

Delay-Throughput Comparison of Single and Multi-Channel Slotted ALOHA Networks

F. L. Lo
fll@hkueee.hku.hk

T. S. Ng
tsng@hkueee.hku.hk

T. I. Yuk
tiyuk@hkueee.hku.hk

Department of Electrical and Electronic Engineering,
The University of Hong Kong, Pokfulam Road, Hong Kong
Tel.: (852) 2859-2686 Fax.: (852) 2559-8738

Abstract – In networks where stations are equally likely to transmit to one another, simple slotted ALOHA is not an appropriate form of multiple access protocol as messages meant for different stations can collide in the common channel. In this paper the performance of a single channel slotted ALOHA network is compared to that of a CDMA ALOHA network where each of the station is assigned a different spreading code with which to receive messages, and thus effectively possesses its own private ‘virtual channel’.

I. INTRODUCTION

After ALOHA multiple access was first introduced in [1], its slotted version was investigated in detail in [2], where both Markov analysis and fluid-flow approximation, later called equilibrium point analysis (EPA), were used. The first technique yields the exact average delay and throughput values of a slotted ALOHA network, while the latter approximates the average values with those obtained from a network at equilibrium. EPA thus yields non-exact values, but these can be accurate when the network is stable and uncongested, as demonstrated in [2,3]. It was also shown in [2] that slotted ALOHA provides high throughput and low delay only when the message generation probability per slot, s , and the retransmission probability, p , are kept very low. Beyond these values, the network easily reaches regions of the system phase space where the throughput drops significantly and delay increases to unacceptable values.

In LANs where applications and databases are distributed over all the machines, and where each station is equally likely to transmit to any other station, simple slotted ALOHA suffers from the defect that messages meant for different stations may collide unnecessarily in the common channel. One way of overcoming this limitation is to use spread-spectrum signalling and assign a different spreading code to each station with which they can receive messages. If the spreading codes are chosen so that the cross-correlation between them is low, then effectively each station has its own ‘virtual channel’ with which it can receive messages. In this paper we will investigate the delay-throughput characteristics of this form of CDMA network through EPA, and compare its performance against that of a simple slotted ALOHA network.

II. NETWORK MODEL

Each station in the network is assigned a different spreading code, as in the ‘receiver-based code’ protocol of [4], or a different phase of the same spreading code, if the system is synchronized,

as in [5], with which it will receive messages. Time is divided into slots and each slot is just long enough to allow the transmission of a single packet or message. Stations are presented with an average of s newly generated message per slot, and we shall assume that $s < 1$, so that we can also interpret s as the message generation probability per slot for each of the station. Stations are equally likely to transmit to one another, and once an idle station is presented with a new message, it will attempt to transmit it during the next slot with probability one.

If no other stations transmit to the same station during that slot, then the message will be received correctly by the intended receiver. Should two or more stations transmit to the same station during a slot, then the messages will collide. Those stations with unsuccessfully transmitted messages will be called blocked, and blocked stations will attempt to retransmit their messages with probability p during each of the slots subsequent to the collision until they are successful. Stations have only one buffer to hold messages, and therefore blocked stations will not accept newly generated messages until they become idle again. A small acknowledgment slot follows each message slot so that stations will know immediately after the slot whether their transmissions have been successful.

Systems are assumed to be ideal, and all imperfections such as finite cross-correlations between different spreading codes, timing jitter, and multi-path fading are not considered.

III. ANALYSIS OF MODEL

Although the model proposed can be analyzed, in principle, by Markov analysis, the amount of computation required for systems with a large number of stations is prohibitive. A simpler approach that can be used is the one first discussed in [2], where it was used to study slotted ALOHA and called fluid-flow approximation. This method is explained in detail as equilibrium point analysis (EPA) in [3]. The idea behind it is that the system in the long run usually operates around its equilibrium point where the inflow of packets is balanced by the outflow of packets.

Consider a network of N stations, each equally like to transmit to any other station. We shall denote the number of idle stations at any time-slot by n_0 and the number of blocked stations at channel k by n_k . Channels with blocked stations we shall call occupied, and channels with none we shall call free.

For simple slotted ALOHA, n_0 and $n_b = N - n_0$ determine the average system throughput completely. In our case, the system throughput depends on how the blocked stations are distributed among the occupied channels. We wish to determine the worst case performance of the system and so shall find the lowest

system throughput given a certain n_0 and n_b . First we need to develop a notation to show the state the system is in. We shall use $\{(s_{11}, s_{12}, \dots, s_{1n_1})(s_{21}, s_{22}, \dots, s_{2n_2}) \dots\}$ to denote the state where stations $s_{11}, s_{12}, \dots, s_{1n_1}$ are blocked trying to transmit to station 1, or through channel 1, etc. For example, the state $\{(0)(1,3)(2)(0)(0)\}$ for a 5-station system shows that stations 1 and 3 are blocked trying to transmit through channel 2 while station 2 is blocked trying to transmit through channel 3. Stations 4 and 5 are, of course, idle. Note also that in the above example, channels 2 and 3 are occupied while channels 1, 4 and 5 are free.

For an idle station i , the throughput due to it is

$$S_i = sN^{-1}(1 - sN^{-1})^{n_0-1} \sum_{k \text{ not idle}} (1-p)^{n_k} + sN^{-1}(1 - sN^{-1})^{n_0-2} \sum_{k \text{ idle}, k \neq i} (1-p)^{n_k}. \quad (1)$$

Note that the exact S_i involves $s/(N-1)$ terms and not s/N , but for large N , $N \cong N-1$, and the above expression is a very good approximation. Given that the system has a certain number of blocked stations distributed among some occupied channels, we can only choose whether a particular channel is free or occupied. To minimize S_i , we should choose channel i to be unoccupied

and therefore eliminate a term $sN^{-1}(1 - sN^{-1})^{n_0-2}$ which is the largest possible in the above summation. Since the number of free channels is greater than or equal to the number of idle stations, we can minimize the throughput of the idle stations by requiring all of them to have free channels, and this is always possible as long as the number of occupied channels is greater than one. Thus, for example, if 2 blocked stations are at an occupied channel and one is at another occupied channel for a 5-station system, the states which will yield the minimum idle throughput are $\{(2,5)(1)(0)(0)(0)\}$, $\{(0)(4)(0)(2,3)(0)\}$, etc.

For a blocked station b trying to transmit through channel k , throughput due to it is

$$S_b = \begin{cases} p(1-p)^{n_k-1}(1 - sN^{-1})^{n_0-1} & \text{if } k \text{ is idle} \\ p(1-p)^{n_k-1}(1 - sN^{-1})^{n_0} & \text{if } k \text{ is blocked.} \end{cases} \quad (2)$$

Since $(1-s/N) < 1$, to obtain the minimum throughput for a blocked station, we should arrange that the station to which the occupied channel belongs is blocked. Thus to minimize the total throughput for blocked stations, we should require that all stations which have occupied channels be blocked. For example, for the case where there are 2 blocked stations distributed one to each occupied channel of a 5-station system, the minimum throughput for blocked stations can be obtained for states $\{(3)(0)(1)(0)(0)\}$, $\{(0)(0)(0)(5)(4)\}$, etc.

From the above discussion, we see that given the system is in a certain state with more than one occupied channel, minimum throughput is obtained by requiring that all idle stations have free channels and all occupied channels belong to blocked stations. Thus the minimum throughput of a certain state depends only on whether idle or blocked stations have occupied or free channels. We can therefore simplify our state notation by writing down configurations as only $\{n_1\sigma_1 n_2\sigma_2 \dots\}$, where n_i is the number of blocked stations trying to transmit through channel i and σ_i denotes the state of the station which owns the channel i , b for

blocked and d for idle. All $n_i = 0$ can be ignored. As an example, the state $\{(2,5)(1)(0)(3)(0)\}$ can be represented more compactly as $\{2b, 1b, 1d\}$.

For the state $\{n_b d\}$ where all the blocked stations are trying to transmit through the same channel, which obviously must belong to an idle station, there is no choice involved and the throughput is

$$S_{out} = (n_0 - 1)sN^{-1}(1 - sN^{-1})^{n_0-2}(1-p)^{n_b} + sN^{-1} \cdot (1 - sN^{-1})^{n_0-2}(N-1) + (n_0 - 1)sN^{-1}(1 - sN^{-1})^{n_0-2} \cdot (N-2) + n_b(1 - sN^{-1})^{n_0-1}p(1-p)^{n_b-1} \quad (3)$$

The first term is due to idle stations transmitting through the occupied channel, the second term due to the idle station with the occupied channel transmitting through free channels, the third term is due to other idle stations transmitting through free channels and the last term is due to the blocked stations transmitting.

Consider the case of a system with n_b blocked stations distributed among several occupied channels so that each occupied channel has n_{b1}, n_{b2}, \dots blocked stations respectively. All the n_0 idle stations have free channels and all the occupied channels belong to blocked stations. This state $\{n_{b1}b, n_{b2}b, \dots, n_{bb}b\}$ has the minimum throughput for the system with $n_{b1}, n_{b2}, \dots, n_{bb}$ blocked stations distributed among b occupied channels. This throughput is

$$S = n_0sN^{-1}(1 - sN^{-1})^{n_0-1} \sum_{n_k \neq 0} (1-p)^{n_k} + n_0sN^{-1} \cdot (1 - sN^{-1})^{n_0-2}(n_0 - 1) + n_0sN^{-1}(1 - sN^{-1})^{n_0-1} \cdot (N - b - n_0) + \sum_{n_k \neq 0} (1 - sN^{-1})^{n_0} n_k p(1-p)^{n_k-1} \quad (4)$$

The first summation is due to idle stations transmitting through occupied channels. The second term is due to idle stations transmitting to other idle stations. The third term is due to idle stations transmitting through the rest of the free channels, and the last summation is due to blocked stations transmitting.

Suppose at least one of the n_{bi} is one, i.e., at least one of the occupied channels has only one blocked station. If we free this channel by moving the blocked station to another occupied channel with m blocked stations, the new state will have a different throughput. The total change in throughput is

$$\Delta S_m = n_0sN^{-1}(1 - sN^{-1})^{n_0-1}[1 + (1-p)^{m+1}] - n_0sN^{-1} \cdot (1 - sN^{-1})^{n_0-1}[(1-p) + (1-p)^m] + (1 - sN^{-1})^{n_0} \cdot [(m+1)p(1-p)^m] + (1 - sN^{-1})^{n_0}[mp(1-p)^{m-1} - p] \quad (5)$$

$$= n_0sN^{-1}(1 - sN^{-1})^{n_0-1}p[1 - (1-p)^m] + (1 - sN^{-1})^{n_0}p\{(1-p)^m[1 - mp(1-p)^{-1}] - 1\}$$

If we let $c = (1 - sN^{-1})^{n_0-1}$ then

$$c^{-1}\Delta S_m = n_0sN^{-1}p[1 - (1-p)^m] + (1 - sN^{-1})p\{(1-p)^m[1 - mp(1-p)^{-1}] - 1\} = A_m + B_m \quad (6)$$

The term A_m is positive and grows monotonically to $n_0 s N^{-1} p$. The term B_m is negative and reaches $-(1 - s N^{-1})p$ as $m \rightarrow \infty$. Thus $c^{-1} \Delta S_m \rightarrow p[(n_0 - 1)s N^{-1} - 1]$ which is negative since $s < 1$. Therefore total throughput can always be reduced by putting single blocked stations onto other occupied channels.

To find the state with the minimum throughput given that there are $n_b = N - n_0$ blocked stations, we can start out with the configuration $\{1b, 1b, \dots, 1b\}$ which has the minimum throughput for the system with one blocked station per occupied channel. When we free an occupied channel by putting the single blocked station onto another channel which already has one blocked station, the reduction in throughput ΔS_I can be easily calculated from (5). If we continue to free another channel, we have the choice of putting the single blocked station onto the channel with 2 stations or any one of the others with just 1. If $\Delta S_2 < \Delta S_1$, throughput is decreased more by putting the single blocked station onto the channel with 2 blocked stations to obtain $\{3b, 1b, 1b, \dots, 1b\}$. As we keep freeing occupied channels with single stations, at a certain point, having two blocked channels will yield a lower throughput than just having one. This point can be obtained by finding the change in throughput by putting single blocked stations onto one channel ΔS_I , and comparing it to ΔS_{II} , the change in throughput obtained by putting single blocked stations alternately between two occupied channels. If n_b blocked channels are cleared before the point where $\Delta S_{II} < \Delta S_I$, we have a choice of either having state $\{(n_b - 1)b, 1b\}$ or $\{n_b d\}$. By calculating the throughput for each, we can decide which state has the lower throughput.

After we find the point where $\Delta S_{II} < \Delta S_I$, if there are still channels with single blocked stations, we can reduce the throughput further by putting single blocked stations alternately among the two occupied channels or putting them onto a third channel. We can compare the change in throughput obtained and find out when distributing blocked stations among three channels yield a lower throughput than distributing them among two. This process can be continued as long as there are remaining single blocked stations. This process, even though requiring a lot of words to describe, can be carried out efficiently with a simple computer program.

Any state where all idle stations have free channels and all occupied channels belong to blocked stations, assuming that there are two or more occupied channels, can be obtained by putting single blocked stations from state $\{1b, 1b, \dots, 1b\}$ onto other occupied channels. Since the above procedure ensures that the reduction of throughput from the configuration $\{1b, 1b, \dots, 1b\}$ is maximum, thus the configuration obtained will have the minimum throughput. For the case $n_b = N$, where all stations are blocked, $A_m = 0$ since $n_0 = 0$, so putting all stations onto one channel should yield the lowest throughput. Of course this is not possible since a station will not transmit to itself, so we need to choose between the configurations $\{\lfloor N/2 \rfloor b, \lceil N/2 \rceil b\}$ or $\{(N-1)b, 1b\}$, where $\lfloor x \rfloor$ means the largest integer not exceeding x and $\lceil x \rceil$ means the smallest integer to exceed x . We should not have more than two blocked channels as ΔS_m is small when m is small,

which means that the reduction in throughput will not be maximum with three or more blocked channels.

Table 1 gives the states for minimum throughput for $N=50$, $S=0.2$ and $p=0.5$. If $n_b \leq 13$, then all the blocked stations should be placed onto one single occupied channel. If $14 \leq n_b < 24$, then the blocked stations should be divided as equally as possible among two channels. If $24 \leq n_b < 39$, then the blocked stations should be divided as equally as possible among three channels, etc.

IV. RESULTS OF EPA OF MULTI-CHANNEL NETWORKS

From the graph of S_m and minimum S_{out} against n_b , e.g., the one shown in Fig. 1, if the minimum S_{out} curve intersects the S_m curve only once at a region where n_0 is large, it is possible to find L , the maximum number of users blocked at equilibrium, and λ , the arrival rate of packets at equilibrium. Using Little's Formula, $L = \lambda W$ [6], we can easily find the average delay W . The values obtained will be the worst case figures, i.e., they show the lowest throughput and highest delay at equilibrium. Table 2 gives the throughput and delay values from EPA as well as simulation for $N = 50$, $s = 0.05$ and various values of p . The simulations are carried out for 100,000 slots.

It can be seen that the values obtained from EPA compare quite well with those from simulation. The disagreement between EPA and simulation occurs where EPA predicts unstable behavior while simulation shows no congestion for $p = 0.5$. This is not surprising considering that the simulation is carried out for a relatively few slots. Furthermore, unlike simple slotted ALOHA where the system throughput can only move along a line in system phase space, and once it crosses the middle unstable equilibrium point is driven inexorably into the congested equilibrium region [2], CDMA ALOHA throughput can jump over a band of values, and thus even when it enters the congested region, the system may not remain there for long.

V. COMPARISON OF SINGLE AND MULTI-CHANNEL NETWORKS

Delays and throughputs of single channel slotted ALOHA systems are obtained by Markov analysis as discussed in [2]. The only difference between that used in [2] and the one we used to derive the results shown in this paper is that we assume that stations are able to receive messages sent to them instantaneously, unlike [2] where a satellite network is analyzed, and thus some time has to be allowed for the message to arrive at the receiver.

Table 3 shows the delay and throughput of a single channel slotted ALOHA network with various values of s and p . For $s = 0.006$, the single channel system becomes congested, i.e., suffer from low throughput and large delay when p crosses about 0.12. For the $s = 0.008$ case, congestion occurs earlier, when p crosses 0.10.

Table 4 shows delay and throughput of a multi-channel slotted ALOHA system for the same set of s and p values. For all the message generation and retransmission probabilities considered, the network is able to handle the maximum average possible

message generation rate, Ns , and all successfully transmitted packets suffer from minimal delays.

VI. CONCLUSIONS

We have shown how EPA can be used to analyze multi-channel slotted ALOHA networks with a large number of stations, and demonstrated that the results obtained agree well with the values obtained by simulations, when the networks are stable. EPA also allows us to decide whether networks are unstable or congested. Lastly, a further advantage of slotted ALOHA networks using CDMA is that they can be overlaid on an environment which already has equipment using narrow band radio signalling. As long as the loading is kept low, the CDMA ALOHA transmissions will look like weak wideband noise to the existing users.

ACKNOWLEDGMENT

This work was supported by the Hong Kong Research Grant Council and by the CRCG of The University of Hong Kong.

REFERENCES

- [1] N. Abramson, "The ALOHA System — Another Alternative for Computer Communications". *Proc. AFIPS Conf., 1970 Fall Joint Comput. Conf.*, Vol.37, pp. 281-285, 1970.
- [2] L. Kleinrock, S. Lam, "Packet Switching in a Multiaccess Broadcast Channel: Performance Evaluation", *IEEE Trans. on Commun.*, Vol. COM-23, pp. 410-423, 1975.
- [3] S. Tasaka, "Performance Analysis of Multiple Access Protocols", MIT Press, 1986
- [4] E. Sousa, J. Silvester, "Spreading Code Protocols for Distributed Spread-Spectrum Packet Radio Networks", *IEEE Trans. on Commun.*, Vol. COM-36, pp. 272-281, 1988.
- [5] K.-W. Yip, T.-S. Ng, "Code Phase Assignment - A Technique for High Capacity Indoor Mobile DS-CDMA Communications", *IEEE 44th Veh. Tech. Conf.*, pp. 1586-1590, 1994
- [6] J. Little, "A Proof of the Queuing Formula $L = \lambda W$ ", *Operations Research*, Vol. 9, pp. 383-387, 1961

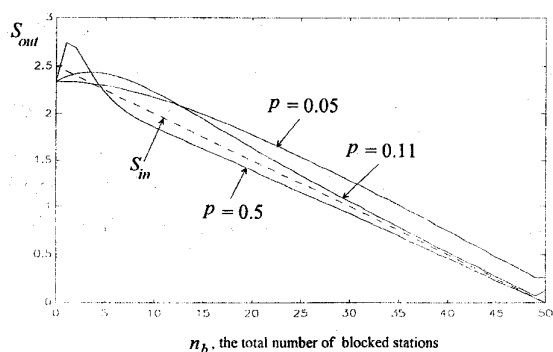


Fig. 1: Graph of S_{out} against n_b for $N=50$, $s=0.05$ and various values of p .

Total n_b	States for min. S_{out}
13	{13b}
14	{7b 7b}
24	{8b 8b 8b}
39	{10b 10b 10b 9b}
47	{16b 16b 15b}
50	{25b 25b}

Table 1: Configuration for min. throughput for $N=50$, $s=0.2$ and $p=0.5$.

P	EPA		Simulation	
	Thruput	Delay	Thruput	Delay
0.05	2.32	1.51	2.39	1.05
0.07	2.36	1.14	2.42	0.68
0.09	2.38	0.88	2.43	0.56
0.11	2.42	0.66	2.44	0.47
0.50	unstable	unstable	2.48	0.16

Table 2: Throughputs and delays of a multi-channel network for $N=50$, $s=0.05$, and for various values of p .

S	P	throughput	delay
0.006	0.04	0.2777	13.32
0.006	0.06	0.2834	9.69
0.006	0.08	0.2863	7.85
0.006	0.10	0.2878	6.97
0.006	0.12	0.0690	557.83
0.008	0.04	0.3383	22.77
0.008	0.06	0.3471	18.98
0.008	0.08	0.3439	20.32
0.008	0.10	0.0411	1092.21
0.008	0.12	0.0084	5835.40

Table 3: Throughput and delay obtained by Markov Analysis for a single-channel slotted ALOHA system.

S	P	EPA		Simulations	
		thr'put	delay	thr'put	delay
0.006	0.04	0.298	0.65	0.299	0.79
0.006	0.06	0.299	0.43	0.299	0.26
0.006	0.08	0.299	0.32	0.300	0.13
0.006	0.10	0.299	0.26	0.300	0.09
0.006	0.12	0.299	0.21	0.300	0.08
0.008	0.04	0.398	0.69	0.401	0.14
0.008	0.06	0.399	0.46	0.400	0.17
0.008	0.08	0.399	0.35	0.400	0.16
0.008	0.10	0.399	0.28	0.401	0.10
0.008	0.12	0.399	0.23	0.401	0.10

Table 4: Throughput and delay for multi-channel slotted ALOHA network obtained from both EPA and simulations over 100,000 slots.