solving a very large system of equations. Numerical results are provided up to $N = 40$, and the optimal frame size is found to be slightly less than the backlog n.

Interesting as it is, the analysis carried out in the cited paper only provides numerical insights up to $N = 40$, because of its complexity, and, therefore, it is of little help in understanding the role of the parameters, and in determining the efficiency for $N > 40$ and $N \to \infty$. The optimal strategy is provided with a table with three parameters; it must be derived off line for each (n, r, c) and stored in the reader, which can be expensive, if not impossible, for very large n. Again, the analysis does not suggests a simpler and efficient alternative strategy.

In the same paper the authors propose that, when n is unknown, an estimate $\hat{n}$ should be used instead. As an example, they use as estimate a slight modification of the one adopted in the Q-algorithm explained above. This estimate is used to set the frame length and to decide the frame restart according to the optimal strategy. The resulting protocol has been referred to as Improved Q-algorithm (IQ). Unfortunately, the authors do not provide the details about the estimation parameters, so that results can not be re-derived; therefore they have been graphically reported in Fig. 2. The performance of IQ is the best among the FR protocols cited, although, again, it seems not definitely better than in the FC case.

III. ANALYSIS WITH KNOWN COLLISIONS

In order to understand the reasons that take advantage of the frame restarting mechanism, we assume that we can exactly track the number $m(k)$ of tags yet to be transmitted in the next k slots of the frame, which implies the assumption of knowing the number of the tags colliding in a slot. This situation can be approached in practice by mechanisms such as the Bayesian

TABLE I
EFFICIENCY AND OPTIMAL SETTING OF DFA PROTOCOLS WITH FRAME RESTART WHEN COLLISIONS ARE KNOWN.

n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
FR	0.6666	0.6207	0.5805	0.5523	0.5311	0.5146	0.5014	0.4907	0.4818	0.4743	0.4678	0.4621	0.4571	0.4527	0.4487	0.4452	0.4420	0.4390	0.4364
r*	2	2	3	4	5	6	7	7	8	9	10	11	12	13	14	15	16	17	18
FR PC	0.0000	0.0000	0.0000	-0.0021	-0.0033	-0.0042	-0.0047	-0.0047	-0.0049	-0.0051	-0.0054	-0.0056	-0.0058	-0.0059	-0.0060	-0.0061	-0.0061	-0.0061	-0.0062
r*	2	2	3	4	5	6	6	7	8	9	10	11	12	13	14	15	16	17	18
FR PT	0.0000	0.0000	0.0000	0.0000	-0.0005	-0.0005	-0.0007	-0.0007	-0.0008	-0.0010	-0.0011	-0.0011	-0.0011	-0.0011	-0.0011	-0.0012	-0.0012	-0.0011	-0.0012
r*	2	3	4	5	6	6	7	8	9	10	11	12	13	14	15	16	17	18	19
FC	0.5000	0.4706	0.4533	0.4414	0.4327	0.4261	0.4208	0.4165	0.4129	0.4098	0.4072	0.4050	0.4030	0.4012	0.3996	0.3982	0.3969	0.3957	0.3947

TABLE II
STRATEGY MATRIX AND GAIN $L_n(m,k) - L_n(n,r)$
FOR FR WITH $n = 10$ AND $r = 8$.

m/k	1	2	3	4	5	6	7
1	-1.41	-0.91	-0.41	0	0	0	0
2	0	-0.90	-0.87	-0.54	-0.12	0	0
3	0	-0.44	-0.79	-0.76	-0.50	-0.16	0
4	0	0	-0.47	-0.69	-0.64	-0.42	-0.13
5	0	0	-0.12	-0.45	-0.58	-0.52	-0.32
6	0	0	0	-0.17	-0.39	-0.46	-0.39
7	0	0	0	0	-0.16	-0.31	-0.34
8	0	0	0	0	0	-0.12	-0.21
9	0	0	0	0	0	0	-0.06
10	0	0	0	0	0	0	0

mechanism in [7]. Anyway, here the knowledge of $m(k)$ is a working hypothesis.

Let $L_n(m,k)$ be the average identification time starting from slot k to frame's end, when at the same slot the backlog is n, and let r be the initial length of the frame. The advantage of FR is that at each slot we can either decide to proceed with the frame exploration, in that case the performance is $L_n(m,k)$, or to restart a completely new frame, where the average identification time is given by $L_n(n,r)$. The best strategy at each slot is, of course, the one that leads to the smallest L_n. Once the best strategy is selected, an optimal choice of r can be determined at each new frame.

Since at the beginning of a new frame the past history does not influence the performance other than in the variables n and r, once the best strategy is determined for an initial number of tags N, $N \leq n-1$, then the performance with $N = n$ only depends on the occurrences in the first or subsequent frames up to the first success. This leads us to the following set of equations

$$L_n(n,r) = 1 + p_1(n,r)L^*(n-1) + \sum_{i \neq 1} p_i(n,r)L_n(n-i,r-1) \tag{6}$$

$$L'_n(m,k) = 1 + p_1(m,k)L^*_{n-1}(m-1,k-1) + \sum_{i \neq 1} p_i(m,k)L_n(m-i,k-1) \tag{7}$$

$$L_n(m,k) = \min\{L_n(n,r), L'_n(m,k)\} \tag{8}$$

where we have denoted with an asterisk the performance of the best strategy and with $p_i(m,k)$ the probability that i out of m tags choose the first slot in a frame of k slots, namely the Binomial distribution

$$p_i(m,k) = \binom{m}{i}\left(\frac{1}{k}\right)^i\left(1 - \frac{1}{k}\right)^{m-i}. \tag{9}$$

Once the best strategy is known with $N = n - 1$, i.e., $L^*(n-1)$ is known, the system in (6)-(8) can be solved for $L_n(m,k)$ and we can find the value $r = r_0$ that minimizes $L_n(n,r)$, thus determining the optimal strategy and $L^*(n) = L_n(n,r_0)$. Then, the procedure can be repeated iteratively for any value of n starting from $L^*(1) = 1$.

Values $L_n(m,k)$ in (6)-(8) can be easily derived for any k and any m starting from $L_n(m,1)$. Since those values can be expressed as

$$L_n(m,1) = L_n(n,r), \qquad m > 1 \tag{10}$$

$$L_n(1,1) = 1 + L^*(n-1), \tag{11}$$

for $m > 1$, the values can be attained iteratively starting from the unknown $L_n(n,r)$. Table I shows the results in term of efficiency, $n/L^*(n)$, up to $n = 20$, together with the optimal setting $r_0 = r^*$ of the frame length. The performance of DFA-FC, is also reported for comparison, while Table II shows the strategy matrix for $n = 10$ and $r = 8$. In this table is reported the gain $L_n(m,k) - L_n(n,r)$, so that when the gains are negative the frame is further explored. We note that function $L_n(m,k) - L_n(n,r)$ can be seen as a plane of height 0 in which a valley is carved along the diagonal axis that from (n,r) leads to state $(1,1)$. The bottom of the valley decreases as k decreases, reaching its minimum in $(1,1)$.

A suboptimal FR protocol, that performs not far from the optimal, is the one that restarts the frame whenever a success occurs, which means replacing $L^*_{n-1}(m-1,k-1)$ in (7) with L^*_{n-1}. We now look for the best strategy of restarting the frame if either empty slot or a collision occurs. Denoted by $l(n,r)$ the average number of slots to the next success, the equations of this protocol can be derived from (6)-(8) as

$$L_n(n,r) = l(n,r) + L^*_{n-1} \tag{12}$$

$$L_n(m,k) = l(m,k) + L^*_{n-1} \tag{13}$$

and the optimal strategy is the one that re-starts the frame if $l(n,r) < l(m,k)$. Making explicit only variable k, we have

$$l_k = 1 + \beta_k l_{k+1} = 1 + \beta_k(1 + \bar{\beta}_{k-1} + (1 + \bar{\beta}_{k-2} + (1 + \bar{\beta}_{k-3} + (...(1 + \bar{\beta}_1 l_n)))). \tag{14}$$

Here β_k is the probability that a success does not occur in slot k, while subsequent $\bar{\beta}_{k-i}$ represent the probability that a

TABLE III

EFFICIENCY AND OPTIMAL SETTING OF DFA PROTOCOLS WITH FRAME RESTART WHEN COLLISIONS ARE UNKNOWN.

n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
FR	0.6666	0.6000	0.5549	0.5247	0.5036	0.4877	0.4754	0.4655	0.4574	0.4506	0.4449	0.4400	0.4357	0.4319	0.4286	0.4256	0.4230	0.4206	0.4184
r*	2	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
FR PC	0.0000	0.0000	0.0000	-0.0008	-0.0009	-0.0008	-0.0007	-0.0008	-0.0009	-0.0009	-0.0010	-0.0010	-0.0010	-0.0010	-0.0010	-0.0010	-0.0010	-0.0010	-0.0010
r*	2	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
FR PT	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0004	-0.0004	-0.0004	-0.0004	-0.0004	-0.0004	-0.0004	-0.0003	-0.0004	-0.0003	-0.0004	-0.0004	-0.0003
r*	2	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
FC	0.5000	0.4706	0.4533	0.4414	0.4327	0.4261	0.4208	0.4165	0.4129	0.4098	0.4072	0.4050	0.4030	0.4012	0.3996	0.3982	0.3969	0.3957	0.3947

success does not occur in slot $k-i$ averaged over the different outcomes up to $k-i$.

$$l_n - l_k = (\beta_n - \beta_k) + (\beta_n \bar{\beta}_{n-1} - \beta_k \bar{\beta}_{k-1}) + \\ + (\beta_n \bar{\beta}_{n-1} \bar{\beta}_{n-2} - \beta_k \bar{\beta}_{k-1} \bar{\beta}_{k-2}) + \dots . \quad (15)$$

The frame should be restarted whenever $l_n - l_k < 0$. However, degree of approximation to this sub-optimum can be attained looking only at some terms of the summation on the right. Stopping at first term, the frame is restarted when $\beta_n - \beta_k < 0$, i.e, when the success probability is increased restarting the frame. This provides the sub-optimal strategy that we call *probability comparison*, (PC), whose results are shown in Table I, where the performance of this suboptimum appears very close to the optimal one. In the same Table we also show the performance of the strategy labeled *probability threshold*, (PT), which restarts the frame when $\beta_k > (1 - p_0)$, i.e., probability of success in the next slot $P_s > p_0$, being p_0 a threshold to be optimized. We see that this suboptimum with $p_0 = 0.33$ behaves still better than PC, and practically coincides with the optimum.

As we can see by comparison with FC, the FR property leads to a gain in the efficiency, which is remarkable for $n = 2$ and diminishes as n increases. For $n \to \infty$ this gain reduces to zero, as shown in the following

Theorem 1. *The asymptotic optimal frame setting is $r = n$ and with such setting we have*

$$\lim_{n \to \infty} \frac{L^*(n)}{n} = e. \quad (16)$$

Proof: Let consider $n \to \infty$, and assume that the optimal setting of r is $r(n)$. Denoting with $s(n) = n/r(n)$ the traffic, we exclude that at optimal setting the traffic is zero or infinite. Therefore, in the following we assume that n and r are related such that $n/r = s$, with $0 < s < \infty$. This means that

$$L_n(n,r) = l_n(s)n + o(n), \qquad n \to \infty. \quad (17)$$

Referring to any finite index i, we have also

$$L_n(n - i, r - 1) = l_n(s)n + o(n), \qquad n \to \infty. \quad (18)$$

Furthermore, $\{p_i(n,r)\}$, i.e., the distribution of the number of tags transmitting in the first slot of the frame, for $n \to \infty$ tends to the Poisson distribution with average $s = n/r$,

$$p_i(n,r) = \frac{s^i}{i!} e^{-s} = p_i(s). \quad (19)$$

With finite traffic s, for large values of i, i.e., proportional to the backlog $i = kn$ for $0 < k \leq 1$, the term in the summation in (6) becomes

$$p_{kn}(s) \cdot (l_n(s)n + o(n)) = \frac{s^{kn}}{(kn)!} e^{-s} (l_n(s)n + o(n)) \quad (20)$$

$$= \left(\frac{se}{kn}\right)^{kn} \frac{e^{-s}}{c_{kn} \sqrt{kn}} (l_n(s)n + o(n)) \quad (21)$$

$$= o\left(\left(\frac{se}{kn}\right)^{kn} \sqrt{n}\right), \quad (22)$$

where (21) comes from Stirling's bound, with $\sqrt{2\pi} \leq c_{kn} \leq e$ for any kn. From (22) we see that the terms with $i = kn$ are negligible, therefore the summation in (6) becomes

$$\sum_{i \neq 1} p_i(n,r) L_n(n-i, r-1) = (l_n(s)n + o(n))(1 - p_1(s)) \quad (23)$$

and, from (6) we have for $n \to \infty$

$$L_n(s) = l_n(s)n = \frac{1}{p_1(s)} + L^*(n - 1). \quad (24)$$

The above is minimized by maximizing $p_1(s)$, i.e., for $s = 1$ or $r = n$, which in the limit yields

$$L^*(n) = e + L^*(n - 1) \quad (25)$$

which is exactly the asymptotic performance of DFA-FC [9]. ■

IV. ANALYSIS WITH UNKNOWN COLLISIONS

The result of the previous section may be too optimistic when an estimate of $m(k)$ is used. Here we use the less stringent hypothesis that assumes we know the backlog n at the beginning of the frame.

The new state, with backlog n and frame r, becomes (c, k), where c is the number of collided slots up to the current slot, anti-correlated with m, and k is as before. We can readily write the equations that correspond to (6)-(8)

$$L_{n,r}(0,r) = 1 + p_1(n,r) L^*(n-1) \\ + p_0(n,r) L_{n,r}(0, r-1) \\ + p_c(n,r) L_{n,r}(1, r-1) \quad (26)$$

$$L'_{n,r}(c,k) = 1 + p_{n,r}^{(1)}(c,k) L^*_{n-1}(c, k-1) \\ + p_{n,r}^{(0)}(c,k) L_{n,r}(c, k-1) \\ + p_{n,r}^{(c)}(c,k) L_{n,r}(c+1, k-1) \quad (27)$$

$$L_{n,r}(c,k) = \min\{L_{n,r}(0,r), L'_{n,r}(c,k)\} \qquad (28)$$

where we have denoted with an asterisk the performance of the best strategy, $p_i(n,r)$ is the binomial (9), with $p_c = \sum_{i>1} p_i$, $p_{n,r}^{(i)}(c,k)$ is the probability that i backlogged tags transmit when state is (c,k), and $p_{n,r}^{(c)} = \sum_{i>1} p_{n,r}^{(i)}$. A closed-form expression of probability $p_{n,r}^{(i)}(c,k)$ can be found in [13].

The equations above can be solved in the same way as (6)-(8), with initial conditions

$$L_{n,r}(c_{\max}, k) = L_{n,r}(0,r), \qquad (29)$$
$$c_{\max} = n/2, \qquad n \text{ even}$$
$$c_{\max} = (n-1)/2, \qquad n \text{ odd}$$

Table III shows the results in the analogous of Table I. In the comparison we see that the loss of knowledge costs to the efficiency about two decimal points starting from $n=3$ up to $n=20$. Even in this case we have evaluated the two suboptimal policies PC and PT of the previous section. Here the best results are again attained with policy PT if the frame is restarted whenever $P_s > 0.37$. Again, policy PT practically coincides with the optimum.

As for the asymptotic performance we still have:

Theorem 2. *The asymptotic optimal frame setting is $r = n$ and with such setting we have*

$$\lim_{n \to \infty} \frac{L^*(n)}{n} = e. \qquad (30)$$

The proof comes immediately by observing that the performance must be no better than the previous case and no worse than with the entire exploration of the frame.

As already mentioned in Section II the results above have already been derived in [11], though the Markovian framework there adopted appears rather more complicated than equations (6)-(8).

V. Unknown backlog

In practical cases N, and the backlog at each frame, are unknown, and the analysis carried out in the previous sections does not apply. If N is known in probability, i.e., is regarded as a random variable, an optimal strategy based on the outcomes observation is still possible. In [12] we have considered the case where N is Poisson distributed, with small average $E[N] \leq 2$. In that case we have been able to attain the best strategy by exhaustive search.

However, if N is large exhaustive search is practically impossible. Furthermore, N can hardly be regarded as a random variable, and must be considered as a constant of possibly infinite range. Therefore, the only way, as suggested in [11], seems to derive an accurate estimate of the backlog n at the beginning of the frame and then using the FR optimal strategy theoretically derived for known N.

In Fig. 4 we have reported the efficiency curve IQ, taken from [11]. For comparison purposes we have also reported the theoretical efficiency with known N, derived mathematically up to $N = 50$ as seen in the previous section, and that we

know asymptotically reaching e^{-1}, as stated in Theorem 1. We have also reported the efficiency limit curve attained with Schoute's protocol in the FC case with $r_0 = N$, which also reaches e^{-1} in the limit. We can see that, at least for moderate values of N, not only the efficiency loss of the proposed protocol with respect to the known-N case is remarkable, but it also presents losses with respect to the FC limit. This must in part be ascribed to the mismatch between r_0 and N, but in greater part to the imprecision of the Q-algorithm in estimating the backlog in this range of N.

We have not been able to replicate the curve IQ, because of lack of details. Therefore, we have devised a procedure able to assess the validity of the sub-optimal PT procedure of Section IV. At each slot of frame i the estimate $\hat{n}_i$ at the beginning of the frame is updated with the Q-algorithm algorithm illustrated in Sec. II and parameters in (2). This estimate is used to evaluate the average number of collided tags in a collided slot, h, assuming the traffic in the slot is Poisson distributed with average $\hat{\gamma} = \hat{n}_i/r_i$:

$$h = \hat{\gamma} \frac{1 - e^{-\hat{\gamma}}}{1 - e^{-\hat{\gamma}} - \hat{\gamma} e^{-\hat{\gamma}}}.$$

Then the number of tags m that remain to be transmitted in the subsequent k slots of the frame is evaluated as

$$\hat{m} = \hat{n} - hc - s,$$

where c and s are respectively the number of collided and successful slots of the frame up to the current slot. The success probability P_s in the next slot is evaluated by the binomial (9) and the frame is restarted whenever it is $P_s < p_0$, where p_0 is a prefixed threshold. The protocol is started by setting $\hat{n} = \hat{N} = r_0$. This procedure is labeled Q1 and its best performance is reported in Fig. 5, which is an enlargement part of Fig. 4, with parameters $r_0 = 16$ and $p_0 = 0.37$. As we can see, the performance of this procedure has a shape close to IQ, although somewhat lower. Changing to $r0 = 8$ provides a higher peak, of about 0.42, but with narrower shape and shifted to the left. The steep descent beyond the maximum must be probably ascribed to an imprecise estimate when N is mismatched with respect r_0.

To avoid the above effect, we have reverted to the Bayesian estimate, as explained in Section II, the FR strategy remaining the same. Its efficiency is shown in the figure with label Bayesian, where the best performance has been attained with parameters $r_0 = 8$, and $p_0 = 0.37$. To avoid the negative effects on the "a posteriori" distribution already mentioned in Sec. II, we initially restart on the first slot as long as a collision is observed by doubling the frame length, turning to the standard procedure when a non-collision is observed at the beginning of the frame. The efficiency peak is lower but the descent is more favorable, as expected. In fact, we expect that as $N \to \infty$, the Bayesian estimate becomes exact, and the efficiency reaches the Schoute's limit e^{-1}.

In trying to raise the peak we have introduced an initial phase that uses the Q-algorithm with the above parameters, switching to the Bayesian when $P_s < p_0$ and at least one
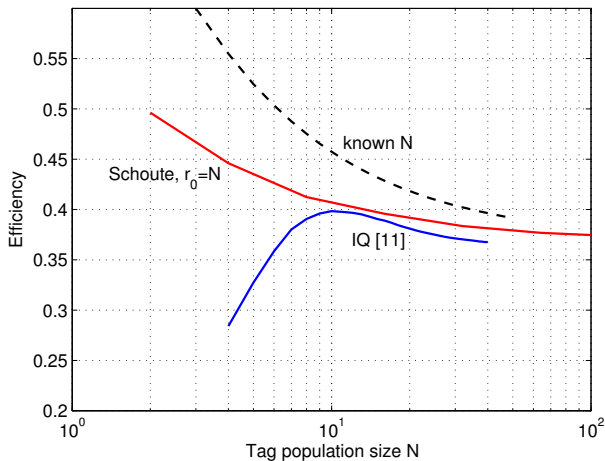
Fig. 4. Efficiency of the IQ protocol compared with the case of precise estimation, known N, and with the FC with $r_0 = N$.
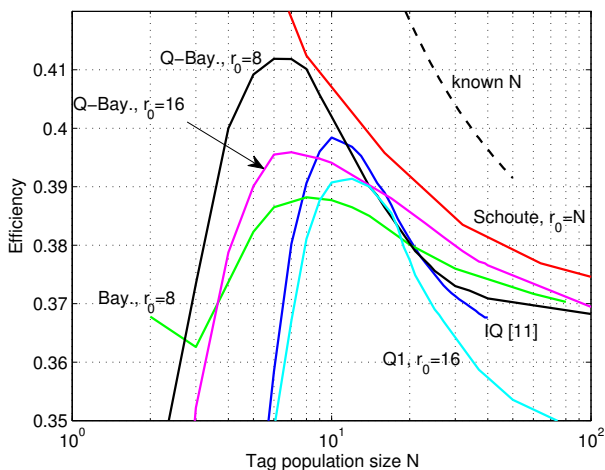

Fig. 5. Efficiencies of different FR protocols.

success has been observed up to the present slot. In the figure we report the curves, labeled as Q-Bayes, obtained with parameters $r_0 = 8$ and $r_0 = 16$ with $p_0 = 0.37$. The former is always better than IQ while the latter presents the best performance for high N, due to the good estimate of the Bayesian method. All these figures are congruent with the IQ and shows the optimal performance of the PT strategy.

In concluding this section, we can not avoid to note that, unlike FC, practical implementations of the FR strategy are very far from the ideal case where N is known.

VI. ASYMPTOTIC BEHAVIOR

We have already mentioned that the Bayesian estimate is expected to converge to n as $n \to \infty$ and to efficiency e^{-1}. However, for large n the Bayesian estimate is not practical. On the other side the simpler Q-algorithm is not proven to converge. Fortunately, a simple and converging protocol

exists. In [9] we have analytically proved that a simple procedure, called *Asymptotically Efficient Estimate* (AE^2), based on Schoute's mechanism, is able to reach the asymptotic e^{-1} limit. This procedure can also be integrated with the mechanisms shown in the previous section and used when $\hat{n}$ becomes large.

AE^2 works as follows. The backlog estimate is updated at the end of frame i, either complete or not, as follows

$$
\begin{aligned}
\hat{n}_{i+1} &= \text{round}\left(H(\hat{n}_i)c_i \frac{r_i}{l_i} \right), \qquad c_i > 0, \\
\hat{n}_{i+1} &= \hat{n}_i - s_i, \qquad c_i = 0, \\
r_{i+1} &= \hat{n}_{i+1}.
\end{aligned}
\tag{31}
$$

where c_i and s_i are respectively the number of collided and successful slots observed in frame i, r_i is the frame length, l_i is the number of observed slots when a new frame is restarted, and

$$
H(\hat{n}_i) = \frac{1 - (\hat{n}_i/r_i)e^{-1}}{1 - 2e^{-1}}.
\tag{32}
$$

As for the increase in l_i, we asymptotically use the law

$$
l_i = \min\{\text{round}((i+1)^b), r_i\},
\tag{33}
$$

with $b > 0$. This setting has been proved sufficient to let the estimate converge asymptotically to the true n. We explicitly note that the above strategy completely replaces the optimal strategies considered up to now.

Briefly, the meaning of the above algorithm is as follows. Law (31) is used to predict the backlog based on the observed collided slot c_i, which is resized to the frame length r_i. The multiplier $H(\hat{n}_i)$ takes the role of the constant 2.39 in (1). When some convergence is achieved and $\hat{n}_i \approx r_i$ we get, in fact, $H(\hat{n}_i) = 2.39$. At the beginning of the convergence phase, where all collisions are observed, $\hat{n}_i$ grows very rapidly, with geometric law, but the frame length (33) increases much more slowly, with a quadratic law if we use $b = 2$, and this assures that the number slots $C(N)$ spent in the convergence phase are of lower order than N, so that $C(N)/N \to 0$ as $N \to \infty$. In the end $\hat{n}_i$ stabilizes, round$((i+1)^b)$ in (33) reaches $\hat{n}_i$ and we get $l_i = r_i$ and $H(\hat{n}_i) = 2.39$. At this point, and for large N we have $\hat{n} \approx n$, and the procedure proceeds exactly as in Schoute with $r_0 = N$ (the convergence phase has negligible overhead), thus reaching efficiency e^{-1}.

In Fig. 6 we have reported both analytical and simulated values (dots) attained averaging 10^4 simulation samples, in the case $N = 10^3$ and $r_0 = 1$. The red dash-dotted line represents the relative error multiplied by 10^3, still very small. For comparison purposes we have also reported the curve referring to Schoute's algorithm. We clearly see the advantage of AE^2: the estimate $\tilde{N}_i$ at first rises sharply reaching N with some overshoot, and right after begins a steady decline with rate e^{-1}. The convergence is reached higher and sooner with respect to Schoute's case, which, in fact, is proven to asymptotically reach efficiency 0.311 [9].
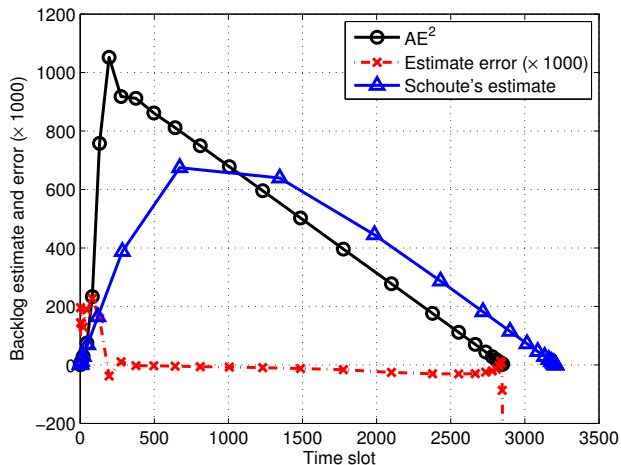
Fig. 6. Average backlog estimate $\hat{N}_i$ at the end of frames versus time slot, for the AE^2 algorithm ($N = 1000$, $r_0 = 1$, $b = 2$). The red dash-dotted line represents the relative error $\times 10^3$ with respect to the analytical values N_i.

VII. CONCLUSIONS

We have presented an analysis of DFA with Frame Restart that provides the optimal FR strategy when the tag population size N is known. The analysis also provides simple sub-optimal mechanisms that do not loose efficiency. Furthermore it shows that asymptotically the FR property presents the same performance as with the plain DFA.

When N is unknown, the performance is determined by the precision of the procedure used to estimate N. Unfortunately, the lack of precision at low and moderate values of N makes the performance comparable with the one of the plain DFA. Furthermore, we provide a procedure that exploits FR to reduce the overhead asymptotically, reaching efficiency e^{-1}, as in case of known N.

ACKNOWLEDGMENT

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*, International Organization for Standardization Std., 2004.

[3] *Class 1 Generation 2 UHF Air Interface Protocol Standard Version 1.0.9*, EPCglobal Std., 2005.

[4] 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.

[5] ——, "A critical survey and analysis of RFID anti-collision mechanisms," *IEEE Commun. Mag.*, vol. 49, no. 5, pp. 214 –221, May 2011.

[6] H. Vogt, "Efficient object identification with passive RFID tags," in *Proc. First Intern. Conf. Pervasive Computing*, 2002, pp. 98–113.

[7] C. Floerkemeier, "Bayesian transmission strategy for framed Aloha based RFID protocols," in *IEEE Intern. Conf. on RFID*, Mar. 2007, pp. 228 –235.

[8] 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*, 2014, accepted for publication.

[9] ——, "An asymptotically efficient backlog estimate for dynamic frame Aloha," *arXiv:1305.0909 (cs.IT)*, 2013.

[10] 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. Science and Engineering*, vol. 6, no. 1, pp. 9 –15, Jan. 2009.

[11] L. Zhu and T.-S. Yum, "Design and analysis of framed Aloha based RFID anti-collision algorithms," in *Proc. GLOBECOM 2009*, 2009, pp. 1–7.

[12] L. Barletta, F. Borgonovo, and M. Cesana, "Performance of frame-Aloha protocols: Closing the gap with tree protocols," in *Proc. MEDHOCNET 2011*, 2011.

[13] B. Knerr, M. Holzer, C. Angerer, and M. Rupp, "Slot-by-slot maximum likelihood estimation of tag populations in framed slotted Aloha protocols," in *Int. Symp. on Performance Evaluation of Comp. and Telecom. Systems*, 2008, pp. 1–7.