Extension of Basic Service Set by Multihop Routing in IEEE 802.11 Wireless LAN Standard

Mustafa Ergen
Department of Electrical Engineering and Computer Science
University of California Berkeley
Berkeley CA, 94720
Email: ergen@eecs.berkeley.edu

Abstract— **In this paper, a multihop routing extension to IEEE 802.11 is presented. Multihop extension preserves the traditional single-hop networks where the service infrastructure is constructed by many access points but it also adds the flexibility of ad hoc networks where wireless transfer through mobile stations is allowed in multiple hop. Multihop extension can reduce the number of required access points and improve the throughput performance. Protocol presented is in compatible with the DCF and PCF medium access methods of IEEE 802.11 standard.**

I. Introduction

Wireless Local Area Networks (WLANs) are designed in order to provide high speed wireless connections. IEEE 802.11 is standardized in 1997 and revised in 1999 [1]. The data rate goes up to 64Mbps in IEEE 802.11a. Although 802.11 covered locations are growing in size, it is still not enough for the nomadic users who seeks less disruption. Giving stations relaying mechanism and enabling multihop routing is one of the ways of increasing the covered area. IEEE 802.11 is designed single hop in the infrastructure mode where *mobile stations* connect to the *access point* (*APs*) directly however IEEE 802.11 allows to extend the protocol to multihop in PCF and DCF modes within the standard [1].

Multihop extension to the existing infrastructure *basic service set (BSS)* allows the stations away from the access point range connect to the Internet. This increases the number of operating stations in the BSS. Transmission power of mobile stations and APs can be decreased. Multiple packets can be simultaneously transmitted within a cell of the corresponding single hop case [4].

The throughput of a multihop extension is analyzed by modelling the packet departure process as a renewal process, in which the renewal point is defined as the time point when all stations in a sub-cell simultaneously sense that the channel is idle[4]. The results lead to three important observations. First, the throughput of multihop is superior to that of the corresponding single hop. Second, the throughput of multihop increases as the transmission range decreases[4].

Several approaches can be taken when multihop protocol is designed. Multihop protocol can be implemented in the network layer or in the MAC layer or in the physical layer.

In the network layer, a network layer multihop protocol uses an ad hoc routing protocol (DSDV, DSR,..etc) for routing in ad hoc domain and Mobile IP for seamless connectivity across subnets [2]. This type of protocol is efficient when the network size is large and movement is across subnets.

If the movement is within the same subset, IEEE 802.11 protocol is enough since it supports mobility in the subnet. A MAC layer multihop standard only uses IEEE 802.11 protocol specification and does not need any protocol above MAC layer. As a result, overhead of Mobile IP and ad hoc routing protocols are not present.

These two approaches above are considered as decoded relaying schemes where each intermediate station decodes and re-encodes the received signal from the immediately preceding terminal before retransmission. When the multihop protocol designed in the physical layer, diversity and undecoded relaying can be leveraged wherein the former transmissions of the same signal from multiple transmitters can be used to estimate the signal on the other hand in the latter, analog signal coming from a transmitter is amplified and transmitted without decoding [3].

In order to consistent with the standard, decoded relaying method is necessary. It also enables access point to know, manage and re-configure the network according to some load metrics.

The organization of the paper is as follows; I compare the probability of error between multihop and singlehop case in Section 2. I give a brief background information about IEEE 802.11 protocol in Section 3. I describe the multihop extension in Section 4. I introduce a link test method in Section 5 and conclude the paper in Section 6.

II. Physical Layer Analysis

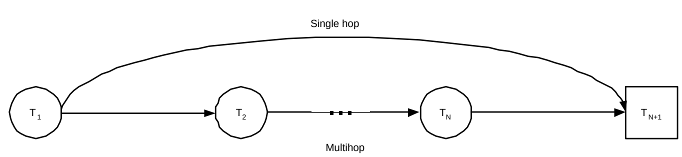

Fig. 1. A Multihop Wireless Communication Channel

The system model for multihop wireless communication channels is composed of a source terminal, a receiving terminal, and an intermediate relaying terminals[3]. In Figure 1 according to Table II the source terminal is identified as T_S. T_I represent the set of intermediate terminals, and T_D represent the set of destination terminals. Therefore, T_R represents the set of all receiving terminals. $T_{P(i)}$ represents the set of terminals that transmit a signal received by receiving terminal T_i.

Research supported by National Semiconductor.

TABLE I
SYSTEM REPRESENTATION FOR MULTIHOP AND SINGLE HOP

	Multihop	Single hop
T_S=	$\{T_1\}$	$\{T_1\}$
T_I=	$\{T_2...T_N\}$	$\{\ \}$
T_D=	$\{T_{N+1}\}$	$\{T_{N+1}\}$
T_T=	$\{T_1...T_N\}$	$\{T_1\}$
T_R=	$\{T_2...T_{N+1}\}$	$\{T_{N+1}\}$
$T_{P(i)}$=	$\{T_{i-1}\}$, i≥2	$\{T_1\}$, i=N+1

Each terminal T_i transmits a signal with complex baseband amplitude given by

$$s_i = \sqrt{\varepsilon_i}a + \beta_i, \tag{1}$$

where ε_i is the transmitted power, a is the information symbol during a particular signalling interval, and β_i is propagated noise. The propagated noise term in (1) is zero for source terminals as well as for intermediate terminals since we decode and encode at each T_I.

Assuming flat fading, each terminal T_i then receives a set of signals with complex baseband amplitudes given by

$$\begin{aligned} r_{k,i} &= \alpha\sqrt{(L_{k,i}/d_{k,i}^p)}R_{k,i}(\sqrt{\varepsilon_k}a + \beta_k) + z_{k,i}, \\ T_k &\in T_{P(i)}, \end{aligned} \tag{2}$$

where α^2 is the free space signal power attenuation factor between the transmitting terminal and an arbitrary reference distance, $d_{k,i}$ is the inter-terminal distance relative to the reference distance, p is the propagation exponent, $L_{k,i}$ is a zero-mean lognormal random variable with variance $\sigma^2_{L_{k,i}}$, $R_{k,i}$ is a complex gaussian (Rayleigh) random variable with mean power $E[R^2_{P(i),i}] = 1$, and z_k, is a zero-mean additive white gaussian noise random variable with variance N_o.

Assuming maximal ratio combining, the received signal to noise ratio at T_i is given by

$$\gamma_{P(i),i} = \sum_{k\in P(i)} \left(\frac{\alpha^2\varepsilon_k|a^2|}{(d_{k,i}^p/L_{k,i}|R_{k,i}|^2)N_o + |\beta_k|^2}\right), \tag{3}$$

where $|R_{k,i}|^2$ is an exponential random variable with mean $2\sigma_{R^2_{k,i}} = 1$.

The probability of outage due to lognormal shadowing when $\beta_k = 0, T_k \in T_{P(i)}$ is given according to the method in [8] by

$$Pr[\tilde{\gamma}_{P(i),i} < \gamma] \approx Q\left(\frac{10log(\overline{\gamma_{P(i),i}}/\gamma)}{\sigma_{L_{z(i)}}}\right), \tag{4}$$

where $\tilde{\gamma}_{P(i),i}$ is the instantaneous received signal to noise ration at T_i averaged over the Rayleigh fading, $\overline{\gamma_{P(i),i}}$ is the expected received signal to noise ratio at T_i averaged over the lognormal shadowing and Rayleigh fading, $\sigma^2_{L_{z(i)}}$ is the variance of the lognormal approximation of (3) determined by Wilkinson‘s method [9], γ is an arbitrary threshold signal to noise ratio that must be maintained at every decoder in order to maintain communication, and $Q(x)$ is $\frac{1}{2}erfc(\frac{x}{\sqrt{2}})$.

The calculation of probability of error is dependent on the modulation scheme employed. For the special case of BPSK, the probability of error under fading conditions when $\beta_k = 0, k \in P(i)$ is given in [8] by

$$P_e(\gamma_{P(i),i}) \simeq \left(\frac{2K-1}{K}\right)\prod_{k\in P(i)}\left(\frac{1}{2\overline{\gamma_{k,i}}}\right), \overline{\gamma_{k,i}} >> 1, \tag{5}$$

where $\overline{\gamma_{k,i}}$ is the expected received signal to noise ratio at T_i for branch k of the diversity combiner and K is the cardinality of $P(i)$.

The channel model for the multihop is given by (1) through (3) with $\beta_k = 0$, $T_k \in T_{P(i)}$. The received signal to noise ratio at terminal T_i is given by

$$\gamma^m_{P(i),i} = \left(\frac{\alpha^2\varepsilon_{i-1}|a|^2}{(d_{i-1,i}^p/L_{i-1,i}|R_{i-1,i}|^2)N_o}\right), \tag{6}$$

The total probability of outage for multihop is given by

$$P_o^m = 1 - \prod_{T_i\in T_R}(1 - Pr[\tilde{\gamma}^m_{P(i),i} < \gamma]), \tag{7}$$

where $P[\tilde{\gamma}^m_{P(i),i} < \gamma]$ is the probability of outage at terminal T_i given a received signal to noise ratio of $\gamma^m_{P(i),i}$.

The total probability of decoding error for the multihop is approximated by

$$P_e^m \approx 1 - \prod_{T_i\in T_R}(1 - P_e(\gamma^m_{P(i),i})), \tag{8}$$

where $P_e^m(\gamma_{P(i),i})$ is the probability of decoding error at terminal T_i given a received signal to noise ration of $\gamma^m_{P(i),i}$. This approximation does not take into account the probability of an even number of individual hop decoding errors in sequence resulting in correct decoding at the destination terminal.

On the other hand, for the single hop case, the received signal to noise ratio at terminal T_{N+1} is given by

$$\gamma^s_{P(N+1),N+1} = \left(\frac{\alpha^2\varepsilon_o|a|^2}{(d_{1,N+1}^p/L_{1,N+1}|R_{1,N+1}|^2)N_o}\right), \tag{9}$$

The total probability of outage for single hop is

$$P_o^s = Pr[\tilde{\gamma}^s_{P(N+1),N+1} < \gamma], \tag{10}$$

and the total probability of error for the single hop case is approximately by

$$P_e^s \approx P_e(\gamma^s_{P(N+1),N+1}), \tag{11}$$

In order to provide a fair comparison with the reference channel, the transmit powers at each terminal are constrained so that the sum of the powers at each hop in multihop is equal to the reference power ε_0 in single hop, $\varepsilon_0 = \sum_{T_i\in T_T}\varepsilon_i$. For the decoded relaying multihop channel, the optimal power distribution based on the upper bound is given by

$$^1\varepsilon_{P(i)} = \frac{\varepsilon_0\sqrt{d^p_{P(i),i}/L_{P(i),i}}}{\sum_{T_j\in T_R}\sqrt{d^p_{P(j),j}/L_{P(j),j}}}, \tag{12}$$

[1] For the proof of (12) refer to [3].

For the decoded relaying multihop diversity channel, the solution to this minimization problem is not as tractable. Although (12) is convex with respect to the transmit powers at each terminal and the solution for any finite number of hops can be calculated using Lagrange multipliers, the resulting equations are extremely complex and not easy to generalize [3].

III. Background

IEEE 802.11 standard defines a network which is composed of *Basic Service Sets (BSSs)* that are interconnected with a *Distribution System (DS)*. *Stations (STA)* in a *BSS* gain access to wired network through an *Access Point (AP)*. Each *BSS* has an *BSS ID* that distinguish them from other *BSSs*. A *STA* access the wireless medium via associating with the *AP* which it is already authenticated. A *station* can be associated with only one *AP* but it can be authenticated with more than one *AP*. The *DS* supports mobility and the integration of *BSS's* in a manner that is transparent to stations.

The basic access method in the IEEE 802.11 MAC protocol is the *Distributed Coordination Function (DCF)* which is known as *Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA)*. In addition to the *DCF*, the standard also incorporates an alternative access method known as the *Point Coordination Function (PCF)*. *PCF* is an centralized scheme which uses a *point coordinator PC* (usually *AP*) to determine which station has the right to transmit.

In *DCF*, a station, before starting its transmission, senses the channel to determine if another station is transmitting. The station proceeds with its transmission if the medium is determined to be idle for an interval that exceeds the *Distributed InterFrame Space (DIFS)*. In case the medium is busy the transmission is deferred by the station until the end of the ongoing transmission. A random interval, henceforth referred to as the *backoff* interval is then selected. The *backoff timer* is decremented only when the medium is idle; it is frozen when the medium is busy. Decrementing the *backoff timer* resumes only after the medium has been free longer than *DIFS*. To reduce the probability of collisions, after each unsuccessful transmission attempt, the *backoff* time is increased exponentially until a given maximum is reached. IEEE 802.11 frames contain a *Duration* field which is for STA to set its *Network Allocation Vector* to the duration specified in the frame in order to suspend its transmission for a period until the medium is busy.

After a successful reception, the standard describes an acknowledgement frame after *Short InterFrame Space (SIFS)* which is less than *DIFS* immediately following the reception of the data frame.This access method is summarized in Figure 2.

The *PCF* is built using the *DCF* through the use of an access priority mechanism that provides synchronous or asynchronous data frames contention free access to the channel. In this case, contention and contention free periods alternate with each other as shown in Figure 3. A *contention free period (CFP)* (during which *PCF* is active) and the following *contention period (CP)* (during which *DCF* is active) alternate. The *Point Coordinator (PC)* when it senses the channel idle seizes the control of the channel by transmitting after it (the channel) has been idle for a time interval *Priority Interframe Space (PIFS)* that is chosen to be smaller than *DIFS* but larger than *SIFS*. The transmission of a

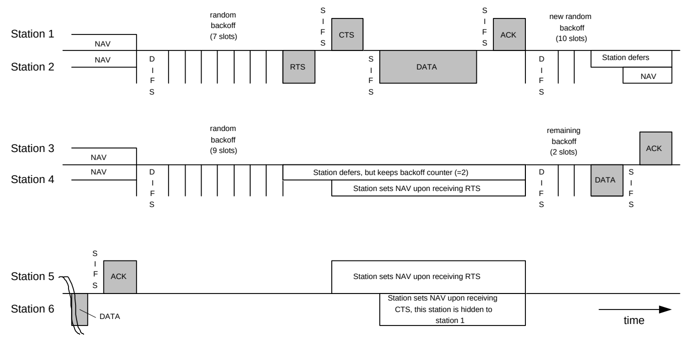

Fig. 2. DCF Access Procedure

CF-ACK frame is an acknowledgement by the *PC* and *CF-End* frame marks the end of the contention free period by *PC*.

The *PC* sends data to stations in *CF-Poll* frames which also achieve the polling function. A "poll bit", if enabled, in the *CF-Poll* frame polls the destination of the *CF-Poll* frame. A station transmits frames after the reception of a CF-Poll frame with the poll bit enabled. The need for separate acknowledgements is avoided by "piggybacking" acknowledgements on subsequent frames (by the setting of the appropriate bit). A station that has no data frame to send following a *CF-Poll* frame (addressed to it) does not transmit a frame and, therefore, the previous *CF-Poll* frame is not acknowledged.

STA frames are transmitted after the channel has been idle for a time interval *SIFS* as compared to *CF-Poll* frames that are transmitted once the channel has been idle for *PIFS*. Thus the *PC* will transmit the next *CF-Poll* frame in case there is no STA frame in time interval *PIFS* after the transmission of the previous *CF-Poll* frame. To minimize collisions during the contention free periods, each station sets its *NAV* equal to the maximum allowable length of the *CFP*. However, a station resets its *NAV* if a *CF-ACK* frame is seen by it before its *NAV* has expired.

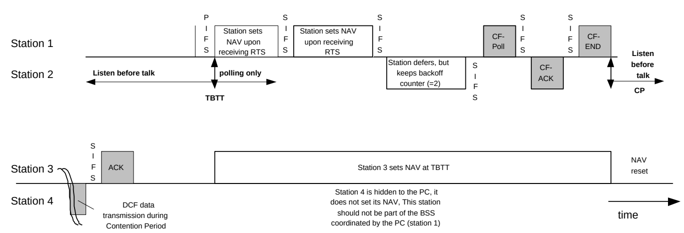

Fig. 3. PCF Access Procedure

IV. System Architecture

We classified the multihop network into three regions as *PCF Active Region*, *Overlap Region*, *DCF Active Region* as seen in Figure 4. Entities in the network are *Access Point, CF-Pollable STA, Non-CF-Pollable STA*.

CF-Pollable STA is a station that is registered to be in the *polling list* of *Access Point*. We assume that all *STAs* that hear AP's beacon are willingly to be in the polling list. *STAs* that are unable to hear beacons but want *AP* service are considered as *Non-CF-Pollable STA*.

PCF Active Region occupies the *AP* and *CF-Pollable STAs*. *Overlap Region* occupies *CF-Pollable STAs* and *Non-CF-*

Pollable STAs that are one hop away from the *PCF Active Region*. *DCF Active Region* is a region where *DCF* rules apply and occupies only *Non-CF-Pollable STAs*.

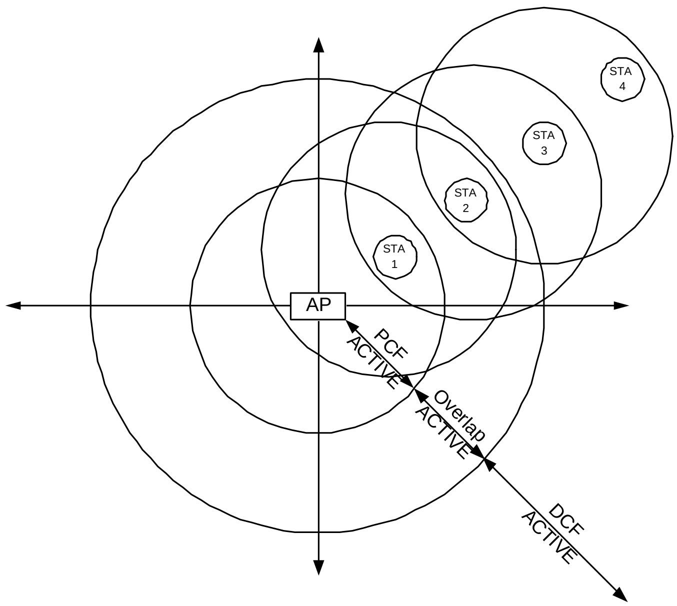

Fig. 4. Network Architecture

Overlap Region introduces rules that prevent interference to the *PCF Active Region* since *Non-CF-Pollable STAs* are hidden terminal for *AP*. *Non-CF-Pollable STAs* in *Overlap Region* are also able to to detect a *Infrastructure BSS* by monitoring the *PCF* related packets in the medium. These kind of packets[2][1] contains *Basic Service Set ID (BSSID)* which is MAC address of AP of the *infrastructure BSS*. *Non-CF-Pollable STA* authenticates with the *Pollable STA* by *open system* or *shared key* in which they use shared key that is assigned by the administrator to the STAs.

Non-CF-Pollable STAs in *Overlap Region* are also capable of detecting *CFP* and *CP* by looking into *RTS/CTS* packets. When a *Non-CF-Pollable STA* in *Overlap Region* detects a *BSSID*, he waits until the CP period and send a *Probe Request* packet to the *CF-Pollable-STA*. *CF-Pollable-STA* is the *STA* who is supposed to send *Probe Response* packet which includes *Beacon Interval*, *Capability Information* and *SSID*[1]. *Beacon Interval* is helpful for *Non-CF-Pollable STAs* in *Overlap Region* to predict the CP. *Capability Information Field* indicates that there is no point coordinator at *AP* and if it is the right *SSID*, the *STA* sends *Association Request* to the *AP* via the *CF-Pollable STA* indicating that the *STA* is *CF-Pollable*, not requesting to be placed on the *Polling List* by setting the *Capability Information Field* of *Association Request* frame format[1]. This registration type allows to activate in the future by reassociation to be placed in the polling list when it moves to the *PCF Active* region. *CF-Pollable STA* forwards the packet to the *AP* in the *CP* or in *CFP* when it is polled. *AP* registers the *STA* as a *Non-pollable STA* and remembers the *CF-Pollable STA* as a gateway for the *STA*. *CF-Pollable STA* also keeps the MAC address of the *Non-pollable STA* in

[2]Data, Data+CF-ACK, Null, CF+ACK, PS-Poll, Association Request, Reassociation Request, Probe Request, Authentication, Deauthentication

TABLE II

ROUTING TABLE FOR STA2

Gateway Node	MAC Addr. of STA1
Destination	**Next Hop**
MAC Addr. of STA3	MAC Addr. of STA3
MAC Addr. of STA4	MAC Addr. of STA3

order to remember his *STAs* under his serve. *AP* sends *Association Response* packet which contains *Association ID* of the *STA* in either broadcast or unicast to the *Gateway STA* of *Non-pollable STA*. If a *Gateway STA* get a packet of one of its children, by looking at the *To DS* or *From DS*[3] fields of the packet, it forwards the packet to the AP or to the corresponding child respectively.

Medium Access Rule for the multihop communication relies on *DCF*. *Non-CF-Pollable STAs* are allowed to send in the *CP* of *CFP*. *Non-CF-Pollable STAs* are also responsible for advertising the *BSS* to the *DCF Active* region by sending periodic Beacon frames. This beacon frames indicate the multihop service in the *Protocol Version Field* of *Frame Control Field* of Beacon Frame. The distinction between the AP beacon and beacon of other STAs in DCF Active Region, *Multihop*1, *Multihop*2, *Multihop*N is introduced as a *Protocol Version*. *Non-CF-Pollable STA* in the Overlap active region starts with *Multihop*1 frame and *STAs* in the *DCF Active* Region increments the *Multihop* level. A *STA* chooses always the lowest multihop level.

A STA keeps a routing table same as in Table II that contains its children and next hop to reach them and its gateway . A STA is responsible to take action to the packets that are destined to or originated from its children. If a STA gets a packet where the *To DS* is set, STA forwards the packet to its *gateway*. On the other hand, If a STA gets a packet where *From DS* is set, it forwards the packet to next hop of the STA according to its routing table. The characteristic of the IEEE 802.11 protocol is that stations react to all type of packets even if it is not destined for itself. Therefore decoding and inspecting each packet does not cause much overhead.

V. MULTIHOP NETWORK TOMOGRAPHY

Multihop Network Management can be done by inspecting the internal loss behavior of each station. The network organization is designed to construct as a tree. AP, the root of the tree, is capable of determining the lossy links by using *Expectation Maximization (EM) Algorithm* [10].

Determining link level loss characteristics of a network is essential in detecting congestion areas, routing faults, anomalous traffic for dynamic routing and video coding applications. Performing these measurements at each station may not be feasible because of the heterogeneous nature and the lack of centralized control. Even if this is possible, computing and communication overhead may not be tolerable.

[3]IEEE 802.11 packet structure specifies the direction of the packet by the To DS and From DS fields. Distribution Service (DS) is supported by the AP in order to handle the wireless network in an Extended Service Set[1].

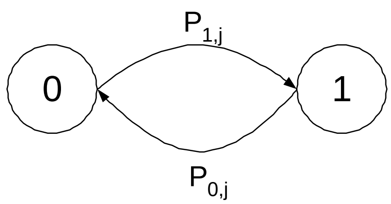

Fig. 5. Markov Chain associated with node j

Root of the tree sends multicast packet (beacon) to its children and each child also relays these packets to their childeren. Leaves of the tree who get the packet send back acknowledgement to the root. Based on the assumption of two-state Markov temporal structure of the loss behavior the transition probabilities in each link can be estimated by using Expectation Maximization algorithm [10].

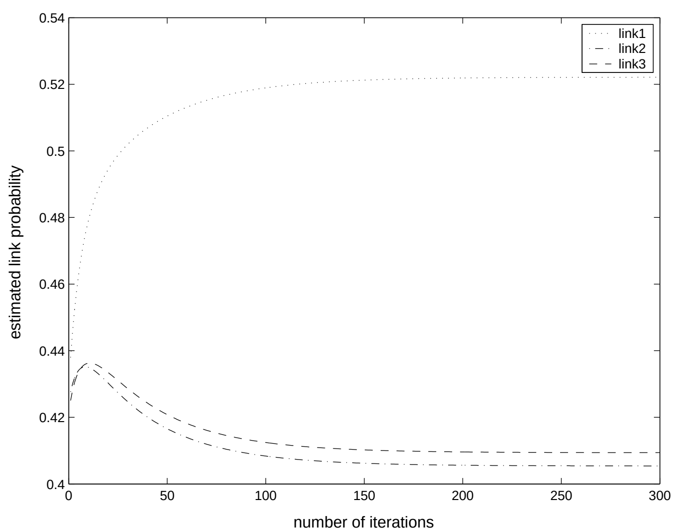

Fig. 6. Convergence of EM with 2-state Markov links of $p_{0j} = p_{1j} = 0.8$

We simulated the network with Markov temporal characteristics of equal transition probability of 0.8 with the EM algorithm for independent link losses. We observe from Figure 6 that the convergence rate decreases and the convergence value is almost equal to the stationary distribution.

By this method, low performance links can be determined and changed with new links in order to increase the performance of the network and have the network under control.

VI. Conclusion and Future Work

We have presented a brief outline of IEEE 802.11 MAC protocol and cited the motivation for a multihop extension. We designed the multihop extension to the IEEE 802.11 protocol compatible with the standard. We introduced a multihop management scheme which finds lossy links and reconfigures the network.

References

[1] *IEEE, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications"*. IEEE Standard 802.11, June, 1999.

[2] Mustafa Ergen, Anuj Puri *MEWLANA-Mobile IP Enriched Wireless Local Area Network Architecture*. IEEE VTC, Vancouver, September, 2002.

[3] John Boyer, David D. Falconer, Halim Yanikomeroglu *A Theoritical Characterization of the Multihop Wireless Communications with Diversity*. IEEE GLOBECOM 2001.

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

[5] Constantine Coutras, Sanjay Gupta, Ness B. Shroff *Scheduling of real-time traffic in IEEE 802.11 wireless LAN*. J.C. Baltzer AG, Science Publishers, Wireless Networks 6 (2000) 457-466.

[6] Mustafa Ergen *IEEE 802.11 Standard Summary*. www.eecs.berkeley.edu/ ergen.

[7] *High Performance Radio Local Area Network(HIPERLAN)*. Type 1; Functional specification, ETS 300 652, Radio Equipment and systems (RES) October 1996.

[8] *Wireless Communications: Principles & Practice*. Prentice Hall, Inc., New Jersey, 1996.

[9] A. Abu-Dayya, N. Beaulieu *Outage Probabilities in the Presence of Correlated Lognormal Interferers*. IEEE Trans. on Vehicular Technology, vol. 43, no. 1, 1994.

[10] S. Coleri, M. Ergen *Estimating Network Internal Link Loss Behavior From End-to-End Multi-cast Measurements*. Submitted to IEEE Trans. on Vehicular Technology