ALOHA PACKET SYSTEM WITH AND WITHOUT

SLOTS AND CAPTURE*

SLOTS AND CAPTURE*

Lawrence G. Roberts
President
Telenet Communications Corporation
Washington, D.C.

As the paper was originally distributed only to other researchers intimately familiar with the area covered by the paper, the paper makes few concessions to the reader along the lines of introductory or tutorial material. Therefore, a bit of background material follows.

ALOHA packet systems were originally described by Abramson ("The ALOHA System--Another Alternative for Computer Communication," Proceedings of the AFIPS Fall Joint Computer Conference, Vol. 37, 1970, pp. 281-285). In an ALOHA a single broadcast channel is shared by a number of communicating devices. In the version originally described by Abramson, every device transmits its packets independent of any other device or any specific time. That is, the device transmits the whole packet at a random point in time; the device then times out for receiving an acknowledgment. If an acknowledgment is not received, it is assumed that a collision occured with a packet transmitted by some other device and the packet is retransmitted after a random additional waiting time (to avoid repeated collisions). Under a certain set of assumptions, Abramson showed that the effective capacity of such a channel is $1/(2e)$.

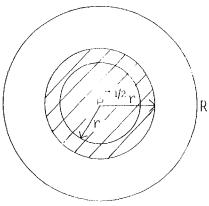(this introductory note is continued at the end of the paper)

Capture

FM receivers will track the strongest of many signals if the power of the next to strongest is down by 1.5 - 3 db, a spec. called the <u>capture ratio</u>. Two major factors affect the power received; the $1/r^2$ distance factor and random individual power variations. The distance factor is major whereas the individual power levels are usually less than 1.5 - 3 db. Since variation will help guarantee that one signal is captured and since individual variations will probably not help much, we will look only at distance variation.

$$\text{Capture Ratio:} \quad CR = 10 \log (P_1/P_2), \ P_1 \text{ is stronger signal}$$

<u>Define</u> $\quad$ Power Ratio: $\quad \beta = P_2/P_1 \quad\quad , \ 0 \le \beta \le 1$

$\quad\quad\quad\quad$ Distance Factor: $\quad r^2{}_1 = \beta r^2{}_2 \quad\quad , \ \beta^{1/2} = r_1/r_2$

Assuming the receiver is surrounded by an equal density population of consoles within a circle of radius R, the only consoles which can interfere with a given station at radius r, are the stations within the circle of radius $\beta^{-1/2} r$.

<u>Define</u>: $\quad \gamma = $ The total bit rate being trans-
$\quad\quad\quad\quad\quad\quad$ mitted by all stations as a
$\quad\quad\quad\quad\quad\quad$ fraction of the total bandwidth.

$\quad\quad\quad \sigma = $ The actual bit rate fraction
$\quad\quad\quad\quad\quad\quad$ being received correctly.

$\quad\quad\quad q = $ Probability of a packet being
$\quad\quad\quad\quad\quad\quad$ received correctly.

<u>Thus</u>: $\quad \sigma = \gamma q$

To compute q for the average station given a transmit rate γ is the task.

There are two cases, asynchronous and synchronous. In both cases we assume a Poisson point process for the totality of stations wanting to start packets of length T seconds at random time with a rate $\lambda = \gamma/T$. Thus, the probability of no conflict within a time (mT) is:

$$q_0 = e^{-\gamma m}$$

However, if only a fraction κ of the stations are in the proper position to prohibit your being captured, the Poisson rate is reduced by κ. At a given radius r, the interfacing stations are those in the area $\pi r^2/\beta$ whereas the total area is πR^2. Thus, the fraction of stations which would cause conflict is the area ratio. Thus:

$$q(r) = e^{-\kappa \gamma m}, \quad \kappa = \frac{r^2}{\beta R^2}$$

The number of stations at any radius r increases linearly with r as the circumference increases. The probability density is therefore:

$$p(r) = \frac{2\pi r}{\pi R^2} \quad , \quad \int_0^R p(r)dr = 1$$

Now, to compute the average q we must integrate the product $p(r)\ q(r)$ over r out to $R\sqrt{\beta}$ and $p(r)\ q_0$ from $R\sqrt{\beta}$ to R, since here all stations interfere.

$$q = \int_0^{R\sqrt{\beta}} p(r)q(r)dr + \int_{R\sqrt{\beta}}^R p(r)q_0 dr$$

$$q = \int_0^{R\sqrt{\beta}} \frac{2r}{R^2}e^{-\frac{r^2\gamma m}{\beta r^2}}dr + \int_{R\sqrt{\beta}}^R \frac{2r}{R^2}e^{-\gamma m}dr$$

$$q = \frac{\beta}{\gamma m}(1-e^{-\gamma m}) + (1-\beta)e^{-\gamma m}$$

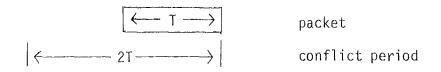Having found the average q, the $\sigma = \gamma q$ relationship will give us σ:

$$\sigma = \frac{\beta}{m}(1-e^{-\gamma m}) + (1-\beta)\gamma e^{-\gamma m}$$

Note that this is the linear combination of two solutions, the solution for no capture (when $\beta=0$) and the solution for full capture, where one packet is always received correctly.

Now it is necessary to look separately at the asynchronous and synchronous cases to determine m.

Asynchronous

A given packet will come into conflict with packets starting within a period 2T since each packet has length T.

←— T —→	packet
←——— 2T ———→	conflict period

Thus, as was derived by University of Hawaii, m=2.

$$\sigma_a = \frac{\beta}{2}(1-e^{-(2\gamma_a)}) + (1-\beta)\gamma_a e^{-(2\gamma_a)}$$

Synchronous Slots

Packets are always put into slots. Each station keeps a clock which counts off slot times and is resynchronized by observing other packets. The slots must have a slightly longer T to prevent errors but this factor will be considered later. The period of packet starting times giving rise to a transmission in a given slot is just the previous slot time. Thus, the Poisson period is T (m=1).

```
|←— T —→|   slot
|←— T —→|   conflict period
```

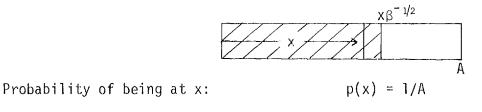$$\sigma_s = \beta(1 - e^{-(\gamma_s)}) + (1-\beta)\gamma_s e^{-(\gamma_s)}$$

Note that these two formulas are identical under the following substitutions:

$$\sigma_a = \frac{1}{2}\sigma_s, \quad \gamma_a = \frac{1}{2}\gamma_s$$

Thus, slot solutions are good for both, keeping in mind that for asynchronous channel use the rates are 1/2 that of the synchronous case. That is, for any given channel bandwidth and capture ratio the peak rate achievable, or the rate at a given delay (dependent on γ/σ) will be only half as good in the asynchronous case as if slots are defined. Slot padding will mean some of this advantage is lost but unless the slots become twice the packet size the slot technique is significantly better.

Other Population Densities

If the population density of stations is not uniform within a circle, the next most realistic ground based case is that of a long narrow strip of land or a circular situation with population density decreasing with radius (like most cities). Both cases result in the same distribution, so consider the long strip of land:

Probability of being at x: $p(x) = 1/A$

Fraction of stations to cause conflict: $\kappa = \dfrac{x}{A\sqrt{\beta}}$

$$q = \int_0^{A\sqrt{\beta}} \frac{1}{A} e^{-\kappa\gamma m} dx + \int_{A\sqrt{\beta}}^{A} \frac{1}{A} e^{-\gamma m} dx$$

$$q = \frac{\sqrt{\beta}}{\gamma m}(1-e^{-\gamma m}) + (1-\sqrt{\beta})e^{-\gamma m}$$

$$\sigma = \frac{\sqrt{\beta}}{m}(1-e^{-\gamma m}) + (1-\sqrt{\beta})\gamma e^{-\gamma m}$$

Therefore, the solution is identical except that $\sqrt{\beta}$ is substituted for β, improving the situation since the capture ratio required for a given performance can be twice that of the previous case.

The worst case is typified by a satellite situation where all stations are almost equidistant from each other (via the satellite). In the satellite case capture can be presumed impossible ($\beta=0$), unless capture ratios considerably better than 1.5 db can be achieved (1.5 db - the power ratio of the extreme case).

Maximum σ for Slots

Except for the unrealizable case where the capture ratio = 0 ($\beta=0$) in which the maximum $\sigma=1$ at $\gamma=\infty$, there is a true maxima for σ.

$$\sigma_{max} = \beta + (1-\beta)e^{-(\gamma_1)}$$

where: $\gamma_1 = \dfrac{1}{1-\beta}$

Delay

The delay encountered by a packet is initially the channel and packet transmission delay assuming no conflict. Call this delay A. Next, if one or more retransmissions occur, each takes some time R from transmission to subsequent retransmission. Since there are (γ/σ) total transmissions on the average, the total delay is:

$$D = A + R(\frac{\gamma}{\sigma} - 1)$$

Satellite Delay   For a satellite the initial delay, A, is the up-down satellite delay (C = .27 sec) plus the packet time, T.

$$A = C + T$$

Retransmission occurs after seeing your own down signal being in conflict, and waiting randomly over 10 packet times to avoid direct reconflict.

$$R = C + 6T$$

Thus: $D = (C + 6T)\frac{\gamma}{\sigma} - 5T$

Typical times for a 50 KB channel, 1400 bit packets are:

$$D = .438\frac{\gamma}{\sigma} - .14 \text{ sec}$$

Ground Radio Delay   Here the propagation delay, C, is much less but the same rule holds for A.

$$A = C + T$$

The retransmission can only occur after acknowledgment. However, this can always be sent within one packet time (ave ~ 1/2 T)   Two

transmission delays occur, however, and a random wait over 9 packet times is assumed.

$$R = 2C + 7T$$

Thus: $D = (2C + 7T)\frac{\gamma}{\sigma} - (C + 6T)$

Typical time for 50 KB, 1400 bit packets at distances under 100 miles are:

$$D = .196\frac{\gamma}{\sigma} - .168 \text{ sec}$$

Operating Point

When σ is pushed to the maximum for any variant, the system is dynamically unstable, as it is if γ is somehow pushed beyond the second solution of the σ,γ equation. Thus, an operating point must be chosen below the maximum σ point, which is dynamically stable and has a reasonable delay. A convenient point to pick, both for operation and for discussing the capacity of the system, is the point where $q = 1/2$ and $\gamma=2\sigma$. This means the delay is fixed if comparisons between variants are considered and that for actual use there are only two transmissions on the average. This also constrains the maximum delay which can grow quickly if $(1-q)$ gets closer to unity.

At $q = 1/2$:

$$\beta = \frac{\gamma m(.5-e^{-\gamma m})}{1-(1+\gamma m)e^{-\gamma m}} \quad , \quad \sigma = .5\gamma$$

where $m = 1$ for slots and $m = 2$ for asynchronous

	Capture Ratio	β	σ_a	σ_s
Infeasible	0	1.0	.40	.80
Good FM	1.5 db	.707	.30	.60
Mod. FM	3.0 db	.500	.25	.50
Poor FM	6.0 db	.250	.205	.41
Satellite	0	0	.173	.347

Slot Padding

For satellite operation, each station knows its delay to the satellite
and the padding can be minimal unless very high bandwidths are used.
For UHF radio with a 25 mile radius the padding is negligible at 50 KB,
10% at 500 KB, and 50% at 5 MB for 1200 bit packets. This is assuming
no delay correction computation as with satellites. Beyond 5 MB either
delay correction must be incorporated or the asynchronous system used
if 1200 bit packets are still desired. Of course, it is at this same
bandwidth that multiple receivers begin to help sort out the conflicts --
so delay correction may be unnecessary to achieve a consistently high
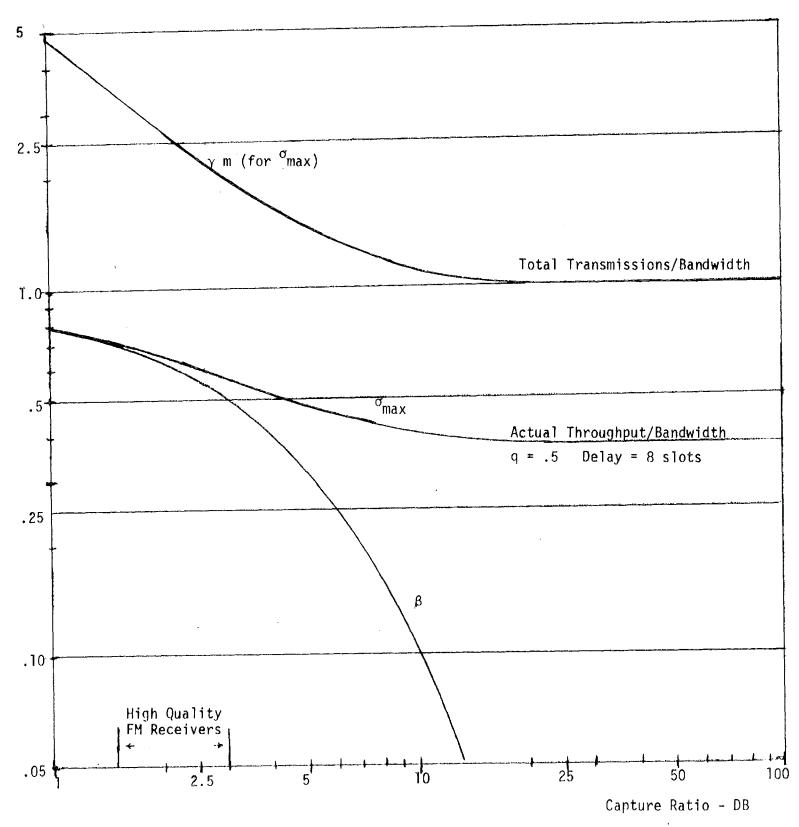channel utilization.

Conclusion

For satellite use, slots will improve the capacity by a factor of two
to .347. For ground radio, a good FM receiver would permit a further
improvement to 60% of the full bandwidth. The delay in this case is
only eight packet times. Slots clearly need not be used at very low
channel utilizations since the asynchronous system will work just as
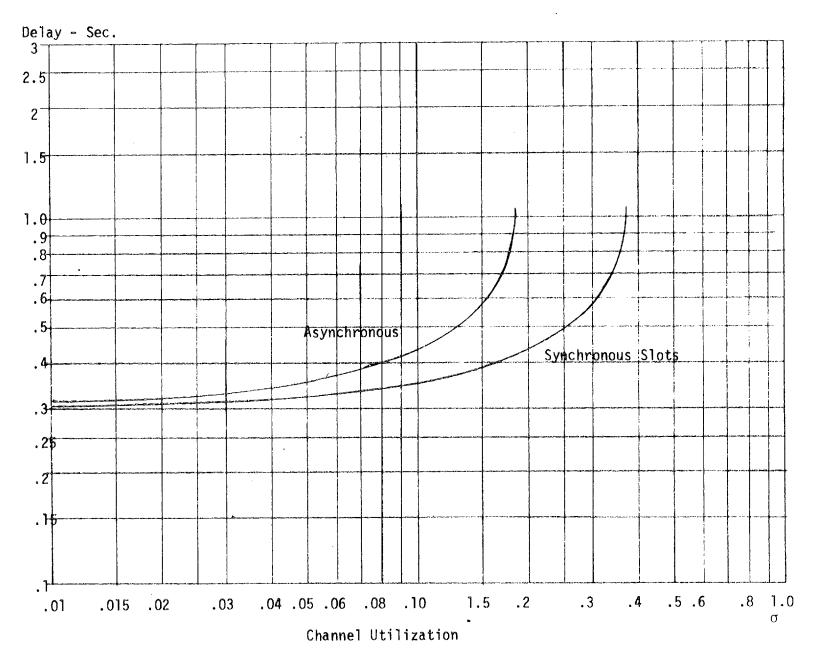well at low rates. Thus, the resynchronization of the slot clock is

not a problem. If no packets come by the clock drifts, no harm is
done -- we just have the asynchronous case.

ALOHA SLOT PACKET SYSTEM

Maximum Channel Utilization vs. FM Capture Ratio
for evenly populated area around station

Maximum Channel Utilization vs. FM Capture Ratio for evenly populated area around station
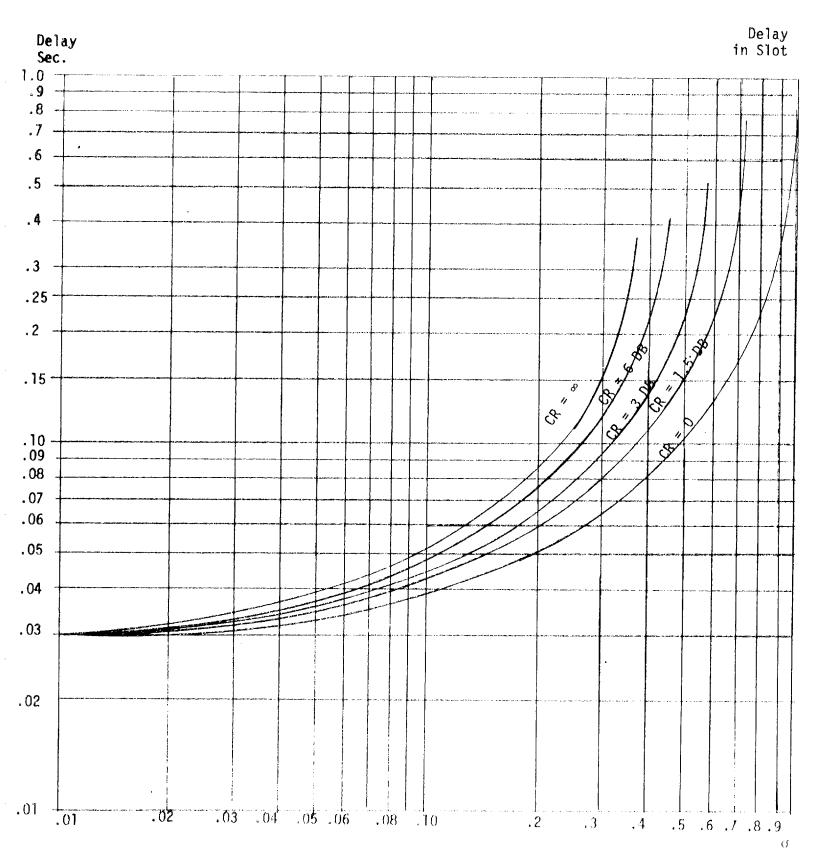
SATELLITE DELAY

Asynchronous and Slot Systems
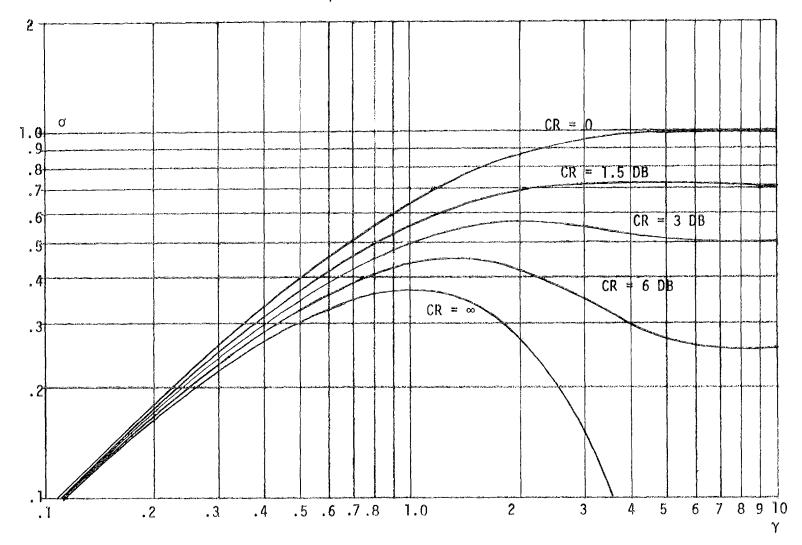## for 50 KB Using 1400 Bit Packets

ALOHA SLOT SYSTEM

Ground Use Delay for 1400 Bits/50 KB

ALOHA SLOT SYSTEM

σ vs. γ

For Capture Ratios from 0-∞

(continued from the first page of preceding paper)

Roberts in the present paper investigates methods of increasing the effective channel capacity of such a channel. One method he proposes to gain in capacity is to consider the channel to be slotted into segments of time whose duration is equal to the packet transmission time, and to require the devices to begin a packet transmission at the beginning of a time slot. Another method Roberts proposes to gain in capacity is to take advantage of the fact that even though packets from two devices collide in the channel (i.e., they are transmitted so they pass through the channel at overlapping times), it may be possible for the receiver(s) to "capture" the signal of one of the transmitters, and thus correctly receive one of the conflicting packets, if one of the transmitters has a sufficiently greater signal than the other. Roberts considers the cases of both satellite and ground radio channels.

(Some of the text for the above background material was abstracted from "On the Capacity of Slotted ALOHA Networks and Some Design Problems," Israel Gitman, IEEE Transactions on Communications, Vol. COM-23, No. 3, March 1975.)