1988 NEAR EAST UNIVERSIT



Faculty of Engineering

Computer Engineering

Dep ar tm ent

GRADUATION PROJECT

SUPERVISOR

Prof.Dr. Fahrettin M. SADIGOGL U

Prepared By: Halit AYDIN

NICOSIA 1997



NEAR EAST UNIVERSTY

GRADUATION PROJECT

COM 400

SUBJECT

PACKET RADIO NETWORK

PRESENT

To my parents

To my brothers

To ali my family

To teacher Prof.Dr. Fahrettin M.SADIGOGL U

To my friejrds

To everybody tolled me

a letter

ACKNOWLEDGEMENTS

I would like to acknowledge to my parents who supported me during my education period who have patiently encouraged me to be the best everywhere

I would like to thank my teacher Fahrettin M. SADIGOGLU who assisted me to get a full picture about, my p.r.oj ect and he has a very strong reason of understanding the topic of 'I'e lecommunications.

I specially appreciate my friends who helped me in preparing outputting the project. I also thank my home mate who provided a healthy and quite environment during my preparing the project.

I continue to thank ali people who got a right upon me.

Halit Aydin

INTRODUCTION

After of computer world discover becomes smaller and smaller everyday scientist make inventions this inventions effect the telecommunications and communications technics new data developed.

in this thesis I introduce Packet radio network This is very shortly interesting subject it is called wireless communication inother world called ALOHA. This communication are based radio waves with frequency band between 300-3000 MHz. Transmission data from removed station host computer is realised to some Frequency of Fl and opposite communication performed is by frequency F2.

EQUATIONS

 $\begin{array}{c} -O & e^{-O} \\ 1 & Pi(k)=i & e \\ 1d & \end{array}$

2 S-G11"G

3 S.m.G.
$$\prod_{i=1}^{n} (1-0j)$$

4
$$\int_{i \cdot l}^{N} \cdot ol = 1$$

5 $Sl \cdot Gl (t - ol)^{Nl}$ $-1 (t - o2)^{N2}$
6 $S2=02 (1-02)^{N2}$ $(1-01)^{N1}$

7 NI 01 + N2
$$02 \cdot 1$$

8
$$\frac{1}{0} R^{+} \frac{L^{+}}{2}$$

9 fit•Pr (nml) Pr(11110) + &(n-0) Pr(1111) + 10 throu@1?UV8NO: $(1 - \delta_{1}t - 1)$

PACKET RADIO NETWORKS

in this chapter we begin our discussion of packet broadcasting networks. Recall that such networks share the following character is tics:

* Packet transmission.

* No switching.

* Reception by many Of all stations.

* Common transm is sion medium.

packet This ch ap ter deals with radio and satellite which exhibit numb er of similarities. The n etwo rks. а character istics of local networks are quite differ ent, For both packet radio and satellite networks, we begin with a discussion of architectur al issues, and then look at medium access control protocols.

PACKET RADIO ARCHITECTURE

The architecture of packet radio networks can be classified as centralised of distributed (Eigure 1) in a centralised network, there is on central transmitter/ receiver attached to a central resource. All other nodes communicat'e only with



(b)Dl1tributed

FIGURE-1 (Basic Packet Radio Architecture)

the central node. Node-to-node communication is a dir.ect, mediated by the central node. The earliest networks followed this model, and were designed primarily to provide terminal access to a central time sharing system. in a centralised system, two radio channels are required. Individual nodes send packets to the central node on channel and the central node broadcasts packets on another. Since radio transmission is by directional, packets transmitted by the central node are heard by all the other nodes. Thus the configuration, is logically equivalent to a multi point line with a primary and a number of secondaries.

PACKET RADIO ARCHITECTURE

network is not appropriate The centralised or the more common situation today of a collection of microcomputers that wish to exchange data, messages, programs, and so forth. The distributed architecture takes full advantage of the On channel is used for all omnidirectional property of radio. transmissions and each transmission is heard by all other nodes. This configuration is logically equivalent to a local area network

discussion above line-of-sight Figure and the assume propagation, recall that the maximum distance between and receiver is slightly more than the line of sight, transmitter of 100 kilometers. Where h is the height of both or a distance in meters, and K reflects a refraction effect. antennas For with two antennas at a height of 10 m, and using a example, nominal value of K= 4/3, the maximum range is 26 km. This represent the maximum radius of a centralised system, and the max imum diameter of a distributed system.

To overcome this geographic limitation, a store-andforward repeater is used (Figure 2) A repeater performs much the same task as a node in a packet





the Distributed

Figure 2 (Packet radio Network with repeaters)

Switched network except that it works with broadcast link rather than point to- point link. in a centralised s system, the repeater accepts packets from the central node and retransmits them to remote nodes. it also accepts packets from these remote nodes and forwards them to the central node. in a distributed

system, the repeater acts as a switch between two sets of nodes, accepting packets form one set for re transmission to the other, an vice versa.

We now turn to a more detailed description of these two configurations, using ALOHANET (ABRA 70), (BIND 75 A) as our example of a centralised system, and a stan dard known as AX.25 as our example of a distributed system.

Centralised Networks: ALOHANET

The first radio network, packet ALOHANET. was developed by the University of Hawaii and became operational in 1970. Its principal objective was to allow user terminals in widely scattered locations to access the university computer system. Traffic was primarily terminal-to-host, but terminaltraffic could be routed via the central node, called to-terminal the menehune (Hawaiian for "imp"). Remote units were of two types. The terminal control unit (TCU) operated with a simple half-duplex terminal and included a buffer, control logic, and transceiver. The programmable control unit (PCU) was a microprocessor-based device for terminal concentration and/or a. computing station.

As a centralised system, ALOHANET requires two channels. PCU- and TCU-to- menehune traffic are on channel \mathbf{f}_1 us.ing a frequency of 407.35 Mhz traffic from the menehune carried on channel f2, at a frequency of 413.475 Ml1z. Both

channels have a bandwidth of 100 kHz and, using PSK, a data rate of 9600 bps. Transmission on both channels uses packets with the following format:

* SYNC (100 bits) : A fairly lengthy synchronisation pattern was deemed advisable to minimise errors.

* Header (32 bits): The header includes the user address (8 bits) repeater address (6 bits) packet type (3 bits), packet length (8 bits), and various other control bits.

* CRC (16 bits): The header is protected with its own errordetecting code.

* Data (640 bits): A maximum of 80 characters can be trans mitted.

* CRC (16 bits): The <lata are protected with another errordetecting code.

The fl (user) channel uses a multi access contention protocol known as ALOHA Each station transmits a packet has <lata to send. when it it then expects to hear an acknowledgment (ACK) from the menehune. However, since each station transmits at will, it is possible that two transmissions will overlap. This is known as a collision; the result is that the menehune receives a garbled transmission. To account for transmission errors and collisions, a user node retransmits a packet if no ACK is received during a random time out interval. The random time interval avoids a second

collision between packets that had originally collided. The time is uniformly distributed in an interval with a minimum of 0.2 s, chosen to allow time for receipt of the ACK, and a maximum of 1.5 s. The lower bound is increased for nodes transmitting through repeaters, to account for the repeater delay in both directions.

The f2 channel is used primarily for two types of packets: acknowledgment packets an data packets prom the ceritr.al resource. Because the timing of ACK's



Figure 3 (ALOHANET Broadcast channel multiplexing) 1s critical, acknowledgment packets have absolute priority. Tw.o queues of packets (Figure 3). As long as the ACK queue is. not empty, the next packet is transmitted from that queue.

implementation, In the original packets from the menehune were not acknowledged by the user nodes, for two First. since the f2 channel was non contention, reasons. there of successful a high probability transmission. Second, was ACKs would increase the congestion of the contention-based f1 channel. Later on, ACKs were added for selected applications such as file transfer.

The repeaters in ALOHANET also use channels fl and 2a. Each repeater has a list of addresses of nodes with which it can communicate. Packets on fl from a user node within the address range are repeated on f1 to the menehune or to the newt in the case of cascad ed repeaters. repeater Packets on f^2 addressed to user node within the address range are repeated on f2 in the opposite direction. The process is depicted in Figure 4 a which shows a repeater that communicates with a central controller (c) and a detail set of user nodes (A). Since radio transmission is omnidirectional. It is clear that the menehune and a repeater should not transmit on f2 at the same time. To avoid this, the menehune will pause after transmission of a packet to a repeater long enough for the packet to be forwarded.

Two more elements of ALOHANET that need to be described are the strategies for routing and flow control. Routing is required if a packet-radio network has more than one repeater. In the case of ALOHANET, a fixed routing strategy is used. The system is set up so that the ranges of addresses for

the various repeaters do not overlap. in a system with many repeaters. A more complex routing strategy would be indicated.

Flow control on f2 requires that the menhune know the input characteristics of the user node. in essence, the menehune waits sufficient. time for a node to absorb a packet of <lata before sending another. Flow control on f1 is normally not a problem for lenghtly transfers from a PCU, ago-ahead packet mechanism is used. The PCU can only send a certain amount of <lata and then has to wait for the go-ahead packet.



Figure 4 (Function of a radio packet repeater)

D is tri bu ted N etwork: The AX. 25 S tandard

An increasingly common application of distributed packet radio is to provide distributed networking among personal computers, including access to centralised computing resources. In most cases these networks consist of amateur radio stations and are open to any user who conforms to the protocol used on a particular network. And are open to any user who conforms to the protocol used on a particular network.

Amateur packet radio networks exist in a number of areas throughout North America, and the number is growing/steadi1y. An effort to standardise these networks has been un.def Way since 1982, under the sponsorship of the American radio Re18.y League (KARN85), (BRUN84), which has produced a standard for a link level protocol suitable for packet-radio networJ.<:s known as AX.25 (ARRL84)

network, drstr ibute An ARRL- type network is а organised into clusters of stations connected buy repeaters. All stations an repeaters share a single frequency for trai:isin.issiôi:i deception. The AX.25 standard does not and specify frequency to be used. Based on FCC-approved ch8.n1:iel availabili ty, the typical network uses the 220-Mhz bend, lislrig FSK and with a bound width of 20 kHz or 100 Khz. A typid8.1 data rate is 4800 bps. As with ALOHANET, a fixed routing scherne is used. In this case, the route to be followed 1S specified by the source station, as explained below.

The link protocol is based on , and very close to, HDEC. The frame format is as follows.

* Flag (1 octet): as in HDLC, 01111110.

* Address (14 to 70octets): explained below.

PACKET RADIO ARCHITECTURE

* Corrtrol: as In HDLC.

* Protocol Identifier (1 octet): specifies what kind of network layer protocol, if any, is in use.

*Information :as in HDLC.

* Frame Check Sequence (2 octets): as in HDLC.

* lag (1 octet): as in HDLC

the AX. are two differences between 25 There frame format and that of HDLC: the protocol Identifier (PID) field and the address field. The PID field is used to designate the layer 3 protocol that is using the AX.25 link protocol. This would allow multiple users of the link layer protocol. For ex amp Ie, ARRL is in the process of specifying a network-layer protocol tailored to packet-radio networks. Another alternative is an Internet protocol. Typically, the layer 3 protocol with in systems on a network is unique and of no concern to the link layer. Thus, at present ,the utility of this field is doubtful.

The most important difference between AX.25 and HDLC 1s in the addressing technique. In HDLC, there are two possible configurations: a point-to-point link with two stations, and a multi droplink with one premary and multiple secondaries. In either case, a single address is sufficient for the operation of

the protocol. This is not true in a packet-radio network, for two reasons:

1- Since the network is a peer, distributed network, both the source and destination stations should be identified; neither is unique. For flow control, error control. and sequence nurnbering, both addresses are needed. To see this, consider a situation in which station A is simultaneously engaged a logical to two other stations, B and C. A is excharging connections AH.25 frames with both B and C, using the HDLC mechatlism. is for flow control and error control. Thus, A rnust keep trac of the send and receive sequence numbers used for its separate connections to both A and B. To da this, each incorning frame rnust identify the sender.

2. If repeaters are involved, these repeaters rnust be specified. In particular, it is the responsibility of the transmitting station to specify the repeater or repeaters that rnust be used to get from source to do destination.

The AX. 25 address field is from 14 to 70 octets long, depending on whether and how many repeaters are used between If the sending source-destination pair. and a particular deceiving stations are in the same cluster (within range of each other), then it is only necessary to specify the source and station addresses. Each is specified using 7 octets. destination Which contains a call sign of up to 7 characters. IF c frame is to go through a repeater, an additional 7-otet address supplied is appended to the end of the address field. This field contains

the call sing of the repeater. It also contains two flag bits of interest: The H (has-been-repeated) bit and th e addressextension bit. If a station wisher to send o frame to a station that con be reached only through a repeater, it includes not only its address and the destination address, but the repeater addre ss as well. The h bit is set to zero. repeater **1gnore** If its address all frames that do not contain its address and retransmits the present the repeater sets the H bit packet. A station receiving a frame with a repeater address whose H bit is set to zero will ignore it. This ds the problem of a station receiving duplicate packets, one from the source and one from the repeater. This is possible since line-ofsight transmission radii of necessity overlap. Figure 4b shows the use of a repe'arer to lirik two Sefir of uSer staliôriS.



Figure 5 (Packet Radio Network with many repeaters)

* Variable data rate links: Data link error dete eti control techniques contribute to transmiss ion overhead an.ô. reduce the effective data rate. Some radio designs allow these techniques, and thus the effective data rate, to significantly in response to link quality variations. This affect the selection of minimum delay routes. * One-way link: This would result from different ambient noise levels, jamming, , or different antenna characteristics at two radios. Most distributed routing techniques require information from any node to which a packet might be sent, but it is unclear how a sender can even learn the existence of an outgoing one-way link.

Common channel effects are those produced by the facr that all link use the same shared transmission channel. In figure 5 for example, stations G. K, and L are in range of each ofhief. Only one of the three may trans mit at a time. A transmissiOn. from one is heard by the other two as well as any othet repeaters within range of the transmitter (e.g. G's transmissiori. is heard buy K and L. And also by, 13,...G, implications of .thi.s situation are:

more The AX.25 protocol allows operation throu Up to one repeater, creating a primitive routing mechanism. address eight repeaters may be used by extending the repeaterad dress but the last will have the field. Each address sub extension bit set to zero. The first repeater address designates the first repeater in the chain. As a frame progresses thr oüğb a H bit chain of repeaters, each successive repeater will set the address. indicting that the frame ın its own has successfully repeated through it.

Whether one or multiple repeaters are employed particular transmission, it is the responsibility of the source

station to design the route and specify the repeater addresses in the frame.

The operation f theAX.25 protocol is essentially the same as that of HDLC. Functions such as flow control, error control, link establishment, and so forth are Ident ic.al. One additional requirement in a packet-radio environment .is the use <Of u medium access control technique, several which described below (ALO HA, P. ALO HA S- ALO HA)

Routing

In packet radio networks with а small number a fix'ed rout.ing scheme is adequate. As these of repeaters. packet radio grcws, we are more likely to see networks with a large numb ert 10) of repeaters and with at 1 some of the stations being mobile. Military requirements clearly reflect this architecture. Figure 5 illustrates the types of configurati ons of In a centralised networks. In a multirepeater net works network, the routing problem is to find a route centralised between the central node F and all work, the routing probl em is to find a route between the central node F and all other nodes. In a distributed network, we have the apparently more difficult problem of finding a route between each pair of repeater s. Note also that in a distributed network, one would need some sort of address for each station of the network, two-level would need some sort of two-level address for each station of the form

(Repeater, station), where "repeater" identifies the cluster of stations local to a particular repeater. This section present an overview of the routing problem for such networks an is based primarily on (CAR8 l)

Each repeater is equivalent to a node, and the local cluster of each repeater is equivalent to multiple stations attached to a of the similarities, one might think node. Because that. the discussed previously would be applicable to routing algorithms multirepeater packet radio networks. However, there are differences. due to the transmission characteristics netwo rks. and the fact that some packet radio nodes may be mobile. the two networks, and the fact that some packet radio nod may be mobile. We can group those differences between the types of networks that affect routing strategy into categories: Link reliability affect and common channel effects.

Link reliability effect are those due to the fact that r links are less reliable than guided-medium link. They subject to fading, multipath and noise interference and, in hotile environment, to jamming. Lower reliability suggests:

*More frequent monitoring of link status: This is required to assure reliable communication. If overhead packets for purpose are exchanged on a regular hasis, routing inform ation could be added. (Piggybacking).

*Variable data rate links: Data link error detection control techniques contribute to transmission overhead an

Some radio designs reduce the effective <lata rate. allow these techniques, and thus the effective <lata rate. to vary significantly in response to link quality variations. This affects the sellection of minimum delay routes.

These would result from-different One-Way link: j amrn in g different antenna ambient levels. or noise routing characteristics radios. Most distributed at two techniques require information from any node to which a packet might be sent, but it is unclear how a sender can even learn existence of an outgoing one-way link.

Common channel effects are those produced by. The that all link use the same shared transmis sion channel. In Figure 5, for example, stations G, K, arrd L are in range of each other. Two as well as any other repeaters within range of the transmitter. (E.g. G's transmission is heard by K and also by B, C, and F). The implications of this situation are.

* Link delays are the same for all output link: With p each node has a queue of packets for each link and knows length of its queues. With broadcast lines, these local view packets waiting to use the common channel. This will affect routing strategies that use queuing delay as a parameter.

* Link are not independent: The amount of traffi c(and hence delay) between one pair of repeaters will affect the delay one other pair wise links. This is not taken into acco

most algorithms for globally generating all routes through the network.

* Routing overhead packets need be sent only once: Many dtstribute d routing strategies require that routing information be sent by a node to each of its neighbours. This would require only one radio broadcast.

It should be clear that the problem of routing in ra packet radio network is even more complex than in a conventional packet-switched network. As yet, there has been little experience or research in this area. As examples, we mention two a:pproaches.

Because of the problems of unreliable links and mobile nodes, it is clear that a highly roust routing algorithm 18 Flooding comes to mind, but is the most wasteful of required. bound width. However, from the point of view of the individual flooding is required! Even if a packet transmitted repeater, bv a destination repeater includes repeater address field. Th packet will be received by all other repeaters within range. On way to reduce waste with a flooding approach would be to to minimise the number of repeaters that actually do retran When a repeater (LIU80) The procedure works as follows. receives a packet to broadcast, it waits a random length of before doing so. If it receives another copy of the packet before its own broadcast, only one of them will relay it. This reduces 5, for but does not eliminate packet duplication. In Figure example, if repeater H broad cast a packet, both C and

rece ive it. Since C and L are not within range of each other, both will rebroadcast it.

Another way to reduce duplicate packets is for each repeater to only forward a packet when it is closer to the destination than the repeater from which it received.

The packet (GITM 76). Suppose that each repeater knew ist hop-count distance from every other repeater Then process would work as follows. When a repeater receives a packet form a station for forwarding, it adds a distanc.e whose value is the number of hops to the destination repeater, and then broadcasts the packet. When a repeater receives packet from another repeater, it does the following:

^{*} If the pack et is for a station local cluster, broadcast it.

* If the packet is to be forwarded, check the distance field. If this repeater is closer to the destination than repeater, update the distance field an transmit.

in Figure 5b, for example, suppose that K broad cast a packet with a destination of D and a distance of 3. Repeaters J,F,G and L receive the transmission, but only J and F have a hop count to D of less than 3. They both broadcast a packet with a distance of 2. Repeaters I, E, A, B, G, K, and J get one or two copies of the packet. Of these only I, E and A are closer to D and they all send a packet to D.

This seems to work fine. The question is how each repeater may determine its ho count to every other repeater. This could be done in several ways. Each station could maintain a distance table which is periodically exchanged with its neighbours. much like ONA and the ARPANET original Another alternative is backward algorithm. а learning algorithm. When a repeater receives a packet for the first me that packet will have traversed the shortest route fto1.n the original repeater. This information could be used to update the distance table in the destination repeater.

Pure ALOHA.andSlotte.d ALOHA

early 1970s. Nôrfüan Abramson (1970,1973a, in the 1973b, 1977) and his colleagues at the University of Hawaii devised a new and elegant method to solve this problem. Their work has been extended by many researchers since then (Binder et al., 1975; Carleial and Hellman, 1975; Ferguso n, 1975b; Lam, 1974; and Roberts, 1973, to name just a few). Although Abramson' s work, called the ALO HA system, used ground based radio packet broadcasting rather than satellite packet the basic idea is applicable to any system in broadcasting, which uncoordinated users are competing for the of a single shared channel. N evertheless, imp ortant there are some ground radio differences between packet broadcasting radio broadcasting (notably the prop agation satellite packet de Iay). We will examine ground radio in general, and the

University of Hawaii ALOHA system in particular, later in this chapter.

'The basic idea of an ALOHA system is simple: Just let the users transmit whenever they have <lata to be sent. There will be collisions, of course, and the colliding packets will be destroyed. However, due to the perfect feedback property of packet broadcasting, the sender of a packet can always find out whether.

Or not his packet was destroyed by just listening downward rain of packets one round- trip time after sending the packet. If the packet was destroyed, the sender just waits a ran dom amount of time and sends it again. The waiting time must be random or the same packets will collide over and over. in locked. Systems in which multiple users share a common channel in a way that can lead to conflicts are widely known as contention systems.

A sketch of packet generation in an ALOHA system is given in Fig.6. We have made the packets all the same length because it has been shown that the throughput of ALOHA systems is maximised by having a uniform packet size rather than allowing variable length packets (Abramson, 1977; ferguson, 1975a; Gaardar; 1972).



Figure 6 (in pure ALOHA Packets are trans mitted at completely arbitrary times)

Whenever two packets try to 0 ccüp'y channel at the same time there will be a collision and bot will be garbled. You should realise that if the first bit of a new packet ov erlaps with just the last bit of a packet almost finished, both packets will be totally destroyed, and both will have to be re transmitted later. The check sum cannot (and should not) disti nguish between a total loss and a near miss. Bad is bad.

A most interesting question is: What is the throu ghput of an ALOHA channel? Let us first consider an infinite collection of interactive users sitting at their terminals. A user is always in one of twostates: thinking or blocked. Initially, all u^{sers are} in the thinking state. Whenever someone decides what to do next, he types a line of text followed by a carriage r^{et}urn. At

this point he is blocked and stops thinking. The microcomputer inside the terminal immediately locks the keyboard to prevent any more input. it then sends a packet containing the line to the satellite and waits R sec to see if it was successful. If so the user7s keyboard is unlocked. If not, the keyboard remains locked, and the packet is retransmitted over and over until it is successfully sent.

Let the "packet time" denote the amount of tim.e needed to transmit the standard, fixed- length packet (i.e., the packet length divided by the bit rate). At this point we assume that the infinite-population of users generates new packets according to a poison distribution with mean S packets per packet time. infinite- population assumption is needed to ensure that S does not decrease as users become blocked.) If S 1, community is generating packets at a higher rate than channel can handle, and nearly every packet packets at a hi rate than the channel can handle, and nearly every packet suffer a collision. For reasonable throughput we would expe S 1.

in addition to the new packets, the users also generate transmissions of packets that previously suffered collis Let us further assume that the probability of k transmis attempts per packet time, old and new combined, 1s Poisson, with mean G per packet time. Clearly, G S. At load (i.e., S O), there will be few collisions, hene retransmissions, so G S. At high load there will be many

collisions, so G S. Under ali loads, the throughput is just the offered load, G times the probability of a transmission being successful that is, $S = O_0$ where Po is the probability that a packet does not suffer a collision.

A pack et will not suffer a collision if no other packets are scnt within one packet time of its start, as shown in Fig7. Under wliat conditions will the shaded packet arrive und amaged tO andtO+l IOF any other user has generated a packet between the end of that packet will collide with the beginning of shaded one, in fact, the shaded packet's fate was already se the long even before the first bit was sent. But due to propagation delay, it has no way of knowing that anoth er packet under way. similarly, any other packet started was already of the shaded between to+a and ao+2. t will bum p in to the packct



Fif~. 7 Vulnerable perind for the 'illaded packet.

The probability that k packets are generated during a given packet time is given by the Poisson distribution:

So the probability of zero packets is just e In aninterval two packet times long, the mean number of packets generated is 2G The probability of no other traffic being in itiated during the entire vulnerable period is thus given by

This result was first derived by Abromson (1970)

offered traffic relation is shown in Fig. The throughput-6-3. The maximum throughput occurs at G = -.5, with S - 1/(2e), which is about 0, 184. In other words, the best we can hope for channel utilisation of 18%. This not very is а result but with everyone transmitting whenever he wants encouraging, to, we could hardly have expected a 100% success rate.



Figure 8 (Throughput versus offered traffic for ALOHA system)

published a method for doubling 1972, Roberts in the of an ALOHA system. His proposal was to divitle time capacity each interval corresponding up into discrete intervals. to one üne way to achieve synchronisation packet. among the users would be to have the satelli te emit a pip at the start of each interval. like a clock. Although the pips would arrive down at 270 ms later, each user would receive the earth the signal at about the same time. By making the time slots slightly larg er time with than the packet times, the variation in propagation position on the earth could be compensated for.

in Roberts method, which has come to be known as slo AL OHA, in contrast to Abr.arn son's pure AL OHA, a terminal is to send .whemever a >.c arriage return is typed. not permitted it is required to wait for the beginning of the next slot Instead, Thus the continuous pure ALOHA is turned into a discre the Vulnerable period is now reduced in half. Since of no other traffic during the same slot as probability packet is e -G which leads to

$S = Ge^{-G}$

As you can se from Fig. 8, slotted ALOHA peaks with a throughput of S= 1/e or about 0,368, twice that of ALOHA. If the system is operating at G= 1, the probability of an empty slot is 0.368 (from Eg. 6-2). The best we can hope for using slotted ALOHA is 37% of the slots empty, 37% successes, and 26% collisions. Operating at higher values of G reduces the number of empties but increases the number of collisions

exponentially. To see how this rapid growth of collisions with G comes about, consider the To see how this rapid growth of collisions with G comes about, consider the transmission ofca test packet. The probability that it will avoid a collision is e -G. The probability of a transmission requiring exactly k attempt (i,e., k-1 collisions followed by one success)

Pk=
$$e -G (1-e-G)$$
 ky

The expected number of transmission per carriage return typed, E, is then

$$E = \sum kPk = \sum ke^{-G} (1 - e^{-G})^{k-1} = e^{G}$$

As a result of the exponential dependence of E upon smal increases in the channel load can drastically reduce its performance

Finite Population ALOHA

The above result is have been obtained usi assumption of an infinite number of users. Abramson (1973a) allows analysed Slotted ALOHA systems with a finite number of users. We now briefly summarise his results.

Let Si be the probability of a successful tran smission generated by user i. Remember that at equilibr ium the throughput rate must equal the rate at which new packet are generated. Let Gi be the total transmission probability (per slot) of user i, including both new packets and retrans missions Clearly The probability that a given slot will contain a successful packet sent by user i is the probability that user i sends a packet, times the probability that none of the n-1 other users sends a packet:

Let us now specialise Eq.3 to the case of n identical users, each hang a through put of Si= S/N packets/slot and a total transmission rate of Gj= G/N packets/ slot, where $G=\Sigma G_i$. Substituting into Eq. (3), we get

Form Fi.8 we see that the maximum throughput for an infinite population slotted ALOHA system occurs at G=1. Abramson (1973a) has shown that this intuitively reaso nable result also holds for systems with a finite number of users. The condition for maxim'um throughput is

Now let us consider two classes of users, for example file transfer users and interactive users. Let there be N_1 of first kind an N_2 of the second, with throughput S_1 and respectively (per user). Then Eq.(3) reduces to

 $S_1 = 0.1^{\circ} 1 - G_1 N_{1-1} (1-G_2)N_2$

For maximum throughput we must obey the constraint Eg.(4), which becomes

We now have three equations in four unkno (S_1,S_2,G_1,G_2) . By using Eq.(6), we can eliminate g2 from two class throughput equations Eq's(5) and (6) to yie parametric equations for the throughput, parameterised by G¹.

As a first example, consider N = 1 and N = 1 And N = 1. This leads to SI=G and are plotted in Fig.9(a). No tic e that when gl is close to O, user 1 hardly ever attempts to send, and user 2 is free to use nearly every slot, so the total throughput is close to one packet per slot. The worst case is attempt to send on every slöt with when both users a probability of 0.5. If that happens there is a 25% chance that there is a 25% user 1 will try and user 2 will refrain. Similarly, The total chance that user 2 will try and user 1 will refrain. to be throughput is therefore 0.5 packet per slot. The collusion yields drawn form this example is that an asymmetric situation a higher throughput than does a symmetric one.

Next let us specialise Eq's.(5) and (6) to the case and

User 1 might be trying to transfer a large file, whe reas the remaining users are doing interactive work. To keep total traffic finite, we must let in such a way remains finite. Letting we get

The condition for maximum total throughput 1s G1+G2=1, which allows us to plot s 1, s and the throughput, sl+s, as a function of G. This plot is give n in Fig9.(b). When G is small, the interactive users are not doing much, and the single " large" user can continuously send packets without collusions, thus achieving


Figure 9 (a- Throughput of a two user ALOHA system. b- Throughput of an ALOHA system with one large and many small users.)

a very high total channel utilisation. As G increas s, the interactive traffic claims a larger partition of the available bandwidth and the large user is forced to send less to maintain the total offered traffic at one packet per slot.

Delay and Throughput of Slotted ALOHA

The throughput is not the only parameter we are inter ested in. The mean packet delay is also important, especially interactive users. Just as with store-and-forward network shall discover that high throughput and low delay are inhe rently

in conflict. Good performance on one of them can be achieved only at the expense of other.

networks, the In store-and-forward delay comes from queuing within the IMPs. In ALOHA, the delay comes from forcing some packets to be retransmitted collisions, over and over. Due to the long round-trip prnpagattöri delay. R. each introduces a delay of at least because it takes that collision long to determine whether or not the packet transmission was If a packet collides three times, the delay between successful. first transmission and successful delivery will be at least 4R. Depending on how long the wait between a collision and the retransmission is, the delay may be considerably more.

To study the distribution of delay tim.es, we will adopt the following model. N identical users each generate new packets according to a poison distribution with a mean packets per slot when not blocked. Retransmissions are not counted i p. If the mean think time is T1 and the slot time is T2. Then Alternatively, p can be thought of as the probability of a given (unblocked) user generating a packet in a given slot. Since the total traffic, new packet in a given slot. In the total traffic, new packets plus retransmissions, must be one packet/sl of for maximum throughput, he total rate at which new packet is are introduced into the system must be less than 1 {)

When a user types a line, he is blocked and steps thinking, as in our previous model. We could allow the terminal to buffer several packets in order to allow the user to continue

imm ediat cly, but this complicates the analysis. Furthermore, it affect the system stability adversely as we shall see later. Our goal is to compute the mean time between a user typing a carriage return and the packet being correctly received by the remote computer. We have used the model of people sitting at than having all nodes be computers) for its terminals (rather value, but the an thropomorphic analysis 1S iuenricai tor computer to computer communication, of course. However, for the sake of generality, we will use the neutral term "st ations" from now on.

A key the model is how the station parameter in randomises its wai tin g times before attempting retransmission. In slotted ALOHA the waiting time consists an integral number of slots to be skipped before trying again. the mean number of slots skipped is short, the chance for same collision occurring again is large. For example, if collide and each waits either zero or one slot w stations equal probability, the chance for an identical collision second time is uO.5. On the other hand, if the retransmiss ore spread out uniformly over the next 100 slots, the chanc the same packets colliding again is 0.001. Of course, the mean delay in the latter case will be much greater than in the former. This is the heart of the throughput- delay trade-off for ALOHA systems.

A reasonable algorithm to use deciding how many slots to skip is to redistribute packets uniformly into the next L slots.

Allowing for a round-trip propagation delay of r slots, the number of slots between the first and second transmission may be R+l,R+2, , R+L, each with probability 1/L.

this randomisation Unfortunately, strategy is difficult to on account of the satellite propagati011 model. largely delay. model which a probabilistic Instead, we will adopt logged" packet hence previously collided packet (called a "b worth) is retransmitted with probability in all slots following the original transmission unit it is sent. in this model the mean distrib uted, with delay before retransmition si geometrically The mean the probability of a k slot delay given by de lay before retransmission is.

Notic e that are incllidirig the retransmission slot itself in the delay. Also Yeni.ember that this delay the time the station is blocked, because a packet may have to be retransmitted many times and this delay corresponds only to the interval between consecutive retransmissions; it does int take into account how many such inter calls are needed.

lam (1974) has shown that the final Using simulation, assumed be tween sensitive mean delay result is to the attempts but not to the shape of the curve. By retransmission equating the mean delay in the probabilistic model to the mean delay in reality, we can use the probabilistic model, dete rmine value of the optimal value for, and later deduce the appropriate L to use in reality.

The of the ALOHA system state can be completely by telling how many stations are blocked. In sate k, described there are k packets backlogged. Each of the backlogged stations may decide to retransmit its one and only backlogged packet, with the probability of transmission being and the probability of skipping the current slot being 1- In addition to the packets N- k mean retransmission traffic of k new packets are busy generating at a unblocked stations collective rate of (N-k) packets per slot.

The sate of the system varies from slot to slot as new packets become backlogged and as backlogged packets are finally transmitted successfully. Unlike our earlier derivation of the M/M/1 queuing system, this is not a birth- death process, because state changes are not always to adjacent states. For example, three new packets could become backlogged during one slot. Figure 10 shows the allowed transiti ons for a three station system. Although the state may increase by more than one during a given slot, decreases are always in units of one, since the backlog can be reduced by at mots one during a single slot. Interestingly enough, the transition from 0 to 1 is impossible, because if there is no backlog and exactly one stat ion decides to transmit, the transmission will always be successful.



An ALOHA system that moves around among 'a finite number of discrete. States in discrete fitne Steps canibe môdeTled using a Markov process (Carleial an Hellm'an, 1975).

To analyse the behaviour of a Markov process, we need to calculate the probability, Pij , that a system in state i moves to state j in the next time slot. For convenience, we introduce that n represents total number of new packets the notation generated by the N-i unblocked stations during the current slot, represents the number of Sim'iIarly, r o::;n:s;N .. **i**. whit retransmissions attempted by the i backlogged stations during with o::;r:s;i, Thus Pr [n=O] is the interval. the same time duri ng the probability that no new packets are generated current slot, and Pr [rz1] is the probability that one or more of the backlogged stations attempt a retransmission during current slot. With this notation, the transition probabilitie s can be written as in Fig. (a). The event probabilities are give Fig. (b) both for case of finite N and p and for the limit

 $N \rightarrow \infty$ p 7 0 , Np 7 S.

As an example, consider a slotted ALOHA syst three users, p=0.1, and a = 0.2. The transition probab ilities are given in Fig. (a). Notice that each row sums to 1.000 because

the probability that the system moves from its current state to *some* state is unity.

Initially, the system starts in state 0, with all stations unblocked . As time goes on, it approaches equilibrium. This does not mean that it is always in the same state', only that the probability of firidiring the system in does not depend time (as it does in the early history) To solve for the equilibrium probability , *ek*, of firiding the system in state k, we must solve the simultaneous linear equations

$$ek = : \prod_{i=0}^{N} |Pik(k) = 0, ..., N|$$

subject to the constraint that |, ek = 1. Consider of the equations for N=3:

$$e_0 = e_0 P_{00} + e_1 P_{10} + e_2 P_{20} + e_3 P_{30}$$

Each term on the right-hand side corresponds to th probability of making the transition to state 0 from one of the states (including 0). The N+1 equations are not independent, which why the additional constraint of requiring the ptobabilities to sum to unity is needed.

The equilibrium state probabilities can also be calculated in another way. If we take the matrix Pij and multiply it by self, we get a new matrix, Pij, (Z) which represents the two-step transition probabilities. For example, p03⁽²⁾ is the probability that an ALOHA system that is in state 0 during a certain slot will be in state 3 two slots later. By multiplying Pij by it self rep_eatedly, we can get the n-step transition probabilities, each matrix multiplication adding one new step. For a well behaved Markov process such as ours, successively higher powers of the transition matrix rapi'dly approach Pij("") which gives the probability of being in state j given that the system was in state 1 long ago. (Figure 11 at -last .page) ObviO-u.sly this 1s independent of i, so all the elements in a column of I'ij (00) are identical. in other words, all the rows are the, same, one gives the equilibrium state probabilities.

Given the equilibrium state probabilities, we can now the mean backlog (i.e., the expected number of stations blo waiting for retrati.smission):

mean backlog"
$$\int_{k=0}^{k} kek$$

To find the mean through put of the system, we need find the throughput for each of the possible states of the system, and then weight them by the equilibrium state probabilities. The throughput in state k is just the probability that a packet is successfully sent given that k stations are blocked and N-k

		Nevu state				At ec:iuilibrium				
		0	1	2	3		0	1	2	3
Origirıal state	0	0,972	0.000	0.027	0,001		0.760	0.131	0.095	0.014
	1	0,162	0.792	0.036	0.010			ihroughput		
	2	0.000	0.286	0.676	0.036		0	1	2	3
	3	0.000	0.000	0.384	0.616		0.243	l 0.306	U.~52	0.3841
		(a)				{d}				

Fig.12. { A three user slot ALOHA system with p=0.1

oc=0.2 (a) State transition probabilities. (b) Equilibrium values of state occupancy and throughput .) are unblocked. A successful tran sm is s ion can happen in one of two ways:

1. One unblocked station sends a new pack et; no retransmi s si 011s.

2.No new packets; exactly one retransmission attemp The probability of this occurring is just

$$f_{k} = \Pr[n=1] \Pr[r=0] + \Pr[n=0] \Pr[r=1]$$

The mean throughput can now be found:

mean throughput =
$$\lim_{k=0}^{N} \mathcal{C}_{k} f_{k}$$

The equilibrium state probabilities and through puts for our three station example are shown in Fig.12 (b)

We now have enough information to find the mean delay as well it is just the mean backlog divided by the mean throughput 1.e

$$\sum_{\substack{k = 0 \\ k = 0}}^{N} \frac{kek}{f_{k}e_{k}}$$

This is essentially Little's result in disguise. Suppose, example, that the mean backlog is eight packets and the 111ean throughput is 0.25 packet/slot. If the stations queued up nicely, it would take 32 slots for a newly blocked station to work its way to the front of the queue. Nevertheless, the system can be seen as a single-server queue. Nevertheless, the system can be seen as a single-server queue in which the queu.ing cfisciplihe is not first come, first served. Instead, a random cüstôl:n.er is plucked from the middle of the queue each time and serveti:

F}gure 13 shows the throughput as a function of G, the del ay as a function of G, and finally the throughput-delay trade-off, using the Markov model, What is important to note is that delay is small as long as G is small. However, as soon as the system begins operating at too high a G value, the delay skyrockets due to collisions, and the throughput falls back. Notice the resemblance of Fig. 13 (a) to the infinite-population model of Fig 13, as well as the resemblance of Fig. 13 (b) to our earlier result E=e0 (also for infinite population).



Fig.13. (Throughput and delay cômputed Markov model with N=25 and oc =0.1.)

Stability of Slotted ALOHA

Some ALOHA system are inherently stable atid others fluctuations inherently unstable. Due to random statistic of unblocked once in a while an unusually large number stations will attempt to send during the same slot. There will be a collision and all the packets will become backfogged. Due to the large number of backlogged packets, thefe <Will be many collisions in the succeeding slots and the throughput will be less than normal. As new packets are generated, most of them also become backlogged . Eventually, all N stations become backlogged, and from Eq. (9) we get

throughput =Noc(l-oc)N-l

Equation (9) suggests that by making or small, we can achieve non negligible throughput even when the system is badly backlogged. This observation is true, but the price we must pay is long delay times. By Eq. (8) a small value of oc corresponds to a large value of L. In other words, the random time waited between retransmission rises quickly with If 1000 backlogged stations spread their decreasing 00 ofa out over a million slots. retransmissions the eh collision will be negligible, but the price paid is terrible response.

To investigate the stability of slotted ALOHA, we n relation between mean backlog and mean thr.oughput. Typically p is a given quantity i.e. the slot time divided he mean time eg the user population Using Eq, (9) it is think straightforward to calculate the throughput as a function of the backlog has reached 60 packets, at oc=O.1 we are attempting to retransmit 6.0 packets/slot, not to mention any new packets. Thus it is not surprising that the throughput is low. For stations and oc=0.22, backlogged the retransmitted pac represent a load of only 1.2 packets/slot, which is only slightly higher than the optimal value of G.

As more stations become backlogged, the rate of packet generation drops linearly. With k stations backlo gged, the input rate is (N-k) p. The input rate as a function of backlog is shown in Fig. 14 (b) for the case of p=0.002 N= 100. Such a plot is referred to as a load line (Kleinrock

Lam,

1,975). The slope of the load line is-p; its x intercept 1s



Fig. 14 ((a) Throughput as " function of backlog. (b) Input as a function of backlog.)

At equilibrium, the throughput rate 'mu equal the input rate. By drawing the throughput curve and the lo on the same graph, we can find their intersections, as shown in Fig 15 . For 'cases are depicted in Fig 15. in the first case, N is small the system is lightly loaded, and equilibrium is achie'ved whit a low backlog. So little traffic is offered that packets nearly always get through the first time and no backlog builds up.

Not the Northeless, it is important to realise what happens if sudden fluctuation in input moves the system to a high backlog, as shown by the dotted vertical



Fig. 15 (At equili brium, in put rate equals throughputStable low delay (b) Bistable. (c) Stable, high delay.(d)Unstable.)

line. The throughput is rrow at point A and the input is at point B. Since A is hfghet than B, then throughpufe:x.ceeds the input area, which drives the backlog down 10 the eqififibrfurn point. Any excursion above equilibrium produced a throughput higher than the rate at which new packets are generated. Similarly, an to the left drops the throughput excursion so sharply that the backlog builds up again. No matter what the backlog is, there driving system inexorably forces the back to the are equilibrium point. If the load line crosses the throughput curve in exactly one place, the system is globally stable and will always return to the equilibrium point after any excursion, no matter how big it may be. Lam (1974) has shown that the values of backlog and throughput at the equilibrium point are excellent approximations to the true mean backlog and mean

throughput calculated by forming a weighted sum over all possible system states.

Now let us be greedy and increase the number of users, so that the load line moves up to the position of Fig. 15 (b). The curv e' in three plac e's. At all load line intersects the throughput three points the inipufnule equals the throughput rate, so each a possible equilibdum po.int; Hqvvevef, i911;lyJwo one represents of them are stable. First consider the low backlog ...equiJib}iµ.m point. For small excursions either way, the situation of iFig, 15 (a) prevails and the system is stable. At the middle equilibrium point, the situation is exactly reversed. If the backlog increases the throughput drops faster than the input rate, so momentarily, the backlog becomes even larger. The situation feeds upon itself until the system reaches the high back log equilibrium points correspond to nearly no one backlogged and nearly everyone backlogged, respectively.

Be gin even more greedy, as in Fig. 15 (c), is like killing the goose that laid the golden egg. The channel is now completely overloaded. The only equilibrium point is when the input rate and throughput are both very close to zero. When there are an infinite number of users, the input rate does not drop as the backlog state, the input rate will continue to exceed the throughput rate and the backlog will grow without bound.

Carleial and Hellman (1975) have expressed the stability problem in a slightly different form. They pointed out that the mean drift of the system in state i is.

Г... ^(j-i)Р _{ії}

States with drift of 0 are equilibrium points. If the slope of the drift curve at the equilibrium point is negative, 'the equi libr'iuni is stable, because increases in the backl'oğ tend to ma.ke the system drift dowtrwafd, arıd viCe \l'ersa.

An interesting question about the bistable sitUation öf Fig. 15 (b) is: What is the chance of the system getting itself into the high backlog stable state? Inductively , the farther down the "hill" the low backlog stable equilibrium point is, the less likely it is that a momentary burst of activity will push the system past the point of no return.

To calculate the probability of an initially eii:pty)sysfem being near the unstable equilibrium point after a givel1 number of slots, proceed as follows, Take the state-transition matrix and make state k absorbing by setting Pkk=l and all the other elements of row k to 0. With this new matrix if the system ever gears to state k, it will stay there. This model 18 an because the chance of spontaneously escaping approximation, region, however small is not zero. from the unstable By computing successive powers of the transition matrix, we get the multi step transition probabilities. The element Pok gives the probability of the system having reached state k starting from an empty system. By plotting Pok against the power of the matrix, we get the probability of reaching transition the unstable state with the corresponding time interval.

16 (a) shows the throughput rate, input rate afrd Figure load line for 25 stations and p=0.01. The corresponding multi step transition probabilities, PoN are show in Fig.16 (b). With = 0.25 there is a 50% chance that the backlog will grow to 2 within 10.000 slots. With $_{oc} = 0.40$ the 50% point has dropped to 1000 slots. With theie pa.r~hieters, the system will collapse of begin sJatt.e(1 .itp\. Ry adjusting ∞ it is within seconds to trade off stability against delay. since the mean possible delay is closely approximated by the backlog divided by the at the equilibrium point, from Fig. 16 it is easy to throughput see that the stability is more sensitive to oc than the del Consequently, system designers have some latitude in choos ing a value o(oc that is small enough to provide good stability, with out extracting too heavy a price in terms of excessive delay.

그 철주 안 없으니



Figure 16 (a- Input throughput for N=25 and P=0.01 b-Probability of the system having reached state 25 within the specified number of slots.)

Controlled ALOHA

The only way to' guaran.lee a stable syslem with a.fixed oc is to restrict the number of users to such low value that load line crosses the throughput curve in only orre place. Doing so is often uneconomical, because it gears the system to wôtst. case behaviour. Put in other words, setting the parameters sb<as to guarantee stable operation under all circumstances, no matter how unlikely, restricts the number of users to a value so small that the system may cost too much to be economigally justified. however, as we can see from Fig. 14 (a), or has an important influence on the shape of the throughput curve, and hence on the stability . If we could somehow decrease oc when the system got into trouble, we could raise the throughput curve above when the backlog is small (to get rid of packets quickly and achieve a short delay time), and small when the backlog is large (to drive the backlog down).

Several methods have been proposed (Lam and Kleinrock, 1975) for controlling _{oc} dynamically in order to achieve short delays when the system is lightly loaded and stable behaviour when it is not. üne method is simply to have each station keep track of how many retransmissions the packet currently in the

output buffer has to its credit. When a packet is first generatedt the counter is set to zero . Upon each collision, it is increased by one. The value of the counter can be used as crude measure of the channel load. The station could have a series of (fixed) ret ran smission probabilities, cci (i.e., random is ation intervals, Lj) subject to =: $+1 \sim =r$ - As the channel load increases, everv station attempting to transmit will: automatically wait longer and longer between successive retransmissions This method has that it relies only on information the advantage directly available to the station (i.e. it does not require the station to monitor the channel other than to see if its own packets have collided). It has the disadvantage, however, that each packet must learn about the channel load all over again.

Lam and Kleinrock (1975) have studied ano.ther controlled ALOHA method that does retain some history, so that new packets can profit from their predecessors experience. In his scheme, there are only two values for oc.a large one that is used when the system is operating properly, and a small one that is used when the system gets into trouble . Whenever the collision rate exceeds a certain threshold value, all stations switch to the lower value and continue to use it until things are back to normal.

A more direct approach to the control problem is based on Abramson's observation that the throughput is maximised at G=1. The goal of this method is to have each station estimate G

and adjust its own value of ∞ upward if G<1 and downward if G>1.

To estimate G, we can use Eq. (1). The probability of an empty slot, P_{0} , is $G_{0}e_{-0}/0!$ which is just e_{0} . Taking logarithms, we have $G=-In P_{0}$. P_{0} can be easily estimated by noting the fraction of slots that" were enlpty diring the recerit past. This requires a shift register with 1 bit per slot. The observation period should be longer than the round-trip time, to prevent instabilities.

When the channel is lightly loaded, we would like α as large as possible subject to the constraint of Eq. When the channel is heavily loaded we want α form G, for example oc=e⁻ o/(R+1). This function ensures that we do not atternpt to reduce L bellow 1, even when the channel is idle. The function is also easy to compute, because e -G is just the fraction of the shift register positions containing 0 (no collision).

Gerla and Kleinrock (1977a) have proposed another method based on estimating G. Their method is designed to handle the case of stations that do not block when a packet is generated, but just keep on going. As a result , two queues may build up in each stations: virgin packets that have not been sent yet and packets that are waiting for retransmission. Different probabilities are used for these queues. They propose that each packet header contain the probabilities currently in use at the sender's station. In this way each station can set its parameters to be equal to the sender's station. In this way each station can

set its parameters to be equal to the mean of everyone else's. a correction term of the form (1-G)D₁, where plus D is an When G > 1, the collection experimentally determined constant. the appropriate oc, when G<l, cc term is negative, decreasing increased. The feedback inherent this proposal ensures that the only change slowly and, furthermore, probabilities that no station gets too far out of step with the others.

Reservation ALOHA

Even the most clever dynamically adjusted control scheme will never get the throughput of a slotted ALOHA channel above /e (except f'or sm.all N) However, at high chann el loads, there are other methods for making good use Of. a sirig le shared channel, in particular, time-division multiplexing. Several researchers have proposed control schemes that act like .riorrna.l or nearly normal slotted ALOHA at low channel utilisation, and move gradually over to some kind of TDM as the channel load grows.

All these methods have one feature in common: some slots are reserved for specific stations. Stations are required to refrain from attempting to use a slot reserved for somebody else. The methods differ in the way reservations to use a slot reserved for somebody else. The methods differ in the way reservations are made and released. For comparison, remember

that in TDM the slots are organised into frames of N slots, with a specific stations.

Binder (1975) proposed a method that starts out with the basic TDM model and adapts to slotted ALOHA for low channel in. TDM, N consecutive utilisation. As slots are grouped together into a frame,\witht _aeachcstationiF'oWning" one frame position. If there are more solutions, the extra slots are not assigned to antenna. If the owner of a slot does not want it during the current frame he does nothing. an empty slot is a signal to everyone else that the owner gas no traffic During the next frame, the slot becomes available to anyone who wants it, on a connection basis. If the owner wants to "his" retrieve slot, he transmits a packet, thus forcing a collision (if there was other traffic). After a collision everyone except the owner must desist from using the slot. Thus the owner can always begin transmitting within two frame times in the worst case. At low channel utilisation the system does not perform as well as normal slotted ALOHA, since after each the collides must abstain for one frame to see if the collision. owner wants the slot back. Figure 17 (a) shows a frame with eight slots, seven of which are owned.

Figure 17 (Reservation scheme (a)Binder (b)Crowther c)Roberts)

One slight inefficiency with this method is that whenever the owner of a slot is through with it, the position must go idle during the next frame to announce that its owner is done. To eliminate. this **wasred** slot, an extra bit .coüld be<added to the header of all packets to anno'unce that the owrter thaf the owner did or did not have any more data for the next frame.

A much more serious problem, however, is that the number of users must be know in advance, or at least bpunded from aboce. If this is not the case, several users could be assigned to the same slot, in the hope that it will not occur too often that both of them claim th'e slot sirmi It'ane ou sly. To arbitrate w'hen this does happen, each user could be given a static priority, with lowerpriorty usersdeferring to higher ones in the case of conflict.

S24

0.1

Crowther al. (1973) have proposed different et а reservation method that is applicable even when the number of unknow and varying dynamically. In station **1**S their method, slots do not have permanent owners, as in Binder's but instead. whenever a transmission is successful. the station making the succesful transmission is entitled to that slot in the next frame as well. Thus aslong as a station has data to send, it can continue doing so indefinitely (subject to some "Please-donot-be-apig" rules). Since it is unlikely that all stations will have long runs of data send simultaneously, this method works

well even when the number of slots devoted to each varying with demand. Figure 17 (b) shows a frame with eight slots. Initially, E is using the last slot, but after two frames, it no longer needs it. It lies idle for one frame, and then D picks it up and keeps it until he is done.

A third scheme, due to Roberts (1973), requires stations make advance request before transmist irig: Each to frame contains one special slot (the last one in Fig. 17 (c)), which is divided into V smaller subslots used to make reservations. When a station wants to send data, it broadcasts a short request packet during one of the reservation subslots. If the reservation is succesful. (i.e., no collision), then the next regulat slot (or slots) is reserved. At all times everyone must keep track of the queue length, so that when any station makes a succesfull reservation it will know how many <lata when any station makes a successful reservation it will know how manu <lata slots to skip before transmitting. Stations need not keep track of who is queued up; they merely need to konw how long the queue is. When the queue length drops to zero, all slots revert to reservation subslots, to speed up the reservation process.

g ser

To see if satellite packet broadcasting worked as well in practice as it does in theory, starting in 1975, DARPA (nee ARPA) began supporting an experiental packet broadcasting system using a 64 kbps Intelsat IV satellite channel (Chu et al., 1979; Jacobs et al., 1978; Jacobs et al., 1979; Kalın, 1979). The paper by Kalın is especially interesting; it describes the legal

hurdles faced by DARPA, even though the project was clearly a research experiment, and was sponsored by the U.S.government.

nctwork, onw as SETNET, uses large ground The satellite stations in Eta m, West Virginia; Goonhilly, England; and ground Tanum, Swcden: and a smalled station in Clarksburg, Marylan. The protoe'o'l üsed, called PODA '(Pr'i ör ity Oriented has two art snrs, FPODA (Ei'x ed PODA), Dcmand Assignment), POIJa), which which is csscntial TDM, and CPODA (Contentioh version of Roberts reservation isa sopliisticated schemc.

Each Iramc is divided into a data part and a control part The relative sizes of for reservations. the two parts are elynamically aeljusted elepending on the load. A portion the data part is reserved for FPODA, providing guaranteed bandwidth for those processes needing it. The rest of the data part is used fcatura of CPODA is the ability to by CPODA. An interesting piggyback reservation rcquests onto data packets, whit the consequence that heavy users of the channel do not have to keep fighting for reservation slots. The reservation information is used in the scheduling of the channel rather than simply using first come, first served. A sophisticated scheduling of the channel rather than simply using first come, first served. А algorithm sophisticated scheduling is needed since packet specch is transmitted, as well as <lata.

523

PACI(ET RADIO

Ground raclio packet broacleasting differs from satellite packet broadcasting in several ways. in particular, stations have limited range, introducing the need for radio repeaters, which in turn affects the routing and acknowledgement schemes. Also, the propagation delay is much less than for satellite broadcasting. In the Follo.wing sections we will look at some of the ways in which ground r.adio differs from satellite packet broadcasting. More details are provided in (Gitman et al., 1976; Kahn, 1977, Kahn et al., 1978; anel Tobagi, 1980b, 1980c).

The University of Hawaii ALOHA system

The first computer system to employ radio instead of point-to-point wires for its communication facilty was the ALOHA system at the University of H'awaii The system first went on the air in 1971, and has been evolving ever since. Since this system is the ancestor of all packet broadcasting system, we elescribe it below in some detail. For a more complete description, see (Abramson 1970, 1973b; and Binder et al., 1975).

The ALOHA system was begun to allow people at the University of llawaii, who were spercad out over seven campuses on four island, to access the main computer center on Oahu without using telephone lines, which were expensive and unreliable. Communication is provideel by equipping each witlı FM radio with station а small transmitter/receiver talk to sufficient 1ange (30km) to the computer center's

transmitter/receiver. Later, a powerful repeater was introduced, increasing the theoretical range to 500 km.

A11 either communication is from а station to the computer center or from the computer center to a station . There is no station-to-station communication. When a packet is received at the compüter ceriter, it is process'ed there. It is not retransmitted for all other stations to hear. This arrangement is from the satellite bro.adcasfin gom.o.del, fundamentally different in which the satellite is merely a big repeater in the sky. Since incoming packets are not rebroadcast, a station has by the central site. As a result, explicit acknowledgements are needed, just as in point-to-point connections.

some initial skepticism After overcoming about the unusual coomunication mechanism, the project was assigned two bands in the UHF part of the spectrum. üne frequency band, at 407 .350 Mhz. is used for inbound traffic, from stations to the central site. The other frequency band, at 413,475 Mhz, 1s used for outbound traffic, from the central site to stations. Transmission is at 9600 bps. The of distinct vhannels for inbound and outbound traffic has important implications for the whole system organization. After several years öf experience, the research group concluded that a single channel probably would have been a better idea.

The original rationale for haing two distinct channels was the fundamental difference in the inbound and outbound traggic. Inbound, many uncoordinated ysers are competing for

access to access to a shared resource. Outbound, a single site is in complete control of the channel, so there is no contention and no collisions. The basic idea is to use the inbound channel on a random access (what is now called pure ALOHA) hasis, and the outbound channel on a straight broadcasting hasis, with each station extracting those packets directed to it form the output stream.

Figure 18 shows the essential elements of the ALOHA system. At the central site is a 32 K HP 2100 minicomputer, called the Menehune (the Hawaiian word for "imp") that is connected to the antenna. All data in or out of the central site passes through it. The Menehune, in turn, is connected go two large computers, an IBM 370 and the BCC 500, as well as to two other networks, ARPANET and PACNET. Each station has a control unit that were hardwired, but later Intel 8080s were used to provide more flexibility. Some stations are connected to concentrators to reduce transmitter/receiver costs.

Packets consist of four parts. First comes a 32-bit headre, among other things, the user identification containing, and the length. To provide good reliability, the header packet **1**S followed by a 16-bit checksum. Next comes the data, up to 80 byte s (640 bits), followed by another checksum. The maximum 32+16+640+16=740 bits. At 9600 packet is bps, the transmission time for the longest packet is 73 msec.

When a station has <lata to send, it just goes ahead and send. This way ogf operating is ALOHA at its purest. When the

Mencliunc correcily receivas a packet, it inserts an acknowledgement packet into the output stream. If a



station does not receive an acknowledgement within a preset time, it assumes that the packet suffered a collision, and retransmits it. The retransmission intervals are distributed between 200 and 1500 mscc, with various distributions having been tried (e.g., uniform, three shorts and then a long, etc.).

Although control of the outbound channel might at first seem trivial, there are some important design issues here too. For one thing, transmitting packets on a first come, first served hasis is not a good idea, because ifan acknowledgement packet is delayed 100 long, the station will assume that a

collision has occurred and send the packet again. Under conditions of heavy load, many unnecessary retransmissions will occur, making the situation even worse. Consequently, the Menehune maintains a separate queue for acknowledgement packets and gives this queue priority over the <lata queue.

The use of store-and-forward repeters (to amplify signals and thus .in creas e .1heir range) introduces some problems. The repeaters use the same frequency for output as for input. This means that a rapeater connot start the retransmission until the input has been completely received and stored; otherwise, thev would overlap and be unintelligible. If the Menehune broadcasted a new packet while a repearter was still forwarding an old one, stations within range of both while a repearter and the Menehune would receive the superposition of the two packets, which is meaningless. Furthermore, while forwarding a repeater is able accept new packets. packet, th e not to Consequently, the Menehune must be careful about the order in which it send packetsi and must the the repearter geometry into accounty when making schedulling decisions.

Another problem unique to ALOHa is the way in which errors are handled. Because packets are not rebroadcast for all to hear, the Menehune must send acknowledgement packets back to the stations. Doing so is straightforward and causes no problems. What does cause problems is the Menehune's lack of knowledge about whether or not its transmissions are being correctly received. in an urban area, such as around Honululu,

there are many electrical devices that can cause static on an FM radio.

The obvious solution is to have the stations acknowledge of computer <lata, as well as vice versa. The trouble reception is, these acknowledgements must be sent in the random access where it' s channel, dog eat dog. The ptobabilty of an acknowledgementbeing delayed for ssub stanci af period of time due to collisions and timeouts is much greater than on a pointto-point line. If the acknowledgement is delayed too long, the Menehune times out and tries again, wasting valuable bandwidth.

The worst part of it is, the station has no way of konowing whether or not its acknowledgement has gottenn through. üne solution would be to have the acknowledgement themselves be This solution is not used. Instead, acknowledged. output are numbered. The Menehune simply refrains packets from sending a new output packet untill the previous orre has been acknowledged. When a station receives a new output packet, it knows that its last acknowledgement has arrived.

Yet another peculiar property of packet broadcasting that was first noticed in the ALOHA system is the need to decouple the transmission speed from the station speed. in а conventional time-sharing system, the computer knows the speed of all the stations and sends <lata to them at the correct speed. it does not try to force feed 300 bps hardcopy matrix

prin crs at 9600 bps, There is also no need to, because each station had a dedicated line connecting it to the computer.

is different. The Menehune in ALOHA, the situation could keep track of the printing speed of the recipient of each packet and adjust its bit rate accordingly, but sending a page of output at 300 bps would be a tremendous waste of bandwidth and would adversely affect all other users, even those with fast The solution used is to have the Mcnuhune CRTs. always at 9600 bps with loca! buffering broadcast provided at the station. Never |hc| css, the Menehuns must keep track of the actual printing speed, to avoid sending more data before the old data have a chance to be printed.

would An a alternative solution be to have the station refrain Irom scnding its acknowledgement until it had finished The trouble with this is that if the acknowledgement printing. were delayed substantially due to collisions, the printer would lic idle, even when there was plenty of work for it. By sending the acknowledgement as quickly as possible (i.e., as soon as the data are received), the printer can be kept busy all the expense of some complexity in the Menehune.

Desing Issues for Packet Radio Networks

in the preceding section we looked at the original ALOHa sysrcm as an cxample of a ground radio packet broadcasting system. In this section we will examine some additional design problems olten present in these systems. The introduction ofa number of repcaters, in particular (to increase the largc geographic coverage of the system), brings with it a number of complications because repeaters store incoming packets and then rebroadcast them on the same frequency. Simultaneous rcception and |ransmission is lhcrefore impossible.

Let us lirst look at the conceptual model of a packet radio system. There are (at least) the situations in which packet radio is altractive as a method of loca! distribution from a central site lo remote stations,

1-The stations are located in arcas where the telephone system is poorly developed or nonexistent; nearly ali rural areas, and most of the Third World falls into this category. Automated weather and scismic data collection stations arc often parachuted into jungles, deserts, and hostilc inountain terrain, which frequently lack the amenities of civilisation, such as iclephone poles.

2-The stations are mobile. A fleet of ships is a good ^example of a group of users that is inherently mobile. Police ^cars, ambulancos, fire engines and taxis are other examples.

3-The stations have a high peak-to-average traffic ratio, Or a low data rate in both cases, the cost of a dedicated line

may make the application uneconomic. Packet radio offers the possibility of sharing a single channel instead of having a large number of channels with fixed (and mostly wasted) capacity.

Although portable FM radio transmitters are usually equipped with a simple whip antenna, which makes them omnidirectional, repeaters may equi pp e d be with or directional antennas, as needed. Repeaters omnidirectional may also be able to adjust their range. If a substantial number of repeaters ate battery powered, the system design must be sufficiently redun band to cope with repeaters whose batteries have gonedeaf, since this will be a regular occurrence.

When a packet radio system contains multiple repeaters, routing becomes an issue again, just as in point-to-point-anforward networks. The most naive routing strategy, just having each repeater act as a transponder, storing and forwarding all incoming packets, does not work at all. Every packet would just bounce back and forth between adj acent repeaters forever, clogging up the channel.

When a repeater receives a packet, it must make a decision. To forward, or not to forward, that is the question. The choice of forwarding or discarding is analogous to an IMP's choice of output line in a point-to-point network. The goals of the routing procedure are the same as in all networks, of course; to achieve a high probability of delivering packets to the ultimate destination and to consume the minimum resources in the process.

Gitman et al. (1976) have suggested three possible routing strategies. The first algorithm is a modified version of Each repeater just forwards flooding. all incoming packets. subject to two restraints. The first restraint is a hop counter, which is included in the header of each packet. Each time the packet is forwarded the hop counter is decremented by one. When this counter gets to zero, the packet is discarded. This rule by itself is sufficient to guarantee that no packet lives forever.

The choice of the initial value of the counter is an important design parameter. If it is too small, packets destined for outlying areas may never arrive. lf it is too large vast numbers of useless duplicates may be generated, wasting bandwidth. The problem is complicated by the fact precious that the central site may only have an approximate idea of the topology, if that much. in Fig. 19(a). the shortest path from the station to the central site is two hops, but if the battery in repeater A has become sufficiently drained, it may not have enough power to reach the central site. in that case a hop limit of two in the packet would prevent the packet from ever arriving. With a hop limit of four, the packet could take the other route. The problem of choosing the hop count is closely connected with the problem of ensuring that a point-to-point network has k node disjoint paths between each pair of nodes. As the hop count is allowed to increase, more and more al terna te paths become feasible, there by making the network

less and Iess vulnerable to repeater failures. If each station has a reasonable idea of how many hops it is form the central site, it can choose a hop limit appropriate to its distance rather than having all stations use the worst-case value.

A second restraint that is need to make flooding a practical algorithm is for repeaters to be available to recognise returns of packets that they have recently forwarded. Consider the example of fig. 19 (b). Suppose that the maximum number of hops for the station shown is set a four.



Figure 19 (a-T'h o shortest route from the station to the central site is shown as a solid line, b-An other repeater topology)

when the packet is answer packet due to another repeater, it appends its identification to the packet. Form the returned answers, the central site can easily determine the shortest path to each repeater. Upon learning the new topology, the central
site then informs each repeater of its new label. A label is just a path description, as shown in Fig. 20.



Figure 20 (A repeater tree and its labeling)

ünce the repeater tree has been formed, each data packet sent by the central site can contain the label of the repeater to which station has been assigned. The packet must also contain a pointer telling which field in the label is the current one. When a packet arrives at a repeater, the repeater checks to see if it is named in the current field. If so, it advances the pointer by one field an forwards the packet. Of not, it just discards the packet.

To see why this incchanism is needed, consider a packet directed to repeater 1310 in Fig. 20. The central site initialises the pointer to the first 1 and broadcasts it. Repeater 1000

knows that it must forward it, and any other repeaters that hear the packet do not forward it. By the time repeater 1300 forwards the packet, the pointer is now pointing at the second 1000 might 1. Were it not for the pointer, repeater detect the packet, see that it was on the path, and forward it again.

path has failed, If repeater along the required all is not failing to get an acknowledgem.ent, lost. The repeater, could set a bit in the header telling all repeaters to adopt the flooding for the packet. It could also announce the failure algorithm to the central site (assuming that it was still connected), requesting the central site to conduct another probe and relabel the tree.

The broadcast nature of packet radio makes a completely different strategy also possible: the repeater could simply advance the pointer one field, turn up its power, blast away, and hope to skip over the failed repeater. This strategy implies a conservative deign, in which repeaters normally operate at less than full power. üne problem, however, is that although the correct repeater may receive the packet, it may be too weak to get the acknowledgement all the way back.

In the third routing algorithm, a repeater only forwards a packet it is closer to the destination than the last repeater that forwarded the packet. Each repeater is assumed to know its distance in hops from every other repeater. This information can be acquired by having each repeater broadcasts its distance table periodically. Each <lata packet contains the identification

of th e destination an d th e sender's distance from that destination. At each hop there is a new sender, and hence a new distance. When a packet arrives at a repeater, the program checks to see if it is closer to the sender. If so, the packet is heading the right way and is forwarded. If not, the packet is heading the wrong way and is discarded.

As an example, consider Fig. 21 When the central site sends a packet to A, it indicates that the distance is three hops. Both J and K receive the initial transmission. J knows that it is only two hops from a, so it forwards the as far away as the sender, so it discards the packet. The packet sent by J is received by F, G, H and k. Of them, F and G are the only ones closer than rwo hops, so they alone forward the packet, each one changing the distance in the header to 1. A will eventually get two copies of the packet. This algorithm always uses the shortest path, but it is more robust that the previous orie, because it concurrently tries all alternate paths whose length are equal to the minimum length. It also consumes more bandwidth. This algorithm is also minimum length. It also consumes more bandwidth. This algorithm is also applicable to station traffic rather !han just station-central stationsite A last point in its favor is its ability to easily adapt to traffic. m The ARPANET obile repeaters. algorithm adapts to new ^topologies.

An important design issue that must be resolved no matter $\frac{1}{2}$ hich routing algorithm is chosen is the way packets are

ackonwledged. There is a basic distinction between hop-by-hop acknowledgement and end-to-end acknowledgement. When the error rate is low, the end-to-end scheine may be sufficient, but when the error rate is high, as it is likely to be in a packet bro adc as ting en vironinen t, hop-by-hop should be used in order to detect lost packets as quickly as possible. üne property a network of packet repeaters has in common with satellite broadcasting is the ability of a repeater to tell whether or not its neighbour received the packet correctly by just listening to the retransmission, In many cases this may eliminate the need for an explicit acknowledgement. However, some kind of acknowledgement is still needed at the end of the path, because neither the station nor the central site retransmits its input.



Figure 21 (A repeater Network)

Up until now we have made the assumption that when two Packet collide, both will be lost. in fact, these is overly Pessimistic. Since many FM receivers are able to extract the stronger of two overlapping packets without This error. phenomenon is known as the capture effect. To see how this Influences the system performance, imagine that there are two classes of stations, one broadcasting at high power, and one at low power, (or one group close by and one group far away, both ^hat the same power). Of two overlapping packets both originate ^from the same group, the receiver cannot disentangle them, and both are lost. If, however, they originate from different groups, the stronger one will be accepted. in half the cases, what was previously counted as a collision is now a success! An ¹ⁿteresting property of the effect is capture that under ^conditions of heavy load, stations close to the central site may generate so much traffic that stations farther out are locked out ^{co}mpletely. This occurs because when a close in and a far out station collide, the close in station has a stronger signal and always wins. Metzner (1976) has analysed the capture effect quantitatively.

CONCLUSION

Communication satellites are fundamentally different from terrestrial store and-forward networks in that they offer a single channel that must be allocated among the users. When

there are many bursty users, FDM and TDM are inefficient ways ^{to} use the channel, which leads to the ALOHA approach.

ALOHA systems, stations broadcast at will, detect the presence or absence of collisions them selves and retransmit if necessary. The key difference between ALOHA and other protocols for controlling a single channel is that in ALOHA there is no centralised channel allocator.

Slotted ALOHA can be modelled as a Markov process. Form this Markov model, we can find the equilibrium state o ccup ancies, the mean throughput, an the mean delay -, as well a s analysing the stability. To prevent bistable ALOHA systems f om moving into the low throughput state, the retransmission pr obability can be dynamically adjusted. By using various re servation techniques, the throughput can be increased well ab ove the 1/e limit of slotted ALOHA.

Packet radio is similar to satellite broadcasting, with the p^{r} oblem of the long propagation delay replaced by the problem t^{he} repeater topology. Routing becomes an issue again, just as in point- to-point networks, and some of the same techniques c^{an} be applied to deal with it. ..

REFERENCES or **BIBLIOGRAPHY**

^Acharya, Arup, and Badrinath, B.R., May 1993. "Delivering mlticast messages in networks with Mobile Hosts", 13 th
 ^{nt1.} Conf. On Distributed computing Systems, Dept. Of Co^t, muter Science, rutgers Univ., New Jersey.

- ^{Bad}rinath, B.R., Imielinski, T., and Virmani, Aashu, December 1992. "Locating Strategies for Personal Communication Networks", Networking of Personal Communications Applications, Depet. Of Coputer Science, Rutgers Univ., New Jersey.
- Davies, Brian H., and Davies, T.R., january 1987. "The
 Application of Packet Switching Techniques to Combat
 Net Radio", Proceedings of the IEEE, Vol. 75, No. 1, pp. 43-

15

 Erdinç, Oğuz, 1988. "Computer Impementation of Packet radio and AX. 25 Amateur Packet Radio Link-Layer Protocol", M. S. Thesis in Electrical and Electronic Engineering, Middele East Technical University, Ankara.

^{Fi}fer, William G., and Bruno, Fredereick j., January 1987. "The Low-Cost Packet Radio" Proceedings of the IEEE, Vol.75, No. 1, pp. 33-42.

- ^{Fi}sher, Jiffrey H., Cafarella, Johan H., Bouman, Charles A., Flynn, Gerard T., Dolat, Victor S., and Boisvert, Rene, Mayl988. "Wide-Bamd Packet Radio for Multipath Environments", IEEE transactions on Communications, Vol. 36. No. 5, pp. 564-576.

¹⁰rzepa, Stan, 1989. Your Gatewav Packet Radio, ARRL, Newington, USA.

¹⁰rzepa, Starı, Decamber 1992. " Packet Radi o Networking Basics: TCP'IP", QST, pp. 111. ben, j ohn, and Tronow, j anet D., January 1987. "The ffat:pa Packet radio Network Protocols", Proceedings of the IEEE, Vol. 75, No. 1, pp. 21-32.

einer, Barry M., Nielsen, Donald L., and Tobagi, Fôtiad, Junuary 1987. "Issues in Packet radio Network Design", Proceedings of the IEEE, vol. 75, No. 1, pp. 6-20.

iner, Barry M., Nielsen, Donald L., and Tobagi, Found, January 1987. "Issues in Packet Radio Network design" Proceedings of the IEEE, Vol. 75, No.1, pp.6-20.

^{yn}ch,Cifford A., and Brownnigg, Edwin B., 1978, Packet

io Networks, Wheaton and Co. Ltd.. Great Britain.

 chal ehn, steve October 1991. "Frequently Asked Qestions for Amaleur Packet radio ", Rec. Radi o. Amtaur Packet Newsgroup, IBM3090, METU, Ankara.

y, and Gupta, Someshwar C., August 1984. "Performance alu ation of a Protocol for Packet radio Network in Mobile uter Communication", IEEE transactions on Vehicular hn ology, Vol. Vt-33, No.3, pp.250-258.