Multi-hop Cellular Networks: The Architecture and Routing Protocols

Ananthapadmanabha R., B. S. Manoj and C. Siva Ram Murthy
Department of Computer Science and Engineering
Indian Institute of Technology, Madras 600 036, India
ananthar@dcs.iitm.ernet.in, bsmanoj@rts.iitm.ernet.in, murthy@iitm.ernet.in

Abstract— **Multi-hop Cellular Network (MCN) is an architecture recently proposed by Lin and Hsu for wireless communications. MCNs combine the benefits of having a fixed infrastructure of base stations and the flexibility of ad-hoc networks. They are capable of achieving much higher throughput than current cellular systems, which can be classified as Single-hop Cellular Networks (SCNs). This work concentrates on MCNs and SCNs using the IEEE 802.11 standard for wireless LANs. We provide a general overview of the architecture and the issues involved in the design of MCNs, in particular the challenges to be met in the design of a routing protocol. We extend the work of Lin and Hsu to enhance the throughput of such networks further. We propose a routing protocol for use in such networks. We conduct extensive experimental studies on the performance of MCNs and SCNs under various load conditions (both TCP and UDP). These studies clearly indicate that MCNs with the proposed routing protocol are a viable alternative for SCNs, in fact they provide much higher throughput.**

Keywords— **Cellular networks, packet radio, IEEE 802.11 DCF, ad-hoc networks, multi-hop cellular networks, routing protocol**

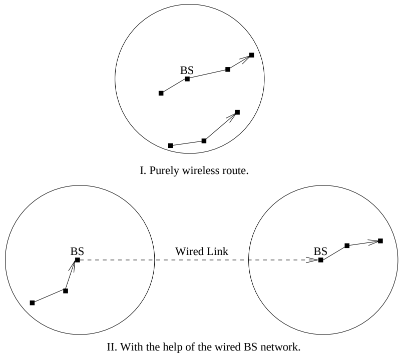

Fig. 1. The different possibilities in establishing routes between mobile stations in an MCN.

I. INTRODUCTION

Over the past decade, both computing and network resources have become inexpensive and at the same time their capabilities have increased manifold. The easy availability of processing power and bandwidth has made possible a variety of applications such as streaming multimedia content (voice and video) delivery and remote access to large information databases and files at low costs. But even today, such demanding applications are confined largely to stationary systems and there is a huge mismatch between the bandwidth availability in wired and wireless networks. The capacity of optic fibers remains largely untapped and hence there is tremendous potential for higher data rates with the development of techniques like Wavelength Division Multiplexing (WDM). Unfortunately, the availability of wireless bandwidth is not expected to increase by much, and the gap will only widen.

In the last few years, research on another kind of wireless network, commonly referred to as ad-hoc or packet radio networks has proceeded independently. Here, Mobile Stations (MSs) can forward packets from other stations to reach the destination in multiple hops. They are particularly attractive because of their low cost of deployment. But such networks do not scale very well and hence their application is limited to very specialized areas like rescue operations, battle fields and traveling groups. As of now, it is very difficult to provide uninterrupted high bandwidth connectivity to a large number of users with these networks and hence they cannot be easily used for Personal Communication Systems (PCS) and the like. Though the limited bandwidth and QoS issues are being addressed in systems such as WAMIS (Wireless Adaptive Mobile Information System), there is still a long way to go before one can even suggest them as a replacement for cellular networks. Initially, service providers could use MCNs for data services and continue with GSM/AMPS for voice calls. Thus, MCNs fit very well into the current state of technology, and allow seamless and easy migration. The architecture, Multi-hop Cellular Network (MCN) originally defined by Lin and Hsu [1] recently comes from a combination of ideas from the two separate directions of research mentioned above. In MCNs, MSs are allowed to communicate without the involvement of the Base Station (BS). The MSs also serve as packet forwarding agents in communication as in ad-hoc networks. The different possible cases for end-to-end transmission in MCNs are summarized in Figure 1.

MCNs provide a method to increase the spatial reuse of the channel by decreasing the transmitting power of MSs without having to pay the penalty for a large number of BSs. The advantages of MCNs over Single-hop Cellular Networks (SCNs) include more frequency reuse, higher tolerance to BS failures, higher scalability than ad-hoc networks and lesser contention and fewer collisions if we use a contention based MAC protocol. MCNs can provide service in ‘dead spots’ in a cell, which are not reachable by the BS in a single hop. They also handle ‘hot spots’ better. Typical examples would be a stadium, a stock market or a conference, where a large number of users assemble and choke the capacity of the cell. Quite a few of these calls may be between devices in the same area, and routing them through the BS is very wasteful.

The subsequent sections of this paper are organized as fol-

Author for correspondence: C. Siva Ram Murthy, e-mail: murthy@iitm.ernet.in

0-7803-7244-1/01/$10.00

lows. In Section II we discuss issues at each layer in the design of MCNs. Section III describes an architecture and a routing protocol for MCNs. We also present some methods to improve the architecture and its throughput further. In Section IV, we present the results from our simulation studies. The behavior of MCNs is studied extensively in terms of various metrics. Finally, conclusions and directions for future research are presented in Section V.

II. MULTI-HOP CELLULAR NETWORKS: SOME ISSUES

A. MAC Protocols

MCNs are more demanding on MAC protocols than SCNs. In SCNs, the BS is involved in *every transmission* of a frame, either as the transmitter or as the receiver. This considerably simplifies the channel assignment, and MAC protocols as simple as slotted ALOHA (used by GSM for access requests) prove to be adequate. But MCNs are closer to ad-hoc and Re-configurable Wireless Networks (RWN) with respect to MAC layer issues. The *spatial fairness* that the MAC protocol provides is to be looked into too.

Although it is possible to provide connection-oriented services in MCNs by partitioning the available bandwidth into channels as in SCNs, in this paper, we use one single channel, packet switched, contention based MAC protocols. The scalability requirements for MCNs are to be met. We have used protocols that have been categorized by Garcia and Fullmer [2] as the Floor Acquisition Multiple Access (FAMA) class of protocols. These protocols are characterized by the use of Request To Send (RTS) and Clear To Send (CTS), also known as virtual carrier sensing, in addition to real carrier sensing. The 802.11 DCF [3] appears to be by far the most mature of these protocols, with low cost hardware implementations readily available off the market.

B. Topology Discovery

The current methods for topology discovery used in ad-hoc networks are typically based on the use of hello beacons which are transmitted periodically. But this method does not scale very well, especially when the density of nodes[1] is very high. When a large number of stations fall within the capture area of a given station, the possibility of the hello beacons themselves colliding and getting lost becomes very high. This can have serious consequences because routing cannot be made robust or efficient without up-to-date information about the location of nodes. Thus, there is a trade-off between the frequency at which neighbors can be notified of the nodes' presence and the bandwidth that is taken up by these hello packets.

C. Mobility Management, Registration and Hand-offs

The interaction between the mobile nodes and the BS is significantly different in MCNs. Therefore, some of the solutions used for registration with the BS, maintenance of location databases etc., do not directly apply to MCNs. The assignment of channels coupled with a good hand-off scheme is perhaps the greatest of the challenges in the design of MCNs. The hand-offs will not only happen between two BSs, but more often between an MS and a BS or from one MS to another.

These problems are greatly simplified in the case of packet radio services. We have decided to stick with connection-less service for two reasons. Firstly, there are a large number of applications for IP services even in the wireless domain. More importantly, considering recent advances in time-bounded services with QoS in ad-hoc networks, we feel that it is realistic to assume that soon it will be possible to support real-time multimedia traffic on this architecture.

D. Routing

Routing in SCNs is fairly simple and clearly does not extend to MCNs. The plethora of routing protocols proposed for ad-hoc networks also fail to provide good solutions for MCNs because they fail to exploit the presence of the highly reliable BSs and the wired backbone. They do not scale very well to a large number of nodes. Ad-hoc routing protocols are mostly aimed at providing stability against path vulnerability, and often trade-off throughput against stability. Paths are much less vulnerable in MCNs [1] and stability of paths is not of major concern. The division of routing functions between the mobile nodes and the BS plays a crucial role in the performance of the routing protocol. The driving principle behind the routing protocol would be to increase the spatial reuse of the channel and thereby increase the throughput.

III. ARCHITECTURE, ROUTING AND ENHANCEMENTS

A. Architecture

Consider cells of radius R, with the BS at the center. Every node transmits data packets with a transmission range of r. We define $k = R/r$. Thus, nodes that do not fall in the capture area of the BS will have to transmit packets over multiple hops to the BS. But in a typical cellular network, a lot of control packets have to be transmitted to the BS and received from the BS. While one solution to this problem could be to flood these control messages such as registration messages, this is counter intuitive to the goal of achieving higher throughput. Also, placing a hard limit on the transmission power will break the system when forwarding nodes are not present.

It is an accepted fact that in cellular systems of the future, the radius of the cells is not bounded by the maximum transmission range of a typical mobile device, but rather by the amount of traffic a cell can support. Therefore, it is reasonable to assume that the transmitters on the mobile devices are capable of transmitting data with range R. So, we have chosen to provide a separate control channel using which nodes exchange control packets with the BS over a single hop. However, the limit on transmission power for the data channel can be placed dynamically by the BS by broadcasting messages on the control channel. This provides a lot of flexibility wherein the MCN can operate as an SCN when the BS thinks it more appropriate to do so. A typical case could be when the number of nodes in the cell is very low, and a lower transmission range will make the network disconnected. But we will not consider such situations in the rest of this paper.

The presence of such a control channel has a lot of advan-

---

[1] We have used the terms "node" and "(mobile) station" interchangeably.

tages for routing. These will become obvious when we describe the routing protocol. A disadvantage of this channel is that we cannot solve the problem of 'holes' in the coverage of a BS. But it should be possible to route control packets over multiple hops when a mobile station is not in capture area of any BS, as an exception. It is just that we expect the bulk of the stations to be able to communicate with the BS over one hop. We do not discuss such exceptions in what follows.

In the architecture described by Lin and Hsu [1], each cell is given a separate data channel. While this does simplify the analysis, it definitely does not lead to optimal use of the available bandwidth. This is especially true of MCNs when the transmission range is reduced to r, which can be much less than the cell radius R. Thus the transmissions in the circle of radius $(R - r)$ centered at the BS *absolutely do not interfere* with transmissions in *any other cell*. The policy of not reusing the channel in an adjacent cell is important in the case of circuit switched, contention free MAC protocols used in cellular telephony, but it does not make much sense in packet switched MCNs. Therefore, we promote sharing of the data and control channels in the entire service area of the network.

B. Registration

This problem is made trivial by the presence of the control channel. Each BS transmits hello beacons periodically. Nodes choose the BS from which the strongest signal is received, and send a registration request (*RegRQ*) to that BS. The BS then transmits an acknowledgment (*RegACK*) to the node. This exchange is very similar to the method used in GSM systems.

C. Neighbor Discovery

The problems associated with announcing the presence of nodes was discussed in the previous section. We now describe a *contention and collision free* protocol for this purpose. There is a separate channel for hello beacons, and this channel is again subdivided into N_{max} channels. Each of these channels have enough bandwidth for one mobile station to transmit hello beacons at a frequency $1/T$. These divisions of the channel could be in principle done in time, frequency or code. But, for the purposes of illustration, we assume that the data, control and neighbor discovery channels use different frequency bands, and that the neighbor discovery channel is divided into time slots.

The assignment of these channels is coordinated by the BSs and the assigned channel is notified to the mobile station as part of the *RegACK*. This scheme imposes a hard limit of N_{max} on the number of mobile stations that can be serviced. In other words, if a mobile station arrives at a cell, and every channel is used by a mobile station in this cell or one of its 6 neighboring cells, the new station cannot be assigned a channel. Given any region of 7 neighboring cells, not more that N_{max} nodes can be supported. However, it is easy to see that N_{max} can be made arbitrarily large without wasting too much bandwidth. Assume that a hello beacon consisting of a node identifier and a checksum is 10 bytes.[2] Now, $N_{max} = 5000$ and $T = 5s$ will require a bandwidth of 80Kb/s, which is not very significant. In any system where nodes are expected to maintain an average transmission rate of at-least 100 b/s, the overhead for $T = 5s$ does not exceed 2 percent, irrespective of N_{max}.

This method provides a simple and reliable solution for neighbor discovery, and scales reasonably well. For networks where this limit of N_{max}, however high it may be, on the number of registered nodes is absolutely unacceptable, we may use other schemes if they satisfy the requirements better, without changing the rest of the routing protocol.

D. MAC Protocols

We have used the Distributed Co-ordination Function (DCF) of the IEEE 802.11 MAC specification [3] for the data channel in this study. This protocol is actually a CSMA/CA (Collision Avoidance) protocol, which uses *virtual carrier sensing*, *immediate positive acknowledgments* and a *slotted binary exponential back-off* for re-transmissions. Virtual carrier sensing is done through an RTS-CTS exchange before every transmission and the use of a Network Allocation Vector (NAV) timer at each transmitting station. Every successful transmission is a 4-way exchange consisting of the *RTS-CTS-Data-Ack* sequence. This protocol does not ensure collision free data transmission without interference from hidden terminals even in the absence of bit errors and mobility (see [4] for details). Sufficient conditions for the RTS-CTS exchange to ensure a collision free data transmission are discussed by Fullmer and Garcia [2]. Every node is required to back-off for the maximum allowed period of transmission of a single frame on hearing any collision. We have also incorporated this feature into the MAC protocol. The same protocol can also be used for the control channel.

E. Routing

Our routing protocol tries to assign as much responsibility to the BSs as possible, for reasons mentioned earlier in Section 2. The BSs are a natural choice for all kinds of databases, including the location database and topology information. Mobile stations access these databases at the BSs through packets sent over the control channel. The computation of routes is also done at the BSs.

E.1 Role of MSs and BSs

Every node is required to participate in the neighbor discovery protocol described above. Each node maintains a table of its neighbors based on the hello beacons it receives from every node in its capture area at a frequency $1/T$. These entries timeout with a period T_n $(> T)$. The other entries in the table include the current received power (Rx_p) and the received power last notified to the BS (Rx_{np}). Whenever there is appreciable difference between Rx_p and Rx_{np}, new entries are added or old ones removed, the node sends *incremental updates* to the BS. Thus the BS has a very up-to-date database about all the links in its cell.

When a node A has a packet to send to a node B, it sends a request for a route (*RReq*) to the BS it is registered to. The BS can then compute a suitable route using the information it has and send a reply(*RRep*) to the requesting node. The BS has to maintain the graph of the nodes in its service area, in some

---

[2] This is a conservative estimate, 5 bytes (4 for address and 1 for check-sum) should be sufficient. This should more than compensate for the synchronization overhead, etc.

suitable data structure. It is also required to exchange this information with the BSs in its adjacent cells. This enables us to compute purely wireless routes between stations that belong to different cells, but lie close to each other at the boundary of the cells. When a BS receives a *RReq* from node A to node B, it uses a shortest path algorithm to compute a suitable route and sends a reply (*RRep*).

If the Dijkstra's algorithm is used with a Fibonacci heap, this algorithm takes $O(n + mf(n))$ time, where n is the mean number of nodes in cell, m is the number of edges and the computation of $w(x, y)$ is $O(f(n))$. The BS has to service $O(n)$ requests in unit time. Hence the computational burden at the BS is $O(n^2 + nmf(n))$. Since different requests are independent, this can be parallelized very easily.

We can cache computed routes and re-compute them only if an update has been received after the route in the cache was computed. A mechanism to cache routes at the mobile stations is discussed in the next sub-section. These two methods can reduce the burden considerably and hence improve the scalability. If a large number of *RReq*s are for nodes outside the cell, it may be a good strategy to maintain a table of shortest routes to *all* the nodes from the BS. The *entire table* can be obtained by running Dijkstra's shortest path algorithm once, and this must be done every time an update is received.

E.2 Route Cache

The traffic due to *RReq* and *RRep* on the control channel can be greatly reduced by the use of *Route Cache* (RC) as in other ad-hoc routing protocols like DSR and AODV [5]. When a node receives a *RRep*, it adds the route to its RC. Whenever an IP packet is received, the recorded source route is reversed and this reversed route to the source of the packet is added to the RC. These entries in the RC time-out after time T_r and the *RReq* has to be sent again.

Broken routes in the RC can cause serious problems such as dropped packets, out-of-order delivery and wastage of bandwidth. The protocol must also support detection and updation of stale routes in the RC. We have used the following mechanism for this. When a route is broken, it has to be removed from the RC of the originating node. Consider the case when the link (X, Y) on a route from A to B is broken. When X receives a packet from A whose next hop is Y, it detects that the link (X, Y) is no longer available either because of a time-out of the hello beacon from Y or because of the maximum re-transmission limit of the RTS. Now, X will send an *RReq* to the BS on behalf of A. The network layer of X rather than dropping the packet, buffers it. The BS responds with a new route from A to B to A and from X to B to X. Thus, the RC at A is updated (see Figure 2). No packets are dropped as X can now empty its buffer and send all the mis-routed packets to B it received in the time it takes for the RC to be updated.

E.3 Edge Weights

We still have not answered the question as to how to compute the weights of the edges, $w(x, y)$ for wireless links. A very naive approach is to assign $w(x, y) = 1$ if y is in the capture area of x and zero otherwise. But with the extensive information available at the BS, we can do a lot better.

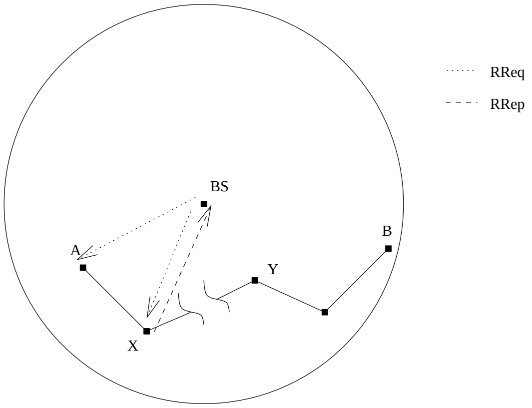

Fig. 2. Packets over the control channel that are used to update RC entries when a stale route is detected.

In order to come up with a good edge weight function, we first identify the nodes affected by the transmission from node x to y. All nodes in the capture area of x or y cannot transmit or receive until this transmission is over. If $N(x)$ denotes the set of neighbors of x, the number of nodes 'blocked' (cannot transmit nor receive) by this transmission is $|N(x) \cup N(y)|$. Therefore, we propose the edge weight function as,

$$w(x, y) = |N(x) \cup N(y)| + 2$$

Two has been added to count the nodes x and y, which are also affected by this transmission. If the list of neighbors are maintained in sorted lists, this can be computed in $O(|N(x)| + |N(y)|)$ time.

IV. SIMULATION STUDIES

A. Simulation Setup

The simulation engine we use is built on top of *GloMoSim*, developed at the University of California, Los Angeles using PARSEC. The radio layer assumes free-space propagation of signals. We have ignored the effect of capture[3]. The MAC layer tries to stick as closely as possible to the IEEE standard [3]. The simulation of the TCP and the UDP protocols are also exact, as provided by *GloMoSim*. All parameters of the DCF are as given by Direct Sequence Spread Spectrum (DSSS) specification of the IEEE standard [3], and we have assumed typical values for other parameters. We simulate packets of a fixed length of 2000 bytes, and the bandwidth of the control and data channels are 1Mb/s and 5Mb/s respectively. We fix $k = 2$ and $R = 500m$.

B. UDP Load

In UDP simulations, we have reduced B_{data} to 2 Mb/s. The traffic is generated according to the Poisson distribution at each node and the destination for packets from the BS is chosen from a uniform distribution. The bandwidth of the data channel has been reduced to ensure that the control channel does not become the bottle-neck. In practice, when we use a route cache, and

---

[3] The ability of a node to receive from another node provided the interference is small compared to the signal.

packets are generated in *trains*, the same control channel can service a much larger data channel effectively. In SCNs, we find that a large number of packet transmissions (almost half) fail even after a CTS is received. Most of the data packet collisions happen from the scenario described in Section III.

We use random way-point mobility (with an average speed of 10m/s and an average pause of 30s). We consider two possible enhancements here. Firstly, we use the enhanced edge weight (EEW) function proposed in Section III, instead of the zero/one edge weight function. Secondly, we mentioned in Section III that the collision of data packets can be reduced by forcing a huge back-off at every node on hearing a collision. In Figure 3(i), we study the effect of these enhancements on the UDP throughput.

For simulation of traffic between nodes in the same cell, we define the parameter, locality of traffic, l as,

$$l = \frac{\text{traffic destined to nodes in the same cell}}{\text{total traffic generated in the cell}}$$

These results have been verified analytically also [4], but we are unable to present the analysis here due to space constraints.

C. TCP Load

TCP load is simulated using FTP sessions initiated between nodes based on a Poisson distribution. We have fixed $l = 1$ for these simulations and all the enhancements described previously have been used. Figure 3(ii) gives the throughput. The RC improves the throughput considerably.

D. Fairness

For the system to be fair, the throughput achieved by nodes should not depend on its position (i.e. its distance from the BS). This depends both on the MAC and the routing protocol. To study the spatial fairness of MCNs and SCNs, we have considered one point for every node from each run of the simulation, with its distance from the BS on the x-axis and the throughput achieved by it on the y-axis. We have plotted the 5th degree least square approximation polynomials for these points (Figure 3(iii)).

V. CONCLUSION

In this paper, we have given a general overview of the architecture, routing and other issues in MCNs. We have extended the architecture proposed in [1], and shown how the enhancements affect the achieved throughput. Routing in MCNs is indeed more complicated than in SCNs. This work recognized the challenges involved in coming up with a viable routing mechanism, and proposed a simple but effective routing protocol for such networks. The hop-by-hop and end-to-end throughput are obtained and these studies show that MCNs indeed provide much higher throughput than SCNs. These studies also give a thorough evaluation of the MAC protocol itself and its suitability for cellular networks. The importance of the traffic locality, l in the system throughput of MCNs is shown through simulations. Better MAC protocols can certainly improve the throughput and fairness of MCNs and SCNs considerably.

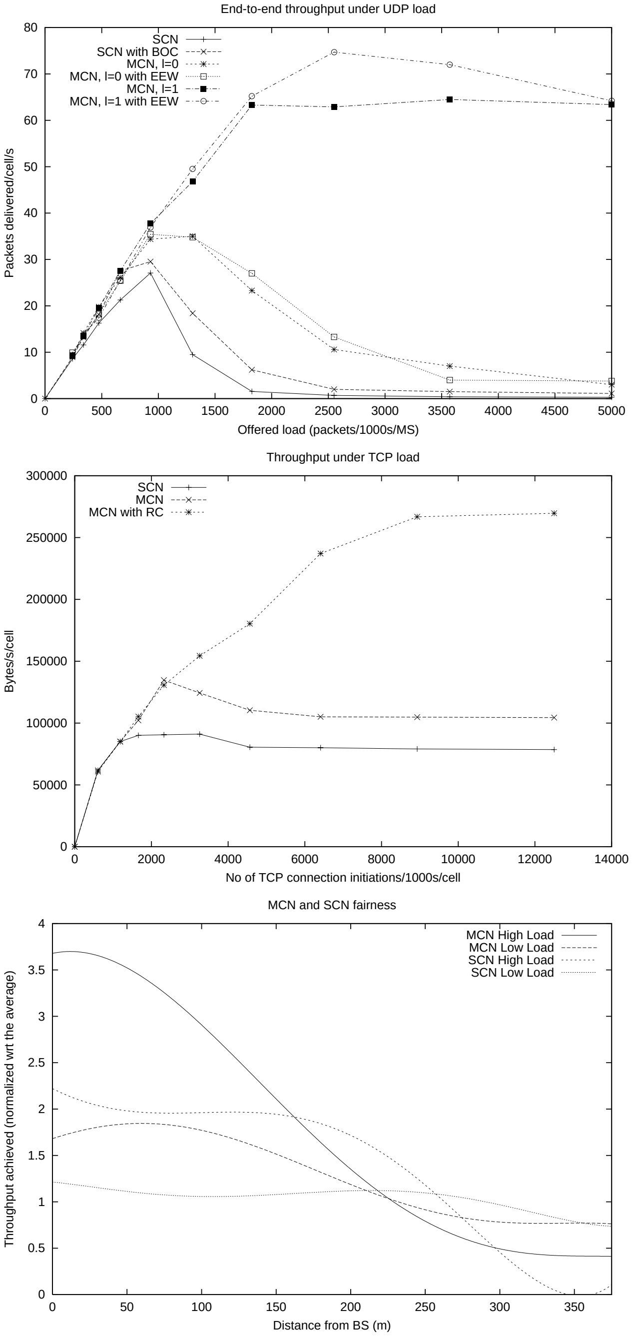

Fig. 3. Throughput under (i) UDP load and (ii) TCP load for MCNs and SCNs and (iii) throughput fairness of MCNs and SCNs (cell radius = 400m).

REFERENCES

[1] Ying-Dar Lin and Yu-Ching Hsu, “Multihop Cellular: A New Architecture for Wireless Communications,” in *INFOCOM 2000*. IEEE, 2000.

[2] C. J. Fullmer and J. J. Garcia-Luna-Aceves, “Floor Acquisition Multiple Access (FAMA) for Packet-Radio Networks,” in *SIGCOMM 1995*. ACM, 1995.

[3] IEEE Standards Board, *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, The Institute of Electrical and Electronics Engineers Inc., 1997.

[4] Ananthapadmanabha R., B. S. Manoj, and C. Siva Ram Murthy, “Multi-hop Cellular Networks: The Architecture and Routing Protocols,” Tech. Rep., HPCN Lab, Indian Institute of Technology, Madras, December 2000.

[5] Charles E. Perkins and Elizabeth M. Royer, “Ad-hoc On-demand Distance Vector (AODV) routing,” in *Second IEEE Workshop on Mobile Computing Systems and Applications*, 1999.