Ergodicity of Controlled and Uncontrolled Multipacket Aloha

Sylvie Ghez, Sergio Verdú and Stuart C. Schwartz

Department of Electrical Engineering
Princeton University
Princeton NJ 08544

1. The multipacket channel model

There are many random access channels, such as with code division multiple access channels [10], radio networks with capture [8], or systems with multiuser receivers [13], for which the simultaneous transmission of several packets does not necessarily result in the destruction of all the the information that was sent, and therefore for which the usual collision channel model [1] does not apply. In this paper, we study a model for an infinite user slotted channel with a multipacket reception capability. It is assumed that the number of correctly received packets in a slot is a random variable that can take any integer value between zero and the collision size, and that depends only on the number of simultaneous transmissions. We define for $n \geq 1$ and $0 \leq k \leq n$: $\epsilon_{nk} = P[\, k$ packets are correctly received $\mid n$ are transmitted $]$. Thus the multipacket reception properties of the channel are described by the stochastic matrix $\mathbf{E} = (\epsilon_{nk})_{n \geq 1, 0 \leq k \leq n}$, to which we refer as the ***reception matrix*** of the channel. Note that this model allows not only collisions but also background noise to be a source of errors, by letting $\epsilon_{10} \neq 0$. An important parameter is the average number of packets correctly received in collisions of size n, denoted by $C_n = \sum_{k=1}^{n} k\epsilon_{nk}$. Denote by A_k be the number of new packets arriving during time slot k, it is assumed that $(A_k)_{k \geq 0}$ are i.i.d. random variables with a finite mean λ.

2. The uncontrolled Aloha algorithm in the multipacket channel

We consider in this section the usual Aloha random access algorithm. New packets are transmitted with probability one immediately after their arrivals, and backlogged packets are retransmitted in each slot with a constant probability p $(0 < p \leq 1)$ until successful reception. By the end of each slot, users receive some feedback telling them if their transmission was successful or not. The number of backlogged packets in the system at time n, X_n, is easily seen to be a homogeneous Markov chain. We define the system to be stable if $(X_n)_{n \geq 0}$ is ergodic and unstable otherwise. All the results here hold under the assumption that $(X_n)_{n \geq 0}$ is irreducible and aperiodic, which is verified in all cases of practical interest. Details and a set of sufficient conditions can be found in [3].

Theorem 1. If C_n has a limit $\eta_0 = \lim_{n \to \infty} C_n$, then the system is stable for all arrival distributions such that $\lambda < \eta_0$ and is unstable for $\lambda > \eta_0$. This also holds if η_0 is infinite: if $\lim_{n \to \infty} C_n = +\infty$, then the system is always stable.

Outline of proof. The proof is based on drift analysis. In general, the drift at state n $(n \geq 0)$ is defined by $d_n = E[\, X_{t+1} - X_t \mid X_t = n \,]$. The one-step transition probability matrix of $(X_n)_{n \geq 0}$, (P_{ij}), can be computed as a function of p, $(\lambda_k)_{k \geq 0}$ and $\mathbf{E}$; and the drifts can in turn be computed from (P_{ij}). The key point of the proof is to to show that $\lim_{i \to \infty} d_i = \lambda - \eta_0$. Then if $\lambda < \eta_0$, the drift is negative when the backlog becomes sufficiently large, which is enough from [7] to guarantee the ergodicity of X_n. If $\lambda > \eta_0$, the drift is positive except maybe for a finite number of states; moreover, Kaplan's condition [6] is shown to be verified in [3], and therefore X_n is nonergodic.

A noteworthy feature of Theorem 1 is that any number of modifications of the reception matrix that leaves $\lim_{n \to \infty} C_n$ unchanged does not affect the stability region. Although this may be surprising at first sight,

This research was partially supported by the Office of Naval Research under Contract N00014-87-k-0054 and by the Army Research Office under Contract DAAL03-87-k-0062.

it can be intuitively explained by the fundamental instability of the collision channel: unless the receiver is perfect (all ϵ_{nn} equal to 1), the backlog will eventually exceed any prefixed value with probability one, thus it is the limit of C_n that determines the stability region.

In almost all practical cases the se· ·ence $(C_n)_{n\geq 1}$ does have a limit. It is however conceptually interesting to examine the case when $\lim\limits_{n\to\infty}$ in $_n \neq \limsup\limits_{n\to\infty} C_n$. It can be shown that the system is stable for $\lambda < \liminf\limits_{n\to\infty} C_n$ and unstable for $\lambda > \limsup\limits_{n\to\infty} C_n$. However to compute the exact value of the maximum stable throughput, it is in general necessary to know completely the sequence $(C_n)_{n\geq 1}$ and examples can be build for which η_0 takes any value between $\liminf\limits_{n\to\infty} C_n$ and $\limsup\limits_{n\to\infty} C_n$.

It has been proved in [9] that the backlog Markov chain for the usual slotted Aloha algorithm is transient, but this result cannot be extended in full generality to our model when $\lambda > \eta_0$. For instance, consider the reception matrix defined by $\epsilon_{nk} = \frac{1}{n^2}$ $(1\leq k \leq n, n\geq 1)$ for which $\eta_0=1/2$. It can be shown [4] that X_n is null recurrent for $1/2<\lambda<R(p)$ and transient for $\lambda>R(p)$, where $R(p) = \frac{1}{p} + \frac{(1-p)}{p^2}\ln(1-p)$ for $p<1$ and $R(1)=1$. In general, the transience region depends in a rather complicated manner on the elements of the reception matrix and on the retransmission probability p; thus in most cases it can only be approximated. However, it can be shown that transience in the nonergodicity region is ensured if the supremum of the elements of the k^{th} column goes to zero faster than k^2. This condition holds for all the examples in [3]-[4], as well as for many real life cases, due to the practical limitations on the receiver capabilities.

3. Optimal decentralized control algorithms

Let us consider now control schemes of the form $p_n = F(S_n)$ with $S_{n+1} = G(S_n, Z_n)$, where p_n is the retransmission probability in slot n, S_n is an estimate of the backlog X_n at the beginning of slot n, and Z_n is the feedback at the end of slot n. Each of the A_{n-1} new packets that arrived during slot $n-1$ is transmitted in slot n with probability p_n. Also, it is assumed in this section that the limit $\eta_0 = \lim\limits_{n\to\infty} C_n$ exists. As in the case of conventional collision channels, it is useful to study first the case of control with perfect state information where the value of the backlog is known prior to the selection of the retransmission probability. Even though this situation is not relevant in practical systems, its study enables us to determine the best throughput η_c achievable by control algorithms of the form mentioned above, as well as the corresponding optimal retransmission probability. For instance, it is well known [2] that for the usual collision channel with the delayed first transmission rule that we are using here, $p_n = 1/X_n$ is the optimal retransmission probability, resulting in an ideal throughput $\eta_c = e^{-1}$.

Therefore, let the retransmission probability be a function of the exact value of the backlog, i.e., $p_n = F(X_n)$. In this ideal case, the system is much simpler to analyze, since $(X_n)_{n\geq 0}$ is a homogeneous Markov chain. We first determine the throughput of the system if we simply set $p_n = A/X_n$, which is analogous to the optimal retransmission probability for the usual collision channel. It is shown in [5] that if A is any positive constant and if $p_n = A/X_n$ for $X_n \geq A$, then the system is stable for $\lambda < t(A)$ and unstable for $\lambda > t(A)$ with $t(A) = e^{-A} \sum\limits_{n=1}^{\infty} C_n \frac{A^n}{n!}$. Note that the value of p_n for $X_n < A$ is left unspecified because it does not affect the throughput. The proof of this result involves essentially the same methods as the proof of Theorem 1.

Theorem 2. There exists a retransmission probability p_n^* that minimizes the drift d_n at state n. With such a retransmission probability, the system is stable for $\lambda<\eta_c$ and unstable for $\lambda>\eta_c$, where $\eta_c = \sup\limits_{x\geq 0} t(x)$, and $t(x) = e^{-x} \sum\limits_{n=1}^{\infty} C_n \frac{x^n}{n!}$.

Outline of proof. If $X_t = n \geq 1$ and if we denote by p the retransmission probability used in slot t, the drift at state n is given by $d_n(p) = \lambda - \sum\limits_{j=1}^{n} \binom{n}{j} p^j (1-p)^{n-j} C_j = \lambda - t_n(p)$, with

$t_n(p) = \sum_{j=1}^{n} \binom{n}{j} p^j (1-p)^{n-j} C_j$. Define $p_n^* = \arg\max_{p \in [0,1]} t_n(p) = \arg\min_{p \in [0,1]} d_n(p)$. The main difficulty here is to show that $\lim_{n \to \infty} t_n(p_n^*) = \sup_{x \geq 0} t(x)$. The theorem basically follows from this equality by invoking as before the results in [6] and [7]. The inequality $\liminf_{n \to \infty} t_n(p_n^*) \geq \sup_{x \geq 0} t(x)$ can be deduced rather easily from the result mentioned earlier in this section. Then we use the fact that $(C_n)_{n \geq 1}$ has a limit to build a new system, which dominates the original one in the sense that its average number of successes per slot is larger than $t_n(p)$ for all n and p. This new system has a maximum stable throughput which is larger than η_c only by an arbitrary $\epsilon > 0$, but has by construction a behavior which is much simpler and enables us to show that $\limsup_{n \to \infty} t_n(p_n^*) \leq \sup_{x \geq 0} t(x)$.

The most important consequence of Theorem 2 is that it is not necessary to compute the value of the optimal retransmission probability p_n^* (which does not have a close form except in some very simple cases). More precisely, if $t(x)$ achieves its supremum, i.e. if there exists A, $0 < A < +\infty$, such that $t(A) = \sup_{x \geq 0} t(x)$, then the simple control $p_t = A / X_t$ for $X_t > A$ yields a maximum stable throughput $t(A) = \eta_c$, meaning that such a closed-loop control is optimal. If $t(x)$ does not reach its supremum, then it can be proved that $\eta_c = \eta_0$, which means that it is not possible to improve the throughput by letting the retransmission probability depend on the channel feedback.

We are now ready to analyze an optimal control algorithm in the partial state information case. To update the backlog estimate S_n, we only need a binary feedback: if slot n is empty then $Z_n = 0$ and otherwise $Z_n = \bar{0}$. As suggested by our earlier results, we consider certainty-equivalence retransmission probabilities of the form $p_n = A / S_n$.

Theorem 3. Assume that there exists $A \in]0, +\infty[$ such that $t(A) = \sup_{x \geq 0} t(x)$. If $\alpha < 0$ and $\beta > 0$ verify the following two conditions,

C1. $\beta > \lambda$

C2. $\beta(1 - e^{-A}) + \eta_c - \lambda + \alpha e^{-A} = 0$

then the control algorithm $p_t = \dfrac{A}{S_t}$, $S_{t+1} = \max\{ A, S_t + \alpha I(Z_t = 0) + \beta I(Z_t = \bar{0}) \}$ has a maximum stable throughput η_c.

Outline of proof. The proof a streamlined version of the method developed in [11] for the special case of the usual collision channel. A detailed proof of a restricted version of Theorem 3 can be found in [5]. The idea is to use the properties of the homogeneous two dimensional vector Markov chain $M_t = (X_t, S_t)_{t \geq 0}$ to build a Lyapunov function whose drift is negative in the entire plane. The basic building blocks of the proof are the drift of the backlog, $c(n,s) = E[X_{t+1} - X_t \mid M_t = (n,s)]$, and the drift of the error in the backlog estimate $\tilde{X}_t = S_t - X_t$, $d(n,s) = E[\tilde{X}_{t+1} - \tilde{X}_t \mid M_t = (n,s)]$. The properties of these drifts that are needed to build a well behaved Lyapunov function depend essentially on the ratio $x = n/s$ between the true value of the backlog and its estimated value. There exists a small $\gamma > 0$ such that if $|x - 1| < 5\gamma$, the backlog estimate being close to its ideal value, we have $c(n,s) < 0$, as in the perfect state information case. If $x < 1 - \gamma$, then the value of the estimate is too large ($\tilde{X}_t > 0$), and the drift of the error is negative: $d(n,s) < 0$. Similarly, if $x > 1 + \gamma$, then $d(n,s) > 0$. The constants α and β in the definition of S_t have been chosen to ensure that this last two properties are verified except maybe in a finite region of the state space. Consider the following Lyapunov function

$$V(n,s) = \begin{cases} n & \text{if } |x - 1| < 3\gamma \\ \dfrac{1+3\gamma}{3\gamma}(n - s) & \text{if } x > 1 + 3\gamma \\ \dfrac{1-3\gamma}{3\gamma}(s - n) & \text{if } x < 1 - 3\gamma \end{cases}$$

We consider separately likely and unlikely events:

$$E[V(M_{t+1}) - V(M_t)\,|M_t=(n,s)] = E[(V(M_{t+1}) - V(M_t))I(\,|A_t - \Sigma_t\,|\le J)\,|M_t=(n,s)] + E[(V(M_{t+1}) - V(M_t))I(\,|A_t - \Sigma_t\,|\ge J)\,|M_t=(n,s)] \tag{1}$$

If $|A_t - \Sigma_t| = |X_{t+1} - X_t| \le J$, then the changes of regions between M_t and M_{t+1} are limited. Using this fact, it can be shown that the first term in (1) is a function of the drifts of $X_t I(\,|A_t - \Sigma_t\,|\le J)$ and $\tilde{X}_t I(\,|A_t - \Sigma_t\,|\le J)$, which are close approximations of $c(n,s)$ and $d(n,s)$ respectively when J is large enough. Then by using the aforementioned properties of $c(n,s)$ and $d(n,s)$, it is possible to show that the first term in (1) is strictly negative outside a finite region of the state space. The second term in (1) basically corresponds either to unlikely events when J is large, or to events that yield a negative drift. Finally, there exists a positive constant μ such that $E[V(M_{t+1}) - V(M_t)\ |\ M_t=(n,s)] \le -\mu$ for all (n,s) outside a finite region of the state space. From [17], this is enough to conclude that $M_t = (X_t, S_t)$ is ergodic.

Theorems 1 and 3 give a complete and general solution of the maximum throughput achievable with uncontrolled and controlled Aloha in a channel with multipacket reception capability. The maximum open-loop throughput is $\lim_{n\to\infty} C_n$, whereas the maximum closed-loop throughput is $\sup_{x\ge 0} e^{-x} \sum_{n=1}^{\infty} C_n \frac{x^n}{n!}$, where C_n is the expected number of successfully received packets in a collision of size n. Applications of these formulas to specific problems in frequency-hopping random access channels, collision channels with power or delay capture and mobile radio networks with pairwise transmissions can be found in [3]-[5].

References

[1] D. Bertsekas and R. Gallager, *Data Networks,* Prentice Hall, 1987.

[2] G. Fayolle, E. Gelenbe and J. Labetoulle, "Stability and Optimal Control of the Packet Switching Broadcast Channel", *J. Assoc. Comput. Machinery,* vol. 24, no. 3, pp. 375-386, 1977.

[3] S. Ghez, S. Verdú and S.C. Schwartz, "Stability of Multipacket Aloha", *Proc. Twenty First Conf. on ISS, The Johns Hopkins University,* March 87.

[4] S. Ghez, S. Verdú and S.C. Schwartz, "Stability Properties of Multipacket Aloha with Multipacket Reception Capability", Princeton University, preprint, 1987.

[5] S. Ghez, S. Verdú and S.C. Schwartz, "On Decentralized Control Algorithms for Multipacket Aloha", *Proc. Twenty Fifth Allerton Conf. on Communications, Control and Computing,* October 1987.

[6] M. Kaplan, "A Sufficient Condition for Nonergodicity of a Markov Chain", *IEEE Trans. Inform. Theory* , vol. IT-25, pp. 470-471, 1979.

[7] A.G. Pakes, "Some Conditions for Ergodicity and Recurrence of Markov Chains", *Opns. Res.,* vol. 17, pp. 1058-1061, 1969.

[8] L.G. Roberts, "Aloha Packet System with and without Slots and Capture", *Comput. Commun. Rev.,* no 5, pp. 28-42, 1975.

[9] W.A. Rosenkrantz and D. Towsley, "On the Instability of the Slotted Aloha Multiaccess Algorithm", *IEEE Trans. Automatic Control,* vol. AC-28, pp. 994-996, 1983.

[10] M.K. Simon, J.K. Omura, R.A. Scholtz, B.K. Levitt, *Spread-Spectrum Communications* , Computer Science Press, 1985.

[11] J.N. Tsitsiklis, "Analysis of a Multiaccess Control Scheme", *IEEE Trans. Autom. Control,* vol. AC-32, pp. 1017-1020, Nov. 1987.

[12] R.L. Tweedie, "Sufficient Conditions for Ergodicity and Recurrence of Markov Chains on a General State Space", *Stoch. Proc. Appl.,* no. 3, pp. 385-403, 1975.

[13] S. Verdú, "Minimum Probability of Error for Asynchronous Gaussian Multiple-Access Channels", *IEEE Trans. Inform. Theory* , vol. IT-32, pp. 85-96, 1986.