39th IEEE VTC,
San Francisco, USA, 1989

INFORMATION DISTRIBUTION FOR NETWORK MANAGEMENT IN A MULTI-HOP MOBILE PACKET-RADIO NETWORK

V. Brass, B. Walke
Department of Electrical and Electronics Engineering
Dataprocessing Techniques
Fern Universitaet of Hagen, West Germany

ABSTRACT: The paper presents an efficient update procedure for stations in a multi-hop S-ALOHA packet radio network. Implicit acknowledgements (ACK/NAK) will be provided for point-to-point and broadcast communications. The simulation results show the mean delay of packets to stations in a 1-, 2- or 3-hop environment. The results indicate the admissible ratio of the generation probability q (packets/slot) to the transmission probability p (packets/slot) of own and relay packets, respectively. Furthermore, the frequency of topological changes and the mean distance from a station, having gained/lost contact to some neighbored station, to other stations needing update of their local data base is investigated through simulation. Our results appear to be valuable for defining management protocols for similar networks.

1. INTRODUCTION

We consider a decentrally organized packet radio network which utilizes the S-ALOHA protocol for channel access. Tasks like network initialization, registration of new stations etc. are served via a common radio channel /5/. All stations communicate point-to-point, whereby a load to the network is generated dependent on the source traffic rate per station. In /1/ an analytical and simulative examination of the one-hop throughput and delay is given. Tasks like registration of new stations entering the network and update of the station-local information base require transmission of broadcast (bc-)packets (one station transmits a packet, which is then forwarded to all other stations in the network). For this purpose a routing scheme is introduced which minimizes both the traffic volume and the end-to-end delay of packets. It is assumed that all stations know the actual network topology and thereby, which stations are in direct radio contact. Therefore, it is possible that bc-packets can be routed through the network such that each station receives a packet only once, and that it will be transmitted only to a well-defined part of their neighbored stations. For bc-packets a new acknowledgement scheme has been developed by which each station gains implicit acknowledgements by observing its neighbors transmissions. ACKs and NAKs for broadcast packets are very efficiently generated by piggy-backing them on each transmitted packet. Without permanent update of their local information, stations don't know the actual network topology in a mobile network. The mean distance from a station, having gained/lost contact to some neighbored station, to other stations which are required to update their routing table is investigated by simulation.

In section 2 the model assumptions and the notation are described. In section 3 the routing problem is discussed and in section 4 a new acknowledgement scheme for bc-packets is introduced. In section 5 the simulation results of our routing algorithm and the new acknowledgement scheme are discussed. In addition, the effects of the mobility of stations on throughput, packet-delay, update-rate etc. are investigated. With the assumptions made for the mobility of stations, events like loss or gain of radio contact between two stations are observed. Based on that we present a proposal for an efficient update of station-local connectivity information, which advantageously supports distributed routing of packets in a mobile environment. Finally, a summary and an outlook to further research are given.

2. MODEL-ASSUMPTIONS

The considered system is characterized by the following parameters:

(a) the total number of stations is n = 50
(b) each station has a transmission radius sr of 2.0km
(c) two stations have a maximal distance of 2r = 10.0km (2r = diameter of the network)
(d) all stations are homogenously distributed in an area with radius r=5.0km
(e) the time axis is divided into slots, the slot-length is equal to the packet length

For the traffic behaviour, the following assumptions are made:

(f) homogenous traffic intensity of all stations with:
   - a source rate of q packets/slot and
   - a transmission probability of p packets/slot, if one or more packets are waiting for transmission
(g) homogenous traffic relationships between stations: each station sends a source packet to any other station with equal probability 1/(n-1).

Figure 2.1 gives an example of the topology of the considered network.

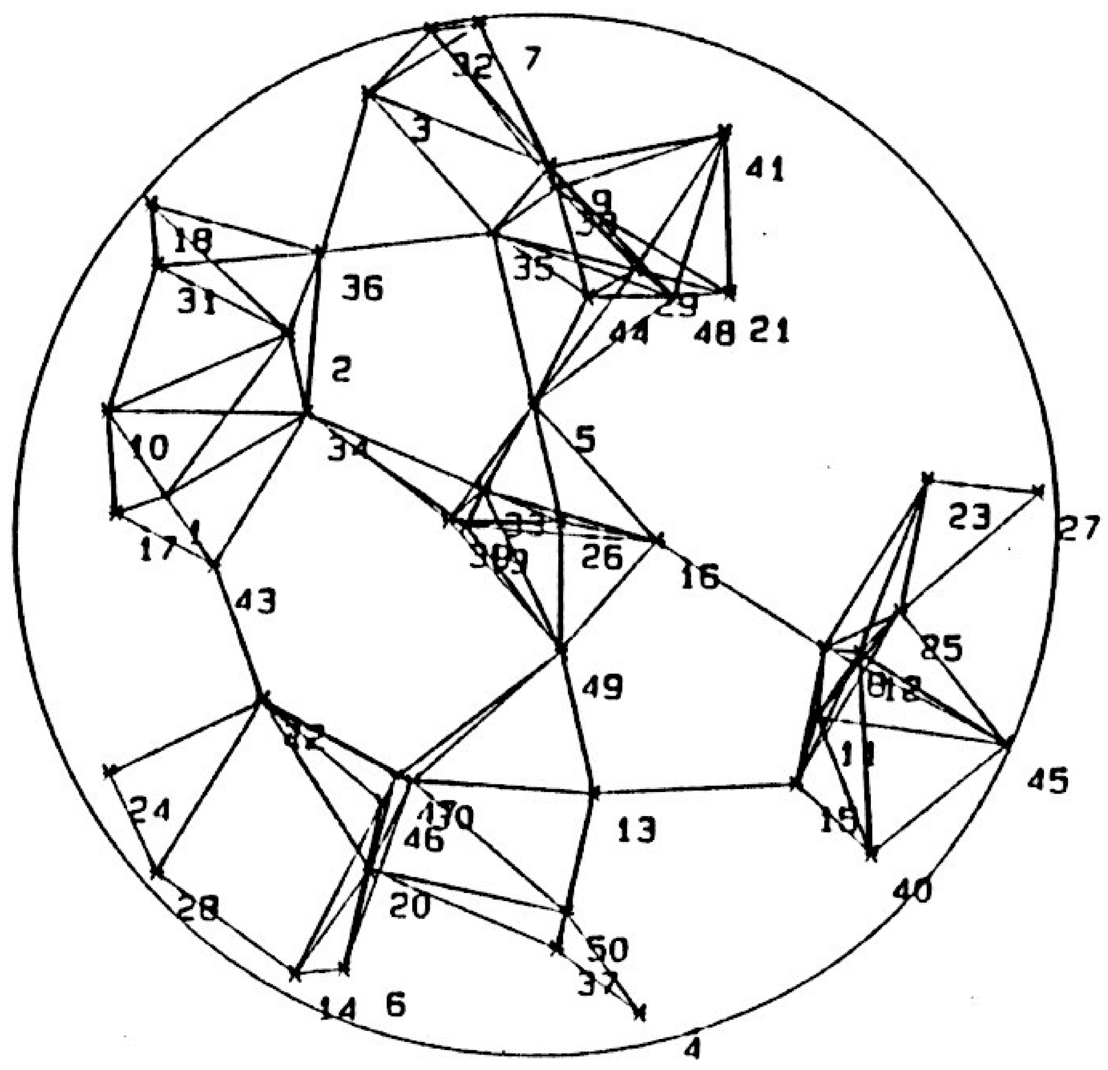
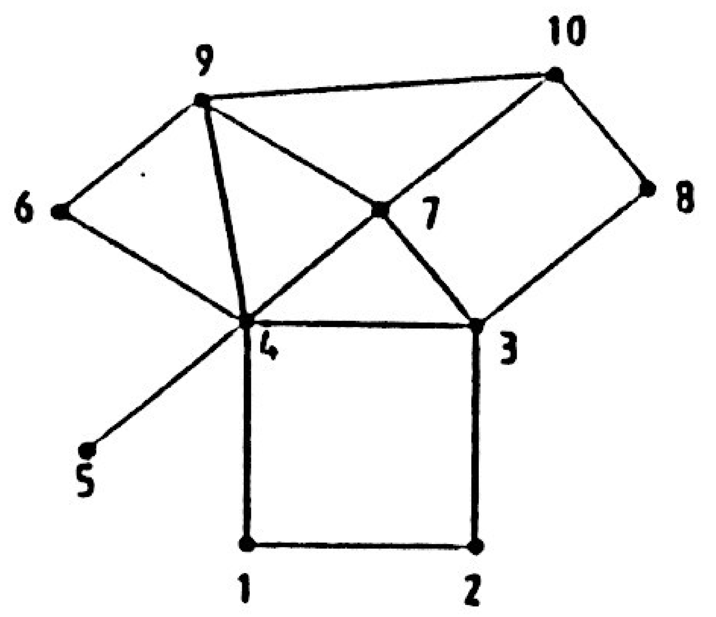


Figure 2.1: Example Network. The solid lines between the stations mark the radio contact.

3. ROUTING-ALGORITHMS

In this section a routing scheme for the transmission of bc-packets is presented. In a multi-hop packet radio network, one or more stations in sequence are necessary to transmit the packets to the final destination (store and forward). Stations are assumed to know the network topology (e.g. figure 2.1). The route to a final destination is determined, decentrally, by each station/relay according to the minimal number of hops necessary /1/. Figure 3.1 gives an example of a routing table of station S4 for point-to-point communication. In addition to the optimal route, one or more alternative relay stations are indicated.



routingtable of S4

sink	relay	
1	1	(3)
2	1	(3)
3	3	(7,1)
5	5	(-)
6	6	(9)
7	7	(3,9)
8	3	(7,9)
9	9	(6,7)
10	7	(9,3)

Figure 3.1: Routing table of station S4

In case of bc-traffic a packet (for example a registration datagram) is sent multi-hop from one station to all other stations in the network. To reduce the traffic volume the following requirements should be satisfied:

(a) completeness: each station must receive the packet at least once,

(b) minimum traffic: each station should receive the packet only once.

If a station has a bc-packet to transmit, in the first hop it transmits the packet to all of its neighbored stations. These stations decide via their own routing table to which of their neighbors the bc-packet must be forwarded in the second hop etc. It is guaranteed thereby that each station receives the packet exactly once. This is possible because each source station of a bc-packet generates a minimal reachability tree and transmits this tree in form of a control-map as part of the bc-packet. The control-map includes each receiver only once. Figure 3.2 gives an example of a control-map for station S1 (see figure 3.1).

R	1	②	3	4	⑤	⑥	7	⑧	⑨	⑩
S	1	1	4	1	4	4	4	3	4	7

Figure 3.2: Control-map for a bc-packet generated by S1

In the first hop the control-map defines all neighbored stations as recievers (R) and marks those stations as senders (S) which are necessary to further complete the reachability tree in some subsequent hops. Thus, in the first hop S1 transmits the packet to S2 and S4 simultaneously but only S4 is marked as a sender. In the second hop S4 multi-casts the control-map as part of the packet to S3, S5, S6, S7 and S9 but only S3 and S7 have to relay the packet to S8 and S10, respectively. The stations marked with a circle don't relay the packet. In figure 3.2 we can see that only S1, S3, S4 and S7 participate in the packet's transmission. The method is quite efficient because each station is assumed to have a complete view of the network topology. A minimum traffic volume is required to reach all stations. Apparently, this minimum is reachable only, if all stations are able to generate valuable reachability trees. In section 5.2 we discuss the effect of mobility on routing and find out that it is possible to route a very high percentage of all packets on the shortest route. It should be mentioned, that transmissions of the complete reachability tree as part of each packet is somewhat redundant, because any receiving station is able itself to generate the respective branch of the tree. If the flooding method (each station has to transmit the packet to all its neighbored stations) is used for the same network topology all stations except S5 participate to transmit the packet. The quantitative improvement of the control-map routing technique compared to flooding is evaluated in section 5.

4. ACKNOWLEDGEMENT-SCHEMES

To guarantee that a bc-packet reaches all stations a hop-by-hop acknowledgement (ACK) scheme is necessary. A transmitting station needs information whether its packet has been received by the desired neighbor(s) or not. Positive acknowledgements (ACKs) and negative acknowledgements (NAKs) are gained by observing the transmissions of the neighbored stations. In /2/ and /3/ a related method for implicit acknowledgements of point-to-point packet transmission is described. This method has the disadvantage that collisions at the sender Si waiting for an ACK during the transmission of a relay Sj will cause loss of the ACK. Such collisions occur if Si itself is active or there are two or more transmissions at the same time in its receiver range. Please note that the respective packet, nevertheless, might savely reach its destination which normally is not in the receiver range of Si.

Normally, a bc-packet is transmitted in the first hop successfully to many (all) neighbored stations. Due to the fact that many stations might relay the packet at the same time, implicit ACKs (NAKs) are then possible only by use of an improved algorithm. The additional traffic load caused by using explicit ACK-packets on the same data channel was already examined in /4/. There, it is shown that an undesirable correlation effect results which enormously reduces the network throughput.

The basic idea of the proposed ACK scheme is, that each station Si prepares a transmission-map (x bits) for its own local use. This map stores during which of the past x slots a packet was transmitted by itself. In the same way Si records in a ACK-map (x bits) during which past slots it received a packet successfully. The ACK-map is piggy-backed to each transmitted packet. All neighbors Sj which receive a packet from Si can definitely decide whether a formerly to Si transmitted packet was successful or not. Thanks to the fact that in a one-hop environment only one transmission can be successful, this ACK scheme is unambiguous. In case of capture a packet numbering scheme must be provided to protect against ambiguous ACKs, where a sender acknowledges only one of some simultaneously transmitted packets which was received successfully by receiver capture. If a sender has not received a packet from the receiver after the expiration of x slots an acknowledgement can no longer be expected.

The following example explains the proposed ACK scheme. Assume S1 (see figure 3.1) transmits during slots t2 and t7 two packets p1 and p2, respectively and marks these two slots in its transmission-map. Assume that S4 does not receive packet p1 due to a collision, but receives packet p2 in slot t7. The following implicit ACKs are supplied:
S4 had during slots t2 and t5, respectively, none or more than one transmission within its receiver radius sr. For this reason S4 marks the slots t2 and t5 in its ACK-map with F(alse), see figure 4.1. Slots t1 and t7 in the ACK-map of S4 are marked with T(rue) because of successful receiptions. By evaluating the ACK-map of S4, S1 is able by comparing the relative slot positions of t2 and t7 with its own transmissionmap to decide, that p1 has collided but p2 was successful. The marked positions t1 and t5 have no relevance to S1. An explicit ACK-packet is sent if a receiver does not have

an own packet or relay packets ready for transmission.

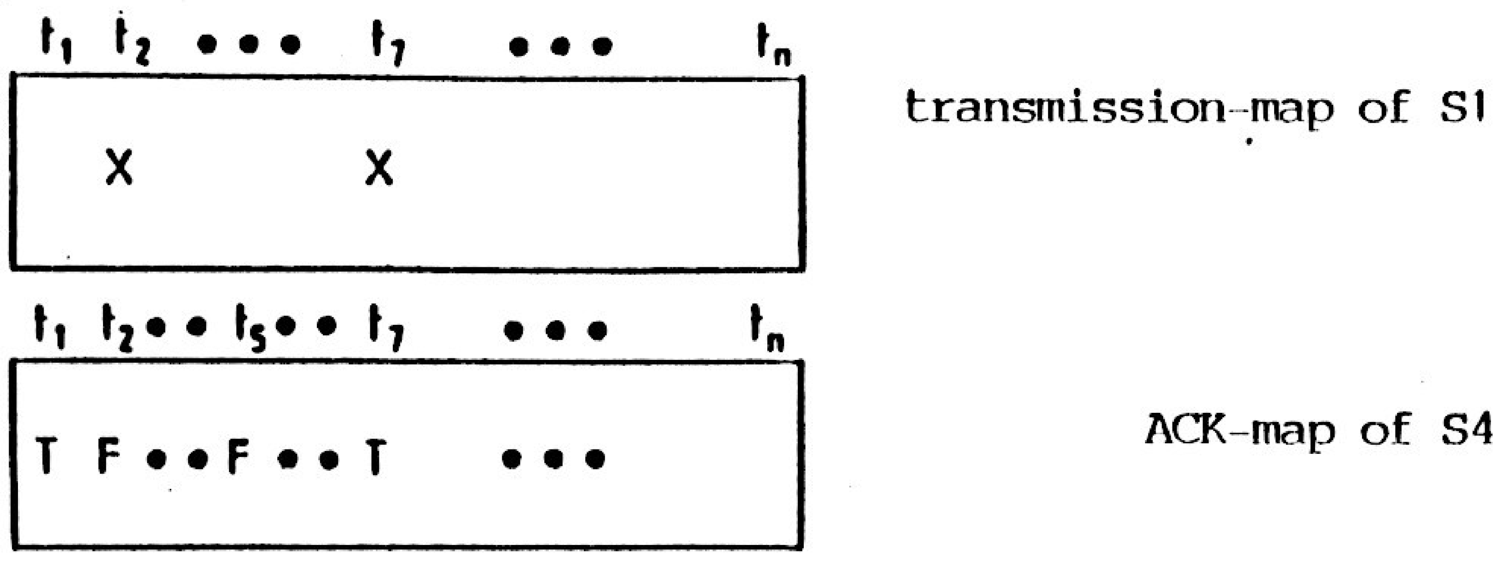


Figure 4.1: An implicit acknowledgement of S4 to S1 (see figure 3.1)

This acknowledgement scheme has the following advantages and disadvantages:

ADVANTAGES:
- the relay traffic needed for point-to-point and bc-communication is used to provide implicitely positive (ACK) and negative (NAK) acknowledgements.
- if a queue of waiting packets exist in a relay Sj, a successfully received packet will not be forwarded immediately. Since each packet of Sj contains its ACK-map, the sender gets its ACK/NAK already with the next transmission of Sj. Thus, in the case of a NAK, the sender must not wait for the expiration of the time-out (x slots) to retransmit its packet.
- each station does broadcast information about the receiver stations during any slot the more often, the greater its traffic intensity is. Apparently, the length of the ACK-map can be optimized to avoid unnecessary retransmissions, which appears advantageous to apply to stations carrying relay traffic.
- each transmission contain ACKs/NAKs for all packets, transmitted during the last x slots, independently of the source station.
- implicit acknowledgements are especially for bc-communication, where explicit ACKs would generate local bursts of traffic and result in a high collision probability of ACKs.

DISADVANTAGES:
- the length of the transmission-map of a sender which corresponds to its retransmission time-out and the length of the ACK-map must be sufficiently long such that the probability of a receiver to transmit at least once is sufficiently high. Otherwise a successfully transmitted packet will be retransmitted unnecessarily.
- the ACK-map in each packet can be interpreted as a small acknowledgement channel which consums capacity of the forward channel.
- in case of small traffic intensity of a station, the probability increases that the receiver must send an explicit ACK-packet to avoid a retransmission. This, however, is acceptable, due to the small traffic situation.

5. RESULTS

5.1 END-TO-END DELAY IN MULTI-HOP NETWORKS

The improvement by using the controll-map for routing was evaluated by comparing it with the flooding method. For this purpose a bc-packet was generated from a randomly chosen station in statistical time-intervals and the mean end-to-end delay was calculated by simulation. A parallel transmission of different bc-packets was not allowed. Figure 5.1 shows the mean delay excluding the time required for acknowledgements.

In case of the flooding method considerably worse results for the mean end-to-end delay are observed, if the transmission probability p is high. This is caused by the fact that through flooding a much higher number of packets are generated and therefore a higher number of collisions result. The packets must be retransmitted very often and the system becomes more saturated.

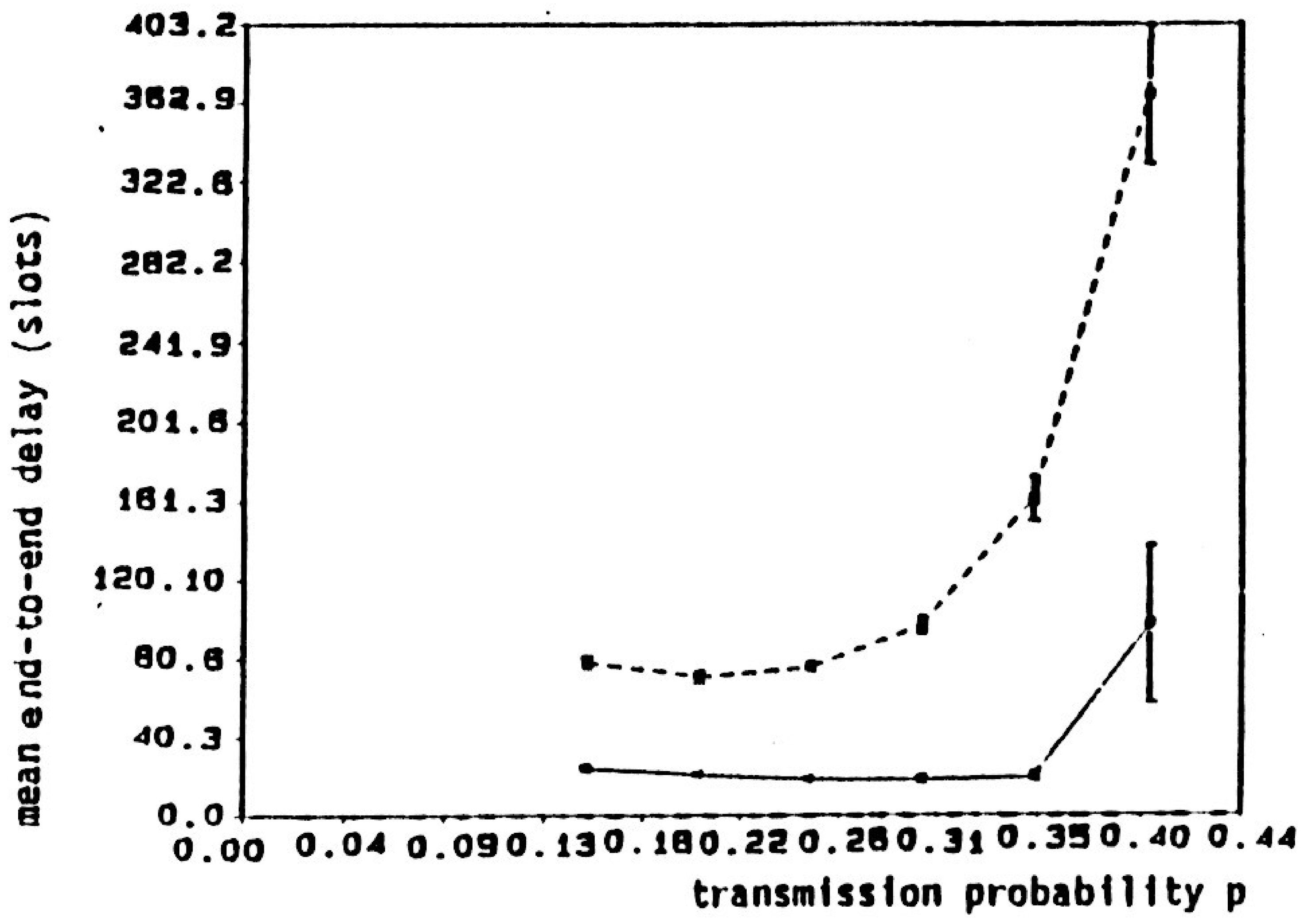


Figure 5.1: Mean end-to-end delay of bc-packets flooding(--); controlled routing(--); Parameter: n = 50, q = 0.0067, p = 0.1-0.4

To evaluate the ACK scheme for bc-traffic, a random station in the middle of the network was defined to periodically generate bc-packets and the mean delay was observed. We differentiated between the mean delays to reach stations which are exactly one, two or three hops away from the source station. This bc-traffic was added to the "normal" traffic where stations communicate point-to-point with the parameter p and q.

Figure 5.2 shows the observed mean delay over the transmission probability p where the source packet rate was fixed to q=0.005. If p 0.15 the result is that in spite of a small traffic volume some relay stations in the system were nearly saturated caused by the small transmission probability p being to small a multiple of the source generation q. If p 0.5 again the waiting-queues grow to infinity caused by an increasing high number of collisions and the system tends to become instable. In a range for p from 0.15-0.40, the packet delay decreases with increasing p. The reason is that with increasing p, a packet must wait a shorter time for a successful transmission. So, depending on the local traffic intensity and the local station density, an optimal transmission probability p exists. It must be recalled that for all stations the same transmission value of p was applied throughout. Please note that the quite high transmission probability (e.g. p=0.4) is effective only if a station has to transmit a packet.
In addition to the mean bc-packet delay for 1-, 2- and 3-hop and the ACK-time, we see the corresponding packet delay excluding the ACK-time (see curve -.-).

5.2 THE EFFECT OF MOBILITY ON THE ROUTING OF PACKETS

Now the influence of mobility of stations on the network topology will be examined. Mobility causes a partial or complete loss of a station's radio contact with its neighbored stations or a new radio contact. We will attempt to show that in spite of mobility a nearly continous and reliable radio traffic is possible. In this case it is necessary to have at each station actual knowledge of topological changes and resulting routes.

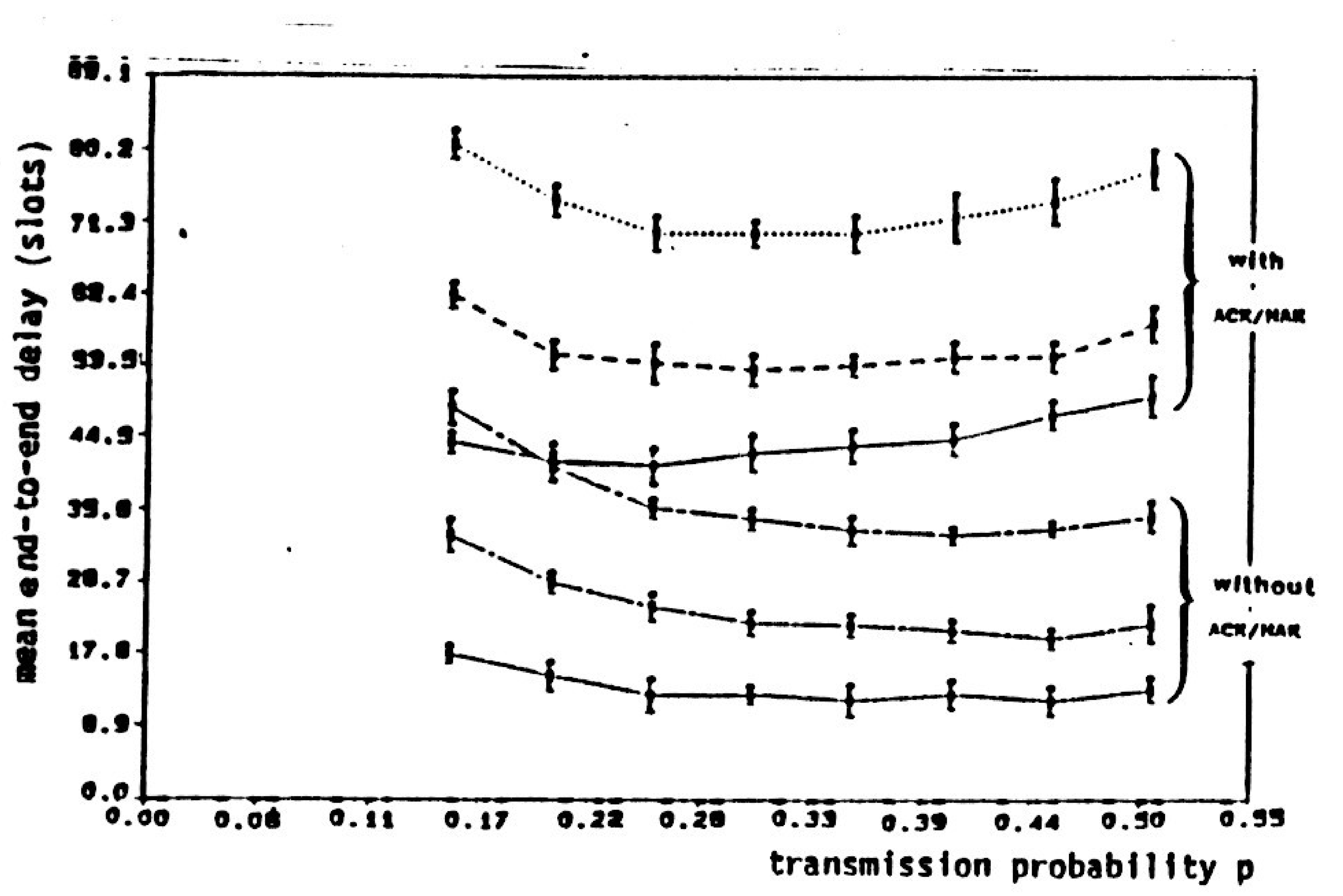


Figure 5.2: Mean bc-packet delay of 1-hop (—), 2-hop(--) and 3-hop (...)
Parameter: q = 0.005, p = 0.15-0.4

The results in the previous sections were gained for networks where stations remain fixed. To get more insight into the problems resulting from the mobility of stations, our model was extended by giving each station a mobility vector with two components: speed and direction of movement. To avoid a random fluctuation the number of stations in our network model which would require much more simulation time to reach statistically confident results, we enforced the stations not to leave a given circular area with radius r. Whenever a station intends to leave the area, it is reflected by the tangent in the point of contact according to law of reflection (billard-effect). The goal of this investigation was to investigate the frequency of topological changes and to find out what amount of network management traffic would be required for updating the station's routing table.

The speed of movement is not measured in km/h, but in meter/slot to establish a relation between the mobility and the slot-duration which gives a basis for more general statements. E.g. a speed of 50 km/h and a packet-length of 1kbit correspond to a speed of 0.87 meter/slot, if the transmission rate is 16kbit/s.

Each topological change - i.e. the loss or gain of radio contact between two stations - is an event which has an impact on existing routes in the network. The two affected stations are so-called <u>immediately</u> affected stations. Ohter stations, which must update their routing tables due to such events, which occurs outside their transmission range, are so-called <u>mediately</u> affected stations. Such stations are distinguished from other stations, which are so-called <u>indirectly</u> affected stations. They previously used or will use in the future the affected route; however, an update of their routing tables is not neccessary. Indirectly affected stations used routes via immediately or mediately affected stations. Furthermore, some stations exist which are absolutely not affected, because of their location with regard to the event. Each event is assigned to a class. Class i (i=0,1..,10) represents all events where i stations are mediately affected. Apparently, in class 0, only the immediate stations are counted. All events with 11 or more mediately affected stations are counted in class 11. Class 12 counts the as either a partition of the network or a reconnection of parts of the network.

Figure 5.3 shows the distribution of the number of mediately affected stations, separately for loss and gain of a radio contact, the discrete probabilities are shown connected. In case of gain of a new radio contact, the curve falls monotonously: In 50% of all gains of a new radio contact, no mediately affected stations exist. More than three mediately affected stations exist in only 8% of all cases; i.e. in 92% of all cases at most three stations are mediately affected.

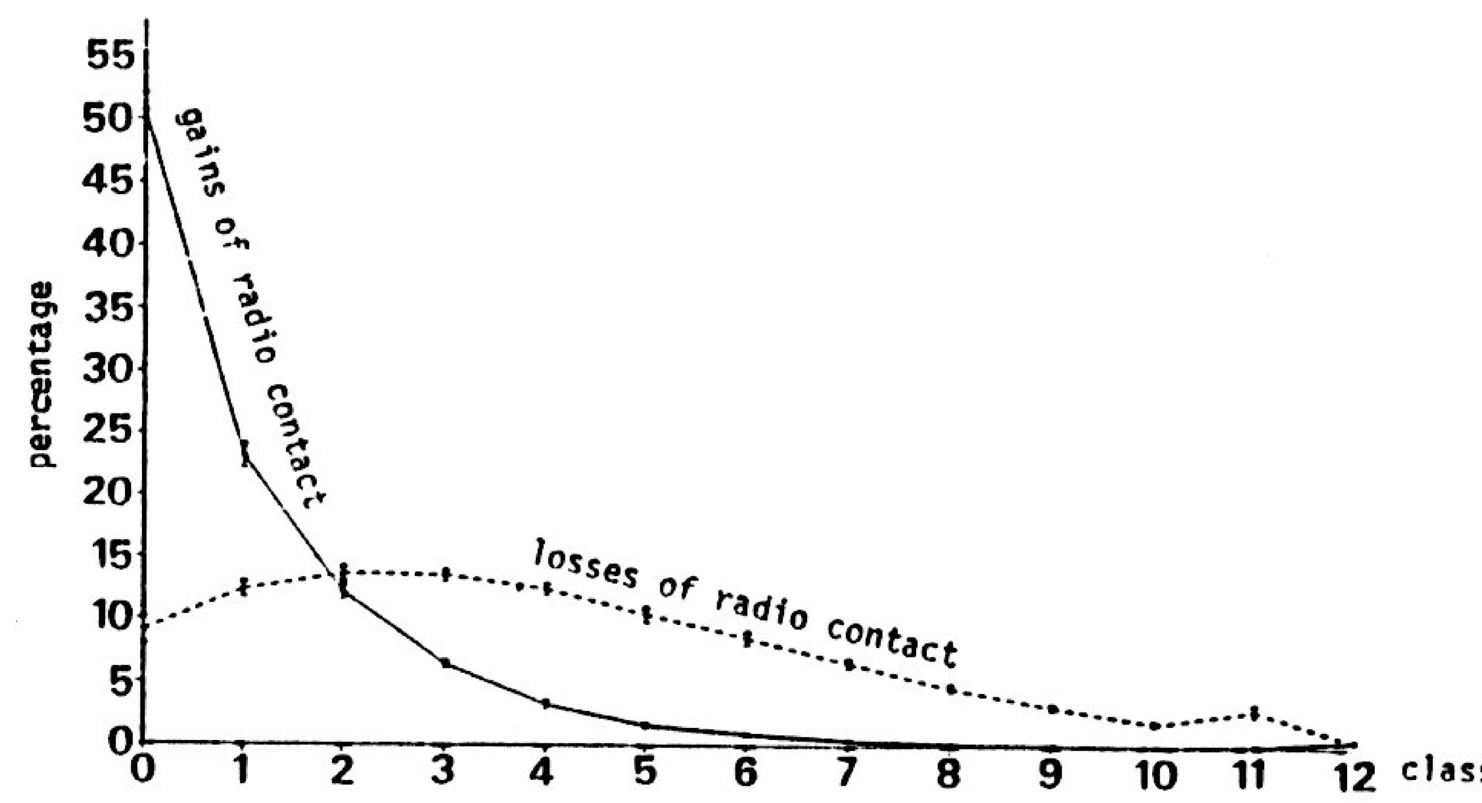


Figure 5.3: Distribution of mediatly affected stations.

In case of loss of a radio contact, the curve starts at 9% (class 0), reaches a maximum in class 2 and 3 at 14% and slowly decreases with increasing number of mediately affected stations. In 50% of all losses at most three stations are mediately affected. It can be seen that, generally, in case of loss more stations are mediately affected than in case of gain of a radio contact. This is caused by the fact that in case of loss more stations have to update their routing table. A new radio contact, which does not alter the actual number of hops on routes for indirectly affected stations, does not require an update of their routing tables.

In addition to the above discussed distribution, the mean distance (in hops) between the mediately and immediately affected stations is of enormous interest.

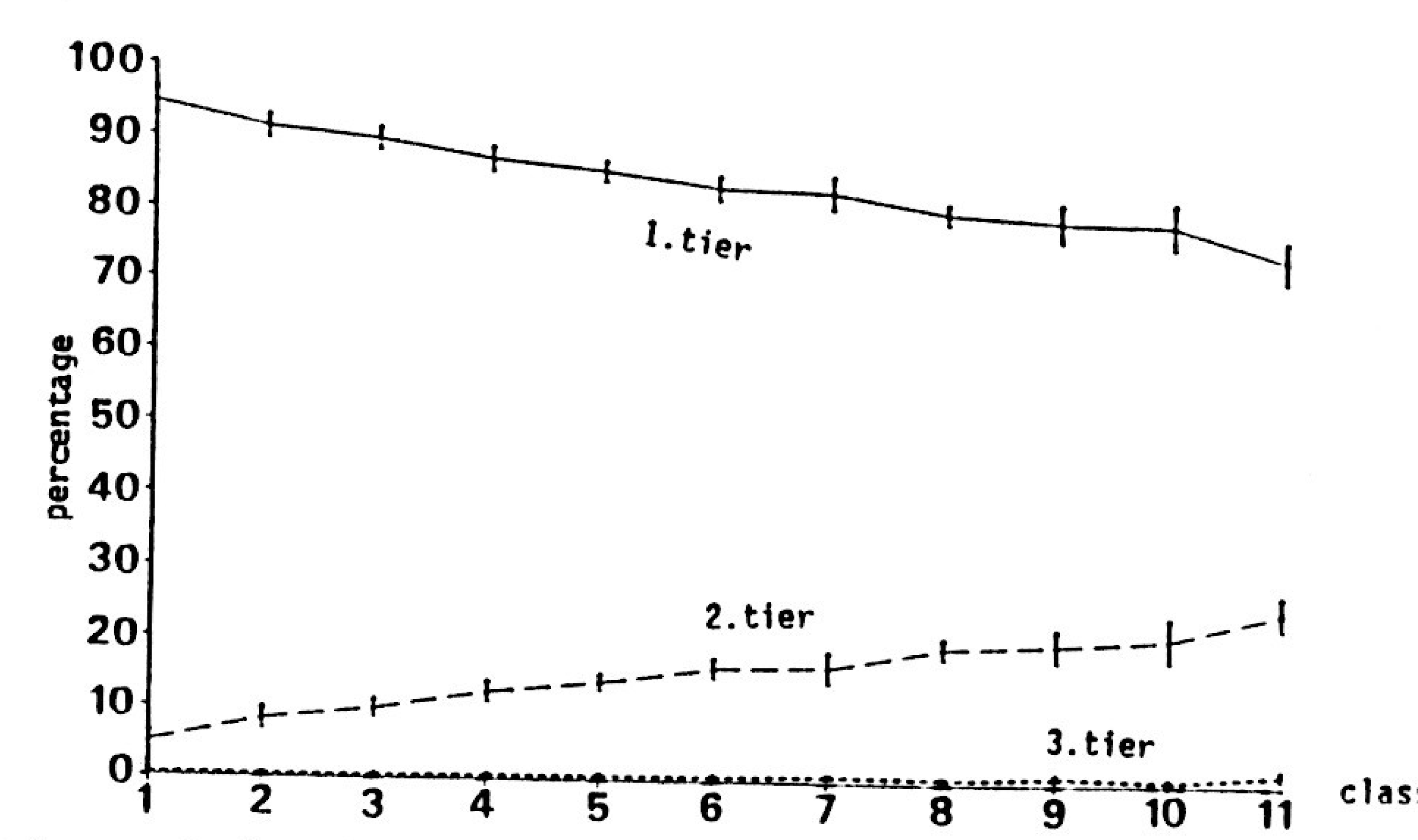


Figure 5.4: Distribution of the number of hops between mediately and immediately affected stations

Figure 5.4 presents results for cases with loss of radio contact and shows that 80%-95% (solid line) of the mediately affected stations are only one-hop (tier 1) away and, therefore, are easily to inform by the immediately affected stations through one-hop bc-packets. The curves representing two- and three-hop stations show with increased number of mediately affected stations an increasing probability of the respective event. It must be remembered however, that the absolute probability of many mediately affected stations is very small (see figure 5.3). Such stations could be informed of the event by extending the update range around the immediately affected stations to two hops. In case of a new radio contact the results are quite similar as in figure 5.4.

To summarize, it was observed that most topological changes only affect a few stations and that these stations are with high probability in the 1st and with low probability in the 2nd tier around the immediately affected stations.

5.3 PIGGY-BACKED UPDATING OF ROUTING TABLES TO SUPPORT DECENTRAL ROUTING FOR MOBILE STATIONS AND DIMENSIONING OF TIME-OUTS

5.3.1 ASYNCHRONOUS ROUTING TABLE UPDATE

In what follows, an asynchronous routing table updating algorithm is introduced, which works absolutely decentrally, does not need any additional packets and guarantees a remarkable high ratio of correctly routed packets. To support the algorithm, each source station does include into each packet a neighbor-map where its actual neighborhood is represented. All relay stations, which route the packet or other stations which hear any transmission, get knowledge of this piggy-backed actual neighborhood of the source station. The major result is that stations, which are mainly used to often relay packets are in control of a quite actual routing information.

Just these are the stations, which need to know of any topological change, because they might be mediately ffected! There exist various probabilities to keep the olume of this piggy-backed information sufficiently small, dependent on the application considered. E.g. for a network of cars moving on a highway using multi-hop packet radio to communicate with some environmental equipment (beacon) it suffices to know where other cars are situated instead of knowing their adress.

5.3.2 TIME-OUT FOR ASSUMING LOSS OF RADIO-CONTACT

A topological change, detected by an immediately affected station, results either from gain or loss of a radio contact to a neighbored station. The gain of a radio contact is detected by "hearing" a new station. A loss of a radio contact must be supposed if a neighbored station was not active during a pre-determined time-out interval. To keep the management oriented traffic volume small we decided not to use explicit "hello"-packet to prove the existence of a silent neighbor. To dimension an appropriate length of the silence-period, the frequency of a station hearing its neighbor was established, depending on the expired time. For this purpose 5 consecutive time intervals were defined, each comprising 30 slots, namely slots 1-30, 31-60, 61-90, 91-120, >120. Figure 5.5 shows the resulting distribution of silence-periods between radio ontacts of neighbored stations. Two source rates, namely q=0.005 and q=0.01 were investigated.

For both source rates q considered stations have radio contact with 85% probability after at most 30 slots. Larger silence-periods are relatively seldom.

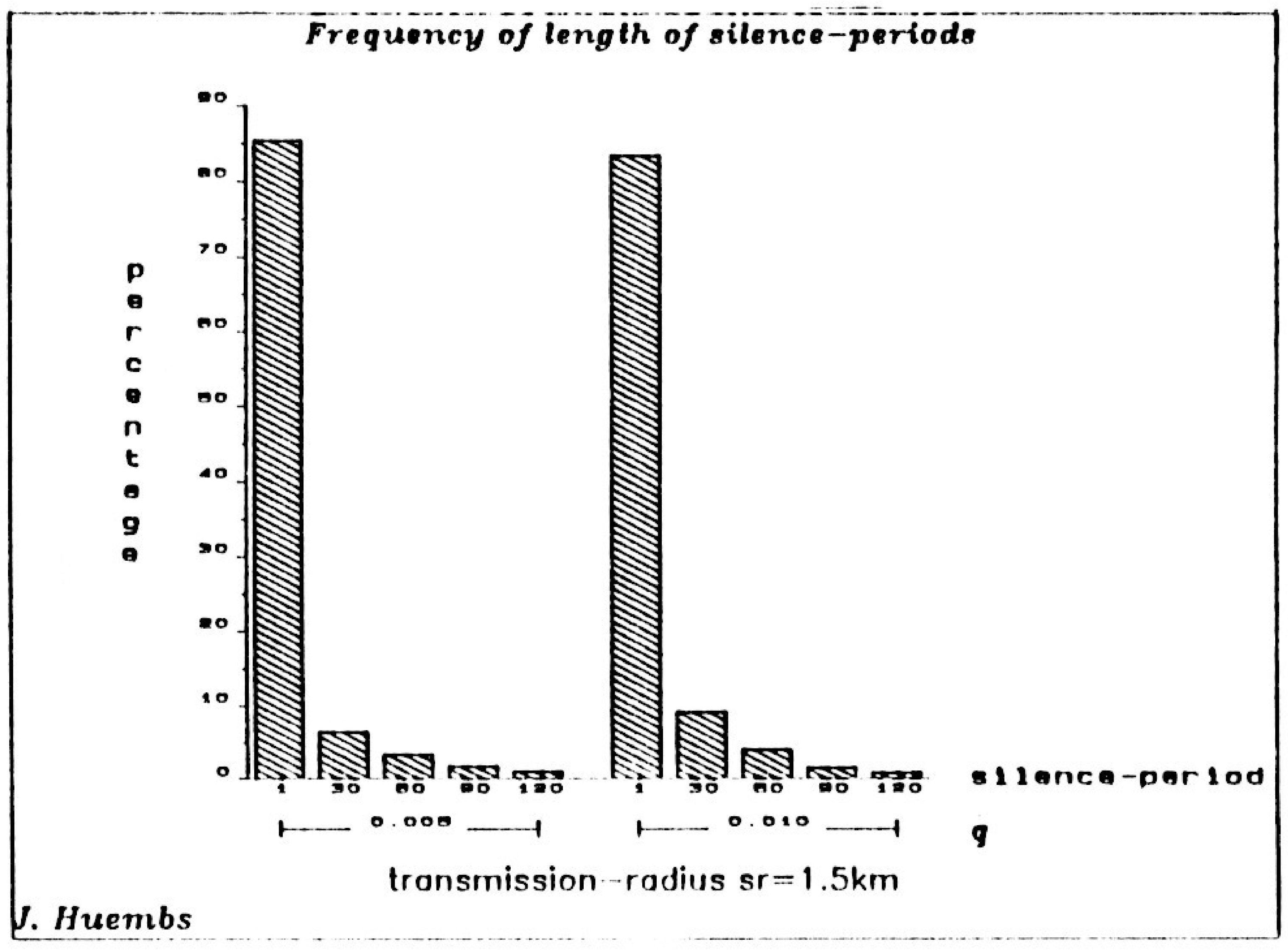


Figure 5.5: Frequency of length of silence-periods between neighbored stations

Furthermore, it was ascertained how often a station attempted to transmit to a no longer existing neighbor depending on the length of the time expired since the last radio contact (see figure 5.6). Here, in addition, the speed of station's movement is considered as a parameter. In case of a small admitted silence-period loss of a radio contact is supposed very soon and packets are seldomly transmitted to no longer existing neighbors. A higher speed generally causes more topological changes per slot and so the number of misrouted packets increases. The frequencies of misrouted packets increase when the admitted silence-period is increased, because the loss of radio contact is detected proportionally later. E.g. at a speed of 90km/h, the misrouting frequency increases from 1.5% (silence-period: 100 slots) up to 10% (silence-period: 600 slots). If the stations only move at 30km/h, the values increase more slowly and in case of a small silence-period nearly all losses of radio contact are detected in time.

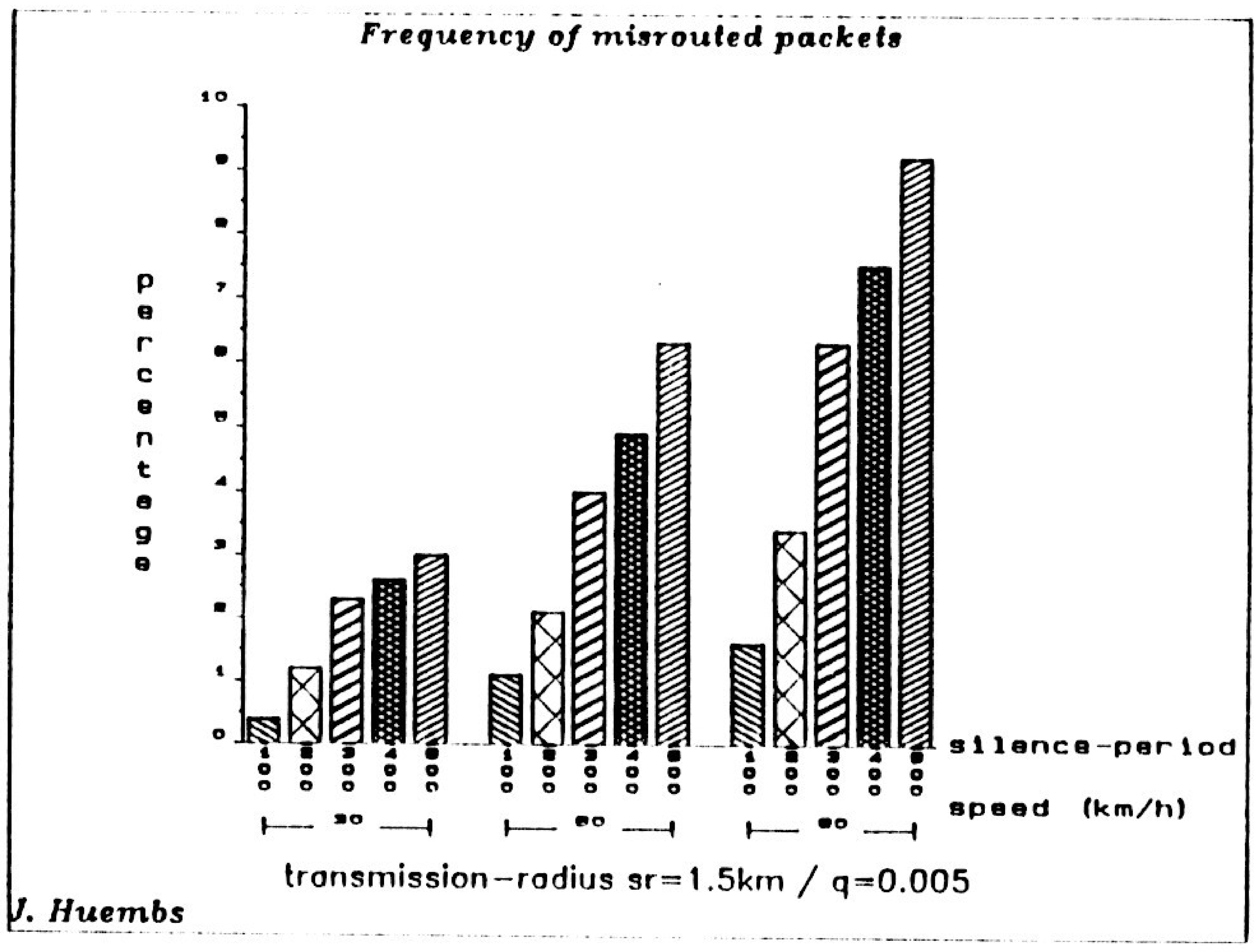


Figure 5.6: Frequency of misrouted packets, depending on the length of the admitted silence-period

In figure 5.7 the quality of our asynchronous update algorithm is demonstrated by showing the number of optimal routed packets depending on the length of the silence-period. Two different speeds of movements were considered and it was found out that 80-97% of packets followed the shortest (optimal) route. It was also found out that packets being not optimally routed made a detour of only 1.4 hops.

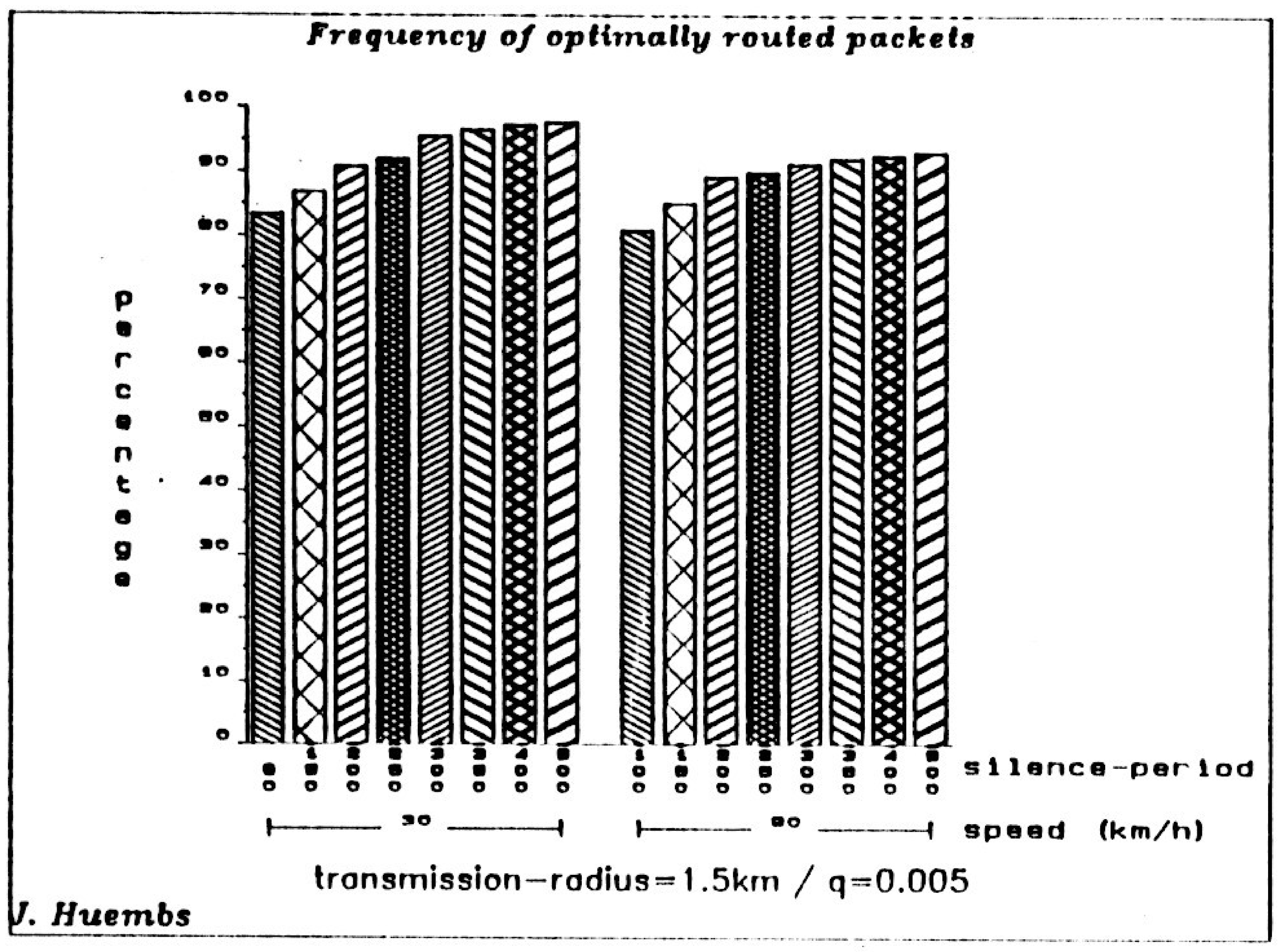


Figure 5.7: Frequency of optimally routed packets, dependent on the length of the silence-period

In case of a large silence-period the percentage of optimally routed packets is impressively high. Otherwise a loss of radio contact is assumed too early and then a relay doesn't compute an optimal route.

These results support the conclusion that our asynchronous routing-table update algorithm is suitable not only for designing a survivable network but also insures good performance measures. The silence-period is an important parameter. If it is less than 200 slots, the number of misrouted packets is too high; if it is more than 300 slots, the ratio of transmissions to no longer existing neighbors becomes too high and no essential increase in optimally routed packets can be gained.

6. SUMMARY

A method to very efficiently distribute bc-packets in a multi-hop packet radio network with stations communicating point-to-point is investigated by simulation. Acknowledgements (ACK/NAK) for a fixed number of past slots are piggy-backed to each packet. Simulation results are presented for the resulting mean packet delay to reach stations which are in distance of 1-, 2- or 3-hop, respectively. The influence of mobility of stations on the mean number of stations affected by a topological change and their distance in hops between immediately and mediately affected stations is analyzed. It is concluded that, due to the quite local influence of topological changes on multi-hop routes, a completely decentral routing of packets should work well. Further a new asynchronous algorithm to update the routing tables of stations is defined and evaluated, which has a considerable small demand on channel capacity. Its effectiveness is demonstrated by a very high percentage of optimally routed packets. Because of a number of significant parameters, the system is very complex. Therefore, not only the presented simulation results should be taken to notice but also the proposed network management algorithms. These algorithms hopefully can be transferred to other related network applications with a similar success. Our continuing research will focus on the optimization of the time-out parameter for implicit and explicit ACK, a dynamic adaption of the transmission probability p dependent to the local queue length and further improvements to the asynchronous update algorithm. In doing so the performance and applicability of the proposed management protocols will be developed further.

The authors gratefully acknowledge the contributions made by Joerg Huembs and Nikos Papadopulos who developed part of the simulation programs and carried out numerous simulation experiments as part of their diploma theses.

REFERENCES

/1/ V. Brass, C. Gotthardt: "On Throughput and Delay-time in S-ALOHA Multi-hop Systems," in Informatik Fachberichte 154, (Ed. U. Herzog) Springer-Verlag Berlin/Heidelberg, Sep. 1987, pp 236-249.

/2/ M.Y. Elsanadidi, W.M. Chu: "Study of Acknowledgement Schemes in a Star Multiaccess Network," IEEE Trans. on Com., Vol. COM-30, July 1982, pp 1657-1664.

/3/ M.Y. Elsanadidi, W.M. Chu: "Simulation study of the behaviour of multi-hop broadcast networks," Computer Science Department, UCLA 90024 (presented in: ACM, Communications Architectures and Protocols, SIGCOM 1983 Symposium, pp 170-177.

/4/ F.A. Tobagi, L. Kleinrock: "The effect of acknowledgement traffic on the capacity of packet switched radio channels," IEEE Trans. on Com., Vol. COM-26, June 1978, pp 815-826.

/5/ R. Briechle, B. Walke: "A Local Cellular Radio Network (LCRN) for Digital Voice and Data Transmission at 60 GHz," presented in: Cellular Communications International, 5-7 Nov. 1985, Online Conferences Ltd., London (Nov. 1985)