Two-way Admission Control and Resource Allocation for Quality of Service Support in Location-Aided Mobile Ad Hoc Networks

S. Sivavakeesar *S.Sivavakeesar@surrey.ac.uk*

G. Pavlou *G.Pavlou@surrey.ac.uk*

Center for Communication Systems Research, University of Surrey, Guildford
Surrey, GU2 7X, United Kingdom

This paper proposes a new QoS-aware medium access control (MAC) protocol in mobile ad hoc networks (MANETs). This takes the unique challenges of MANETs into consideration, and works in conjunction with the location-based forwarding strategy. This novel protocol is based on the legacy IEEE 802.11, and thus can be relatively easily integrated into existing systems. It is adaptive and network-aware depending on the type and intensity of traffic, and relative mobility patterns of nodes. In addition, it makes use of the point-coordination-function (PCF) of IEEE 802.11 in a distributed fashion for the first time in multihop MANETs. Our strategy enables two-way admission control for improved performance, whereby the forwarder-node selection algorithm allows previous hop nodes to perform implicit admission control using locally available information, while a selected forwarder-node performs explicit admission control depending on its current load. Analytical results confirm the performance improvement of our strategy.

I. Introduction

There is a growing need to support Quality of Service (QoS) in mobile ad hoc networks (MANETs), however, this is challenging. This task often requires acceptable channel conditions, QoS-aware mechanisms for channel access, identification of proper forwarding (transit) nodes, as well as measures for congestion prevention and management in those nodes. The absence of fixed infrastructure in MANETs means that there is no dedicated agency to manage and regulate the scarce channel resources. This problem gets exacerbated when the network grows in size, and additional problems such as increasing node-density and large number of nodes have to be faced and tackled [3].

A MAC based on the distributed coordination function (DCF) of IEEE 802.11 a/b is mostly prevalent in MANETs Although multiple non-overlapping channels exist in the 2.4 GHz and 5 GHz spectrum, most IEEE 802.11 based MANETs today use only a single channel [7]. In addition, despite significant advances in physical layer technologies, today's IEEE 802.11 still cannot offer the same level of sustained bandwidth as their wired brethren. More over, the advertised 54 Mbps bandwidth for IEEE 802.11 a/g is the peak link-layer data rate. When all the overheads – MAC contention, handshake packets such as request-to-send (RTS), clear-to-send (CTS) and ACK, and packet errors – are considered, the actual net bandwidth available to applications is almost halved. In addition, the DCF of IEEE 802.11 a/b/g has been primarily designed to support asynchronous best-effort traffic [7]. Since this MAC is based on random access method of carrier sense multiple access with collision avoidance (CSMA-CA), its ability to support QoS especially when the contention is high is very small. The contention from multiple users to access the common medium using a random access technique often results in unavoidable packet collisions, unbounded delay, and increased jitter. The time required to resolve collisions is a function of the network load. In addition, the DCF's "capture-effect", extensive use of control frames, the use of a binary exponential backoff scheme, and the time-varying nature of the bandwidth often results in the provision of insufficient bandwidth for time-sensitive applications [1][2]. In other words, due to interference from adjacent hops and problems due to hidden- and exposed-nodes, a bandwidth loss of up to 50% per hop is possible depending on the network topology [2]. Because of these reasons, a number of research works have questioned the suitability of DCF-based MAC for QoS support [4][5][8][9][10] in single radio multihop ad hoc networks. Fortunately, the IEEE 802.11 b/g standards and IEEE 802.11a standard provide 3 and 12 non-overlapped frequency channels respectively, which could be used simultaneously within a neighborhood [7].

In addition, in order to support the transmission of real-time traffic, a polling-based scheme called the point coordination function (PCF) was introduced in IEEE 802.11 [6][7]. In order to support both asynchronous and time-sensitive multimedia traffic, our MAC approach is based on a hierarchical strategy that utilizes the DCF- and PCF-based operations of the IEEE 802.11 for the first time in multihop MANETs after being modified to accommodate MAC-level service differentiation. Due to the fact that both the proposed MAC and location-based forwarding strategy work on the same principles – both use the local behavior to achieve a global objective – in this paper we combine our MAC scheme with a location-based forwarding strategy [3]. We then show how our MAC protocol enables two-way admission control facilitated through localized promiscuous listening, and localized mobility and load predictions.

II. Related Work and Our Motivation

It is difficult to compare different MAC protocols, as each has been developed with a different architecture and application in mind. A scheme termed soft reservation multiple access with priority assignment (SRMA-PA), is presented in [8]. It is a time division multiple access (TDMA) based MAC protocol that allocates stations to different time-slots. This scheme does not take asynchronous data traffic into consideration, as all data transmissions are required to reserve slots irrespective of whether they are real-time or best-effort traffic. Also there is a possibility for higher priority traffic to starve lower priority traffic by "confiscating" the slots already reserved by it. Since in this approach, the channel is time-slotted, and a slotted system requires network-wide time synchronization, which is relatively easy to achieve in infrastructure-based networks by using the base station as a time reference. This task becomes extremely difficult in distributed networks such as multihop MANETs [14]. A MAC approach that combines an allocation-based (TDMA) protocol and a contention-based (CSMA-CA) protocol is proposed in [9]. In this scheme, the number of slots in each frame is dependent on the number of nodes in the network, and hence each slot belongs to a single node only. The higher the number of nodes in the network, the larger the frame size. This leads to unbounded delay for time-sensitive applications. Similar approach is followed in reservation CSMA-CA [5]. In this scheme, CP and CFP alternate, and the CFP is based on TDMA. Since there is no node to regulate the common medium, this scheme may lead to a "stretching" problem [6]. It also requires proper time-synchronization, and each node is supposed to maintain a "slot-table" that indicates whether each slot is "reserved" or "available". Another MAC protocol that considers multiple channels is proposed in [10]. It combines code division multiple access (CDMA) or frequency division multiple access (FDMA), and TDMA to create a contention-free MAC, termed the sequenced neighbor double reservation (SNDR). Since it mainly considers time-slot allocation to make it contention-free, it fails to support asynchronous data traffic and requires complex slot-synchronization.

Although the DCF-based operation of the IEEE 802.11 is meant for best-effort traffic, there have been efforts that investigated differentiated services at MAC-level [11]. Service-differentiation is achieved by setting different values for contention-window (CW) - values of minimum (CWmin) and maximum (CWmax) - for different traffic classes. There is, however, no explicit guarantee of the level of service differentiation. There have also been some proposals to make DCF to be "per-stream-fair", as the DCF of legacy IEEE 802.11 tends to be unfair due to the "capture-effect" [12]. Fairness is achieved by dynamically modifying the CW of each traffic type by the source. The fairness approach does not, however, guarantee QoS support. In summary, each work presented above has its own drawback(s), and does not have the capability to provide MAC-level QoS for multimedia traffic in multihop MANETs.

III. Proposed MAC Mechanism

The proposed protocols are targeted for an unslotted multihop system, which is typical for mobile ad hoc wireless networks. In our scheme, both the DCF and PCF of IEEE 802.11 are used for the first time in multihop MANETs after being modified to accommodate MAC-level service differentiation. Although the PCF does require a centralized node, we describe next how this can be achieved in multihop MANETs using a novel strategy. The motivation for this work comes from the observation that the PCF-based operation offers a "packet-switched connection-oriented" service which is well suited for voice as well as multimedia traffic. The "connection-oriented" aspect of the PCF would allow the network to provide throughput, delay, and possibly jitter guarantees [6]. These guarantees, however, are *soft* as opposed to *hard* due to the unique nature of MANETs. The *soft* QoS means that there may exist transient time periods when the required absolute QoS guarantee is not possible due to non-availability of suitable forwarding nodes or network partition. However, this is minimized to a greater extend with our more distributed forwarder-node selection algorithm in moderately-densed MANETs as will be discussed in section III.D.

III.A. Need for the Creation of Multiple Parallel Channels

Although relatively simple, as explained in sections I and II, the use of single radio based on 802.11 DCF is not suitable for multimedia traffic in multi-hop MANETs [1][2]. On the other hand, supporting real-time applications in any network necessitates the availability of predictable resources. This is simply possible – especially in wireless networks where bandwidth is not abundant – with a central agency, which can have a control over the scarce channel resources for efficient and fair sharing. On the other hand, the very basic requirement of an ad hoc network is that it should not rely on any central node. However, some form of an agency to manage the channel resources is still required for QoS support. In order to accommodate the above mentioned mutually conflicting requirements, multiple parallel channels (multiple radios) are used in our scheme in order to improve capacity and scalability. Accordingly, each node is assigned a unique receiver-based channel [4], and each node behaves as a central node (AP) with respect to its own unique channel (medium).

In this receiver-based channel-assignment scheme, any sender has to transmit data using the receiver's unique channel, and hence, under normal circumstances each node uses its own channel to receive data from other nodes [4]. In addition, there is a common channel, which all nodes can use to disseminate and acquire mostly neighbor and routing related control messages. Accordingly, under normal circumstances each node in our scheme has to monitor its own unique channel and common channel for the reception of data and control frames respectively. However, there may be an exceptional case, where any node may be required to transmit data on the common channel as will be explained later. These channels are assigned to nodes dynamically in a conflict-free manner using the common channel. Since the unlicensed spectrum using IEEE 802.11 is extremely limited, an intelligent channel assignment scheme can lead to a proper coordination of the spectrum utilization which in turn mitigates coexistence/interference problem and increases the spectral efficiency [2]. In order to accommodate the situation in which any node can receive multiple transmissions initiated by different sources, IEEE 802.11 (both DCF and PCF) is used on top of each unique channel as depicted in Fig. 1 of [1]. Although the use of multiple receiver-based channels in MANETs has already been proposed in [2][4] and references therein (some of them use circuit-switching concept), the way the DCF and PCF co-exist on each unique channel (packet-switching concept) makes our strategy unique and different from previous approaches. Due to statistical multiplexing, our strategy improves the utilization of wireless bandwidth without compromising QoS support. The common channel, however, can support only the DCF of IEEE 802.11. Each node is expected to regulate and schedule its own unique channel. As will be explained below, the duration of CP and CFP on each unique channel is variable and adaptive depending on the traffic type and intensity of traffic within each type. This technique is to conserve the bandwidth available on each unique channel. Hence, with the use of multiple parallel channels, the PCF-based operation of the legacy IEEE 802.11 is adopted on each channel in order to support QoS in multihop mobile ad hoc networks. In our strategy, transmission by any node A to another node B has to be on the receiver's (B's) unique channel (see Fig. 1). This requires that each node maintain a channel-assignment table mainly for its one-hop neighbors, so that it can find out the channel associated with its intended next-hop node (receiver) at the time of transmission.

III.B. Protocol Description

Our QoS-aware MAC protocol is based on a hierarchical approach consisting of two sub-layers on each unique channel (i.e., this is not the case for the common channel) [1]. Like in the IEEE 802.11 standard, the lower sub-layer of the MAC protocol is called the DCF. This is to provide the fundamental access method in order to support asynchronous data traffic. The upper sub-layer (called PCF) is implemented on top of DCF to support real-time traffic through the "association process" [6][7]. Once a node A becomes "associated" with its forwarding-node B, the node A would not need to contend the unique medium (channel) of B any more during the whole session as long as B is within the transmission range of A. Our MAC protocol has the following three components on each unique channel in order to support QoS for real-time traffic as depicted in Fig. 1 of [1]: i) Admission control, ii) QoS-mapping, and iii) Resource reservation. In this work, two different service classes - high-priority (HP) (e.g. voice) and best-effort (BE)- are

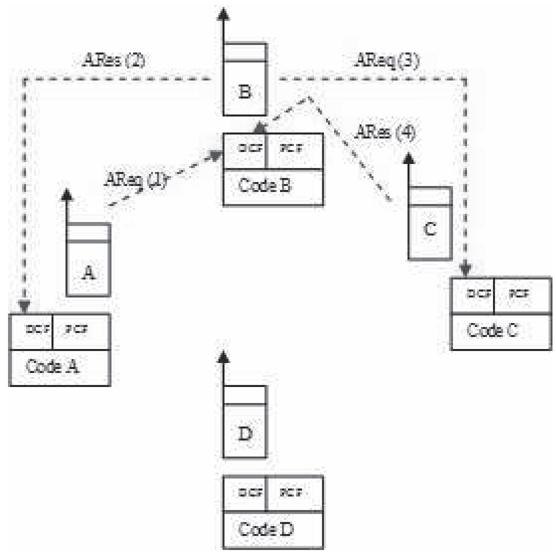

Figure 1: Working Mechanism of Our MAC

considered. The mode of operation of the proposed MAC on any unique channel switches between pure DCF-mode and combined (DCF+PCF)-mode depending on the traffic types the channel handles. On the other hand, the common channel always supports only the pure DCF-mode irrespective of the traffic types it handles. At bootstrap, our MAC on any unique channel sticks to pure DCF-mode by default as long as the traffic types the channel supports are all best-effort.

On the other hand, whenever a node A that is to initiate or relay high-priority (HP) data traffic has become associated with another node B, then the unique channel associated with B has to switch to the combined (DCF+PCF)-mode. We try to demonstrate the working mechanism of our MAC-protocol using a simple example. In Fig. 1, suppose that node A initiates a high-priority data transmission to destination C via an intermediate node B. Since source node A needs to transmit data of HP-type, it has to send an "Association Request" (AReq) frame to the forwarder (transit or next-hop) node B selected by our forwarder-node selection algorithm that will be described in section III.D [3]. This AReq frame is normally sent during the CP of a transit node's (B's) own unique channel, and hence uses the DCF-based access technique (see Fig. 1). As soon as node B receives the AReq frame, it has to send an "Association Response" (ARes) frame to the requesting node A during the CP of the latter's (A's) own channel. However, the AReq or the ARes frames can be transmitted as a piggyback to a DATA-frame during the CFP of a unique channel, if the sending node has already been associated with the other node. The transmission of ARes is, however, subject to our MAC-level admission control and QoS-mapping process as explained later in section III.C. At the same time, node B has to create a polling-list and include node A in it. At the start of CFP on B's channel, node B has to begin polling node A. In this way, any node (B) should be able to emulate the functionality of PC, and in our approach such a node is referred to as a virtual point coordinator (VPC). Since node B is an intermediate node, it has to forward the packet to its destination or the next forwarding node. Accordingly, it would soon send the AReq frame to node C, which is here assumed to be the destination, on C's own channel. After sending the ARes frame on node B's channel, node C will act as a VPC for node B, and has to be ready to poll node B at the start of CFP on C's unique channel. If node C were to send packets back to node A, then it would follow the same process as node A, but in the opposite direction. In this way, nodes along a particular path (or route) become polling-list members and VPCs of each other. This demonstrates as to how PCF-based operation is supported in a distributed manner in a multihop MANET. It is thus important that whenever a node (A) initiates or relays data traffic to another node (B), it has to be on the latter's (B's) unique channel. If the traffic type is of high priority, the source (A) can transmit when the node (B) polls, provided A is associated with the unique channel of B. This is the case even when a node transmits an ACK for the high-priority packets it receives correctly on its own channel. The forwarder-node selection and hence the channel selection is determined by the forwarder-node (next-hop) selection algorithm that will be described in section III.D. In case the algorithm of node A is unable to find a forwarder-node, then node A has to rely on the common channel for data transmission irrespective of its type.

In this way, only the high-priority traffic is allowed to use the CFP for data and ACK transmissions, and CP for the "associatione" process (i.e., AReq and ARes transmissions). On the other hand, best-effort traffic can only use the CP for data transmissions. Hence, provisioning network resources in our scheme uses two techniques, i) resource reservation during CFP (using PCF-functionality), and ii) prioritization during the CP. Since AReq, ARes and best-effort data traffic share the contention-period (CP) of any unique channel in most cases, the objective of the priority-based technique is to provide service-differentiation by allowing faster access to the medium to high-priority traffic classes [11]. As in the IEEE 802.11

DCF, priority access to the wireless medium is controlled through the use of an IFS. A new IFS termed Reservation IFS (RIFS) is defined, and its value is selected such that SIFS < PIFS < RIFS < DIFS [7]. To initiate new data transmission, RIFS or DIFS is used to contend for access to the medium during CP, depending on the traffic type. High-priority traffic uses RIFS before sending the AReq, while DIFS is used to gain access right for best-effort asynchronous traffic as in the IEEE 802.11 DCF. Since the RIFS is smaller than the DIFS, the high-priority traffic class has priority over the best-effort traffic that uses DIFS. An ARes frame is sent by any VPC node B (refer to Fig. 1), after SIFS during the CP of the unique channel of the requesting node A. When a collision happens or the unique medium associated to any node is sensed busy in the "association" process, the back-off time is calculated using the following modified equation [11].

$$Back\_off\_time = \left[2^{c+i} * rand()\right] * slot\_time \quad (1)$$

The constant c takes two different values depending on the traffic class, the parameter i is the transmission attempt number, and rand() is a random function with a uniform distribution in [0,1]. For the high-priority class, constant c takes the value 3, and the parameter i will be in the range of (1, MAX_ASSO_REQ_RETRY_LIMIT). For best-effort, the constant c takes the value 6 and the parameter i ranges from 1 to 4. This ensures that the high-priority class still enjoys priority over best-effort traffic even during the collision-resolution period [7][11]. The system constant MAX_ASSO_REQ_RETRY_LIMIT depends strongly on the characteristics of the real-time applications the ad hoc network supports and the extent of node mobility.

Whenever a node B becomes a VPC, the duration of CP and CFP on its unique channel is variable and adaptive depending on the intensity of traffic in each class. This is to conserve bandwidth available on the unique channel of any VPC. The presence of variable length CP and CFP should not, however, adversely affect jitter experienced especially by real-time applications. For this purpose, each node maintains a real-time active counter (RAC), which is associated with its own unique channel and each of the unique channels of its one-hop neighbors. This RAC is used to record the total number of active real-time sessions that are associated to a unique channel of a particular node at a point in time. When a node becomes a VPC for the first time, its channel switches from pure DCF-mode to the combined (DCF+PCF)-mode as soon as possible, i.e., such a new VPC initiates the CFP in its own unique channel by broadcasting a beacon-frame. On the other hand, if a node has been a VPC for more than a superframe period, then it should initiate the CFP every CFP Repetition interval (CFPRate) in order to minimize delay jitter [6]. In this process, a VPC uses RAC associated with its own channel in order to calculate the time required to support real-time traffic, and to allow sufficient time for the complete polling operation. Suppose each polling operation by a VPC takes T_{poll} seconds for a complete data transfer and a VPC supports n number of voice-sessions, then the total time required to support n number of sessions is given by equation (2):

$$T_{RT} = n{*}T_{poll}{+}PIFS{+}T_{Beacon}{+}SIFS{+}T_{CF\_End} \quad (2)$$

In equation (2), T_{Beacon} and T_{CF_End} are the time needed for beacon-frame transmission and time needed for CF_End frame transmission respectively [7]. Hence, exactly T_{RT} seconds before the end of the current superframe, a VPC that requires supporting n real-time sessions will initiate a beacon-frame in its own medium. The RAC associated with a unique channel of a specific VPC is broadcast as part of its beacon-frame, so that its one-hop neighbors that are interested in that particular node's unique medium can become aware of the RAC, and hence CFP and CP. Also, with RAC any node becomes aware of the current load of a unique channel associated with a given one-hop neighbor node, and it is important in forwarding-node (next-hop) selection and implicit admission control processes as will be explained in section III.D. The VPC finishes the CFP with CF_End frame broadcasting. Also when a VPC node operates in its contention period, other one-hop neighbor nodes have to make sure that their transmissions (best-effort data or AReq or ARes) will not extend the CP period of that VPC (otherwise it will cause a "stretching" effect [6]). For this purpose each neighbor that is interested in a particular VPC has to monitor the current RAC, beacon-frame and CF_End frame transmissions of the latter in order to locally maintain the CFP and CP of that VPC's unique channel. Once each neighbor has become aware of CFP and CP of a VPC's own unique medium (for this purpose each neighbor has to use equation (2), the former has to determine whether its pending transmission can be achieved well before the end of current contention-period. In case the time needed for any neighboring node to make its transmission attempt is not enough in the current DCF period, then it has to defer its transmission and wait for the next CP.

Because of the way in which transmissions take place, our strategy can completely eliminate the exposed-terminal problem on unique channels. However, such a problem is minimal as far as the common channel is concerned. On the other hand, the hidden-node problem is completely avoided when a unique channel operates in PCF, however, its effect is minimal when the channel operates in CP (when compared to the single radio DCF-based operation of IEEE 802.11a/b) [14]. In MANETs, due to mobility a node that has become associated with another may move out of its transmission range at any time. In order to accommodate this situation, each VPC dynamically maintains its polling-list as follows. If a polling-node (VPC) finds that it has not received (i.e., not heard) any transmission from one of its polling-list members for time period greater than POLLING_LIST_TIME_OUT, then that node address will be deleted from the former's polling-list immediately. This is how a "disassociation" process is performed in our scheme [6]. This strategy leads to efficient bandwidth management, and this occurs whenever nodes move out of each other's range or have finished their data transmissions. In our scheme, however, the forwarder-node selection algorithm tries to ensure that the forwarding-node (next-hop) to be selected will remain connected with the requester for a longer period as will be explained in section III.D. As mentioned before, each node is expected to monitor its own unique channel and common channel constantly for data and routing related information. As long as any node does not intend to initiate data transmission, there is no need for it to monitor the unique channels belonging to its one-hop neighbors. On the other hand, any node should start monitoring channels (media) belonging to its one-hop neighbors, whenever the former intends to transmit data, and the forwarder-node selection algorithm is used in this process. The number of channels any node should monitor for data transmission depends on the type and intensity of the traffic to handle, and its relative velocity with respect to its one-hop neighbors. There should be, however, a proper trade off between the number of channels a node can monitor simultaneously and the complexity of its receiver circuitry. For evaluation purposes, we assume that each node can monitor as many channels (subject to the maximum number of neighbors a node has at any moment) as it likes simultaneously.

III.C. MAC-Level Admission Control and QoS-Mapping

Traffic regulation is essential in networks that need to satisfy absolute QoS guarantees. This is achieved in our scheme with a two-way admission control mechanism. The explicit admission control is explained in this section, while the implicit admission control will be explained as part of the forwarder-node (next-hop) selection process in section III.D. Network congestion is difficult to resolve when real-time traffic, sensitive to both latency and packet loss, is present, without jeopardizing the QoS expected by the users of that traffic. Call Admission Control (CAC) is a strategy used to limit the number of callers into the network in order to reduce network congestion, therefore enabling the system to provide the desired QoS to incoming as well as existing calls. Since the PCF is utilized in our scheme to support real-time traffic, it has to limit the number of voice calls it can support when any unique channel operates in PCF. The maximum number (N_{max}) of high-priority traffic sessions that can be supported in the longest CFP, given a constant superframe size T_{SF}, is given by equation (3) [6][7]. Although in our mechanism, the durations of CP and CFP are variable, the CFP can grow up to a maximum value in order to safeguard the best-effort traffic. On the other hand, if CFP were allowed to increase arbitrarily up to the length of a superframe, the high-priority traffic would starve the best-effort traffic, and lead to "unfairness". For this purpose, each superframe in any unique channel of a VPC should contain the minimum CP given by T_{cp_min}. The high-priority traffic in our approach is assumed to be a time-sensitive periodic interactive voice service, which is generated using a constant bit rate (CBR) source for convenience.

$$N_{max} = \frac{T_{SF} - T_{cp\_min} - T_{ovhd}}{T_v} \tag{3}$$

T_{cp_min}, T_{ovhd}, and T_v are the minimum duration of CP, overhead involved for beacon and CF_END transmissions, and time to send a voice packet generated over a T_{SF} respectively [7]. For a particular HP application type, the above parameters are constant, and hence N_{max} tends to be constant in every node. In other words, the VPC can poll to a maximum of N_{max} number of times (or nodes) within a CFP on its own channel. Depending on the intensity of the high-priority traffic load, any node can request a VPC to poll it for more than once within each superframe period (T_{SF}) of VPC. The MAC-level QoS-mapping module of a particular node calculates the number of times it has to be polled by any VPC. This calculation is based on the bandwidth requested by the network-level QoS mechanism, and the latter should pass this information to the MAC for this purpose. Any node can inform any VPC as to how many times it has to be polled by it during each T_{SF} of the latter through the AReq frame – the AReq frame format is modified in order to accommodate this in our scheme. Whenever a VPC receives an AReq frame from any of its one-hop neighbors, its admission control module will check whether its CFP period is fully utilized (i.e., whether the RAC of its own channel has already reached N_{max}). If not, the VPC is required to send the ARes frame, and allocates the required bandwidth; here allocation means adding to the polling-list and this implies how many times the requesting node has to be polled within each T_{SF} of a VPC. This information is conveyed to the requesting node by the VPC through the ARes frame - the ARes frame format is modified for this purpose. If the maximum number has already been reached, then the VPC should not respond to any AReq. In this case, the requesting node should look for another appropriate forwarding node, after having tried for MAX_ASSO_REQ_RETRY_LIMIT. If, on the other hand, only part of the requested bandwidth can be supported by any VPC, it has to inform this to the requesting node through the ARes frame. The requesting node, in this case, has to look for another appropriate forwarding node for the unsupported bandwidth.

III.D. Forwarding-node (Next-hop) Selection Algorithm and MAC Functionality Adaptation based on Mobility

This section explains as to how our scheme enables previous hop nodes to perform implicit admission control and how the MAC functionality adapts depending on relative mobility predictions. This hop-by-hop strategy in mobile ad hoc networks is simple, scalable and effective with location-based forwarding mechanism – hence the adoption of location-based forwarding with our MAC [3]. However, it is the forwarder-node selection criteria, as formulated by equation (4) that play a major role in the efficient operation of our MAC. The MAC and the routing algorithm interact to find a neighbor as a forwarder-node (next-hop) as follows.

$$\Omega_{MI} = \frac{[LET_{MI} - LET_{TH}] * [N_{max} - RAC_I]}{d_{MI} + d_{IB}} \tag{4}$$

Let $N(M)$ be the neighbor set of node M, and M currently have a packet to be forwarded, d_{MI} be the distance from node M to any of its one-hop neighbors I (I $\epsilon\, N(M)$), d_{IB} be the distance from any node I (I $\epsilon\, N(M)$) to the packet's destination B, LET_{MI} be link expiration time of M with respect to I (I $\epsilon\, N(M)$), and RAC_I be the current load (number of voice calls) of a unique channel belonging to node I (I $\epsilon\, N(M)$). LET_{TH} is a system parameter. The criterion used in the forwarder selection algorithm is given by equation (4). The geographical distances can be determined from the location information of every node. The selection algorithm considers the currently available bandwidth during the CFP of the unique channel belonging to a one-hop neighbor, link expiration time (LET) and relative locations of the node-pair under consideration. Any one-hop neighbor I of node M that has the highest value for Ω_{IM} of equation (4) can be chosen as a forwarding-node. In other words, any one-hop neighbor that has enough bandwidth, higher probability to stay connected with the requesting node, and that lies very near to the destination would be chosen by the forwarder selection algorithm. This strategy enables a previous hop node M to perform implicit admission control for the forwarder I (I $\epsilon\, N(M)$), as node M prefers a node I having lower value for RAC on I's unique channel. In addition, this strategy enables a proactive way of traffic regulation, as M always tries to avoid an overloaded neighbor I. Since a location-based forwarding mechanism is adopted, there needs to be a recovery mechanism to resolve the "local-maximum" problem. As governed by the denominator of equation (4), in case of a "local-maximum" problem, forwarding would follow the least-backward-progression technique (LBF), and hence will tackle this problem [3]. If, previous hop node M's speed with respect to any I ($\epsilon\, N(M)$) is so high that Ω_{MI} is less than Ω_{TH} (Ω_{TH} is a system parameter and it takes a value of zero in our work), then M should not rely on the unique channel belonging to any I. Instead, M has to rely on the common channel for data transmission irrespective of the traffic type. This is the only exception where any node uses the common channel for data transmission, as the common channel is normally intended for the dissemination of routing related control information.

In order to determine the LET, each node is equipped with a mobility prediction module. This collects the motion parameters (e.g. velocity, radio propagation range) of two one-hop neighbor nodes M and I ($\epsilon\, N(M)$), and uses them to determine the time duration these two nodes will remain connected [15]. The predicted time is the link expiration time (LET) between two nodes, and this is used in equation (4). Every node becomes aware of the motion parameters related to its one-hop neighbors through the "HELLO-" message dissemination on the common channel. Our location service [3], which mainly operates on the common channel (for location updates and location queries) and unique channels (for location replies), enables nodes to register their locations and acquire locations of their desired communication partners.

III.E. Channel Assignment

Channel-assignment, which is a typical example of graph-coloring problem, however, is not included in this paper due to space limitations. However, it can be similar to the one proposed in [2]. This new assignment approach does not require modifications to IEEE 802.11. Further, this work considers a scenario in which each node possesses multiple interfaces, but the number of available interfaces is less than the number of available channels - and hence leads to reduced complexity [2].

IV. Capacity Analysis of the Proposed MAC Assignment

This section presents an analytical model that derives the saturation throughput/capacity of our QoS-aware MAC protocol in multihop mobile ad hoc networks. The obtained result is then compared against that of the DCF-based operation of the legacy IEEE 802.11 [7]. This analysis demonstrates the throughput increase of our proposed MAC scheme. This increase is important in a network where almost every node at a moment has high-priority packets for transmission, and hence affects the QoS.

Although the proposed QoS-aware MAC protocol is targeted for an unslotted and multihop system - which is typical for mobile ad hoc wireless networks - we assume that nodes operate in time-slotted mode in order to simplify our analysis. Prior studies in the literature show that the performance of MAC protocols based on carrier sensing is much the same as the performance of their time-slotted counterparts [13]. In these studies, it is assumed that the length of a time-slot equals one propagation delay and the propagation delay is much smaller than the transmission time of data packets. In our analysis, the length of each time-slot is denoted by τ However, τ is not just the propagation delay, because it also includes the overhead due to transmit-to-receive turn-around time, carrier sensing delay and processing time. Hence, τ in effect represents the time required to know the event that occurred τ seconds ago. The transmission times of AReq, ARes, RTS, CTS, DATA and ACK frames are normalized with respect to τ, and are denoted by l_{AReq}, l_{ARes}, l_{RTS}, l_{CTS}, l_{DATA} and l_{ACK} respectively. In deriving the throughput of our protocol, it is assumed that nodes always have packets in their buffers to be sent, and the destinations are chosen randomly from their neighbors: i.e., the analysis is based on the heavy traffic assumption. The network model envisaged here is, thus, one where nodes are involved with multiple parallel real-time traffic transmissions. In addition, in our network model, the nodes are Poisson distributed over a plane with density λ such that the probability $p(\iota, A)$ of finding ι nodes in an area A is given by:

$$p(\iota, A) = \frac{(\lambda A)^{\iota}}{\iota!} e^{-\lambda A} \tag{5}$$

Assuming that each node has the same transmission and receiving range of R, the average number of nodes (represented by $\tilde{\mathcal{N}}$) within a circular region of radius R is equal to $\lambda \pi R^2$. i.e.,

$$\tilde{\mathcal{N}} = \lambda \pi R^2 \tag{6}$$

In our network model, since nodes are distributed over a plane according to Poisson distribution, the probability of having N nodes within the range R of a node X is $\frac{e^{-\tilde{\mathcal{N}}}\tilde{\mathcal{N}}^N}{N!}$, where $\tilde{\mathcal{N}} = \lambda \pi R^2$. Let p be the probability that a silent node is ready to transmit in each time-slot, where the parameter p is slot independent. Although a node may have packets ready for transmission, its actual transmission attempt in a particular slot depends on the collision avoidance and resolution schemes used as well as the channel's current state. In order to take this fact into consideration, the probability – denoted by p' – that a node actually transmits in a slot is defined in equation (7). Here it is assumed that p' is independent of any time-slot in order to make the analysis tractable. Hence, $p' = p * Pr$ *(At least one channel is sensed idle in a slot)*, and

$$p' = p * \Pi_I \tag{7}$$

Since we consider N number of nodes for our analysis, the total number of channels available in our scheme for any node to communicate with its one-hop neighbors are $(N-1)$ unique-channels plus one common-channel. Therefore, the total number of channels available to any node X to communicate with its one-hop neighbors are N. Hence,

$$\Pi_I = \left[1-(1-\Pi_{I_{unique}})^{N-1}\right] \sum_{y=1}^{N-1} \varphi + \Pi_{I_{common}} [1-\varphi]^{N-1} \tag{8}$$

$\Pi_{I_{common}}$ is the probability that the common-channel is idle in a time-slot, $\Pi_{I_{unique}}$ is the probability that a unique-channel is idle in a time-slot. Let $\varphi = Pr(\Omega(t) < \Omega_{TH})$ be the probability that the forwarder-node selection criterion as given by equation (4) of node X with respect to its one-hop neighbor Y is less than the threshold value, where X $\neq$ Y and X,Y ϵ N-1 and $Pr_{xy}(\Omega(t) \geq \Omega_{TH}) = (1-\varphi)$ be the probability that an instantaneous criterion value Ω_{XY} of node X with respect to its one-hop neighbor Y is greater than or equal to a threshold value Ω_{TH}. The limiting probability will be derived shortly after considering the common-channel and unique-channel cases. To simplify the channel model, two key assumptions are made here [13].

1. The channel is modeled as a circular region, in which nodes can communicate with each other, while they have weak interactions with nodes outside the region. By weak-interaction, it is meant that the decision of inner nodes to transmit, defer and back off is almost unaffected by that of outer nodes and vice-versa. Thus, the channel's status is only decided by the failed and successful transmissions within the region.

2. Failed handshakes initiated by nodes within the region to outside nodes are also considered here, as this has a direct effect on the channel's usability to other nodes within the region. For this purpose, a new circular region is defined with a radius R' which falls between $R/2$ and $2R$. When $R' = R/2$, all nodes are guaranteed to hear each other within the circular region, and all the direct neighbors and hidden nodes are included when $R' = 2R$. Thus R' can be expressed $R' = \alpha R$ where $0.5 \leq \alpha \leq 2$, and α needs to be estimated.

With these assumptions, the unique-channel or the common-channel of a node can be modeled by a four-state Markov chain when it uses the DCF-based operation. Since only a unique-channel uses the PCF-based operation, it can be modeled by a two-state Markov chain during the contention-free-period (CFP) as depicted in Fig. 2 and Fig. 3. "Idle" is the state when any channel around node X is sensed idle, and its duration is τ in the case of DCF-based operation of the channel (i.e., $T_{idle_{DCF}} = \tau$), and can be either 0 or τ depending on the intensity of the high-priority traffic that a neighbor node handles during the CFP of its own channel (i.e., $T_{idle_{PCF}} = \{0, \tau\}$)."Long" is the state referring to a successful two-way handshake (RTS-DATA or CTS-ACK) being taking place for the best-effort traffic during the contention-period. During the CFP, the "long" for the high-priority traffic is represented by a two-way handshake (POLL-DATA).

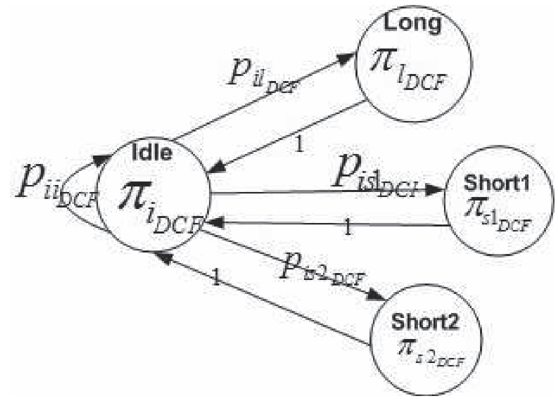

Figure 2: Markov chain model for a channel during the DCF period

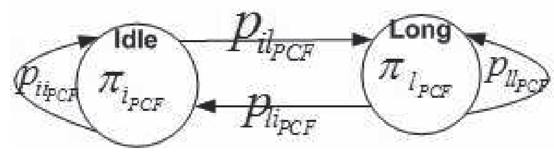

Figure 3: Markov chain model for a channel during the PCF period

In our approach, we consider a two-way handshake (RTS-DATA or CTS-ACK) during the CP for best-effort traffic for complete data transfer, as opposed to the conventional four-way handshake (RTS-CTS-DATA-ACK) [7]. The reason is that if any node X tries to transmit best-effort DATA to another node Y, the former has to send RTS and the subsequent DATA on Y's unique-channel, while the receiver Y has to send CTS and ACK on X's unique-channel. For simplicity, we assume that the channel is in effect busy for the duration of the whole handshake, thus the busy time T_{long} is:

$$T_{long} = \begin{cases} T_{long_{CP}} & \text{for BE in CP} \\ T_{long_{CFP}} & \text{for HP in CFP} \end{cases}$$

In the above equation $T_{long_{CP}} = (l_{RTS} + l_{CTS} + l_{DATA} + l_{ACK} + 4\tau)/2$ and $T_{long_{CFP}} = (l_{poll} + l_{DATA} + 2\tau)$. Assuming $l_{RTS} \approx l_{CTS}$, and $l_{DATA} \approx l_{ACK}$ the maximum value that T_{long} can take is given by the following equation.

$$T_{long} = \begin{cases} (l_{RTS} + l_{DATA} + 2\tau) & \text{for BE in CP} \\ (l_{poll} + l_{DATA} + 2\tau) & \text{for HP in CFP} \end{cases}$$

This can be approximated further to the following value by assuming $l_{RTS} \approx l_{POLL}$.

$$T_{long}(max) = [l_{RTS} + l_{DATA} + 2\tau] \quad (9)$$

"Short1" is the state when multiple nodes transmit RTS or AReq frames at the same time-slot on a particular unique-channel during the time the channel uses the DCF-based operation, and their transmissions collide (in the case of the common-channel this occurs due to multiple simultaneous RTS transmissions). The busy time of the channel $T_{short1_{DCF}}$ is therefore $T_{short1_{DCF}} = (l_{RTS} + \tau) \approx (l_{AReq} + \tau)$, assuming $l_{RTS} \approx l_{AReq}$. "Short2" is defined as a state when one node initiates a failed handshake with another node outside the region on a particular channel (unique or common) during the time the channel uses the DCF-based operation [13]. Even though a CTS packet may not be sent due to the collision of the sending node's RTS frame with other frames originated from nodes outside the region or due to the deferring of the receiving node to other nodes, those nodes overhearing the RTS as well as the sending node do not know whether the handshake is successfully continued, until the time required for receiving a CTS frame elapses. Therefore during the T_{short2}, the channel is in effect busy and unusable by other neighboring nodes. The $T_{short2_{DCF}}$ is here taken as:

$$T_{short2_{DCF}} = [l_{RTS} + l_{CTS} + 2\tau] \qquad (10)$$

The transition probabilities of the above Markov chains need to be calculated in order to derive the saturation throughput. During the DCF-based operation, no node is allowed to transmit continuously, and thus any channel needs to be idle between subsequent transmissions [7]. Therefore, the transition probabilities from state "long" to "idle" or from state "Short1"to "idle" or from state "Short2"to "idle" are all 1. However, obviously this is not the case during the PCF-based operation. During the CFP, each node takes control of its unique-channel and tries to resemble the functionalities of a point-coordinator. Since the neighboring nodes are accessing the unique-channel of a particular node in an organized way using the polling-based scheme, no simultaneous transmissions by nodes are allowed in a time-slot of a particular unique-channel. Therefore, during the PCF, the Markov chain model for any unique-channel consists of two states "idle"and "long" as shown in Fig. 2 (i.e., no "Short1" or "Short2"states during the PCF-based operation of a channel). In our scheme, only the neighboring nodes (one-hop neighbor) that are within a receiving range R of a particular node X can use the unique-channel of node X. Therefore, the transition probability $P_{ii_{DCF}}$ of any node X from "idle" to "idle" state in the case of the contention-period is the probability that none of the neighboring nodes of X transmits in a particular time-slot on the channel (it can be the unique-channel of X or the common-channel), and is given by:

$$\begin{aligned} P_{ii_{DCF}} &= \sum_{i=0}^{\infty}(1-p')^i \frac{N^i}{i!}e^{-N} \\ &= e^{-p'N} \end{aligned} \qquad (11)$$

Now the transition probability $P_{il_{DCF}}$ from "idle" to "long" during the DCF-based operation needs to be calculated. For such a transition to happen, one and only one node should be able to complete one successful two-way handshake while other nodes do not transmit. Assuming i number of nodes around node X, and let p_s be the probability that a node begins a successful two-way handshake at each time-slot, the $P_{il_{DCF}}$ during the contention-period can be given by the following equation.

$$\begin{aligned} P_{il_{DCF}} &= \sum_{i=1}^{\infty} ip_s(1-p')^{i-1}\frac{N^i}{i!}e^{-N} \\ &= p_sNe^{-p'N} \end{aligned} \qquad (12)$$

The transition probability $P_{iS1_{DCF}}$ from state "idle" to "short1" during the contention-period is the probability that more than one node transmit RTS or AReq frames in the same time-slot. It can be derived as follows.

$$\begin{aligned} P_{iS1_{DCF}} &= \sum_{i=2}^{\infty}[1-(1-p')^i - ip'(1-p')^{i-1}]\frac{N^i}{i!}e^{-N} \\ &= 1-(1+Np')e^{-p'N} \end{aligned} \qquad (13)$$

The transition probability from state "idle" to "short2", $P_{iS2_{DCF}}$, for the DCF case can be calculated as follows. Let $\pi_{i_{DCF}}$, $\pi_{l_{DCF}}$, $\pi_{S1_{DCF}}$ and $\pi_{S2_{DCF}}$ be the steady-state probabilities of states "idle", "long","short1" and "short2" respectively during the DCF-based operation. From Fig. 2, it can be deduced that,

$$\begin{aligned} \pi_{i_{DCF}}P_{ii_{DCF}} + \pi_{l_{DCF}} + \pi_{S1_{DCF}} + \pi_{S2_{DCF}} &= \pi_{i_{DCF}} \\ \pi_{i_{DCF}}P_{ii_{DCF}} + 1 - \pi_{i_{DCF}} &= \pi_{i_{DCF}} \end{aligned}$$

$$\pi_{i_{DCF}} = \frac{1}{2-P_{ii_{DCF}}} = \frac{1}{2-e^{-p'N}} \qquad (14)$$

$$P_{iS2_{DCF}} = \lfloor 1 - P_{ii_{DCF}} - P_{il_{DCF}} - P_{iS1_{DCF}} \rfloor \qquad (15)$$

Now the transitional probabilities associated in the case of PCF-based operation of any unique-channel need to be calculated. In Fig. 3, the unique channel of a particular node remains idle under two circumstances only, when it uses the PCF-based operation.

1. All the one-hop neighbors of the node under consideration have finished their high-priority traffic. (i.e.,: have nothing to transmit).
2. The forwarder-node selection criterion Ω_{IM} as given by equation (4) of any one-hop neighbor node I with respect to the node M that is under consideration is greater than the threshold Ω_{TH}

With these facts, the transition probability $P_{ii_{PCF}}$ from state "idle" to "idle" when the channel uses the PCF is given by,

$$P_{ii_{PCF}} = \left[\frac{1-p}{N-1}\right]^{N-1} + [1-\varphi]^{N-1} \qquad (16)$$

Also, the transition from state "idle" to "long" when a unique channel uses the PCF-based operation can happen only when,

1. Exactly one node uses that channel while other one-hop neighbors do not make transmission attempts.
2. The Ω-value of the node that owns the unique-channel under consideration with respect to the transmitting node is below the threshold value.

Hence, the transition probability $P_{il_{PCF}}$ from state "idle" to "long"during the contention-free-period of the unique-channel which is owned by node X is given by:

$$P_{il_{PCF}} = \frac{p}{N-1}[1-p]^{N-2}\sum_{y=1}^{N-1}\varphi \qquad (17)$$

It can be further shown that the transition probability $P_{li_{PCF}}$ from state "long" to "idle" when the channel uses PCF-based operation is equal to $P_{ii_{PCF}}$. Also, the probability for the channel to remain in the "long" state – i.e., transition from state "long" to "long" – can also be shown to be equal to $P_{il_{PCF}}$. Let $\pi_{i_{PCF}}$ and $\pi_{l_{PCF}}$ be the steady-state probabilities of states "idle" and "long" respectively during the PCF-based operation. From Fig. 3, it can be deduced that:

$$\begin{aligned} \pi_{i_{PCF}} P_{ii_{PCF}} + \pi_{l_{PCF}} P_{li_{PCF}} &= \pi_{i_{PCF}} \\ \pi_{i_{PCF}} P_{ii_{PCF}} + (1-\pi_{i_{PCF}}) P_{li_{PCF}} &= \pi_{i_{PCF}} \\ \pi_{i_{PCF}} &= P_{li_{PCF}} \\ \pi_{l_{PCF}} = 1 - \pi_{i_{PCF}} & \qquad (18) \end{aligned}$$

The limiting probability $\Pi_{I_{unique}}$ that the common-channel around any node is found idle has been derived in [13] as follows. Since the common-channel uses only the DCF-based operation, there is no contention-free-period in the common-channel.

$$\Pi_{I_{common}} = \frac{\psi_1}{\psi_2} \qquad (19)$$

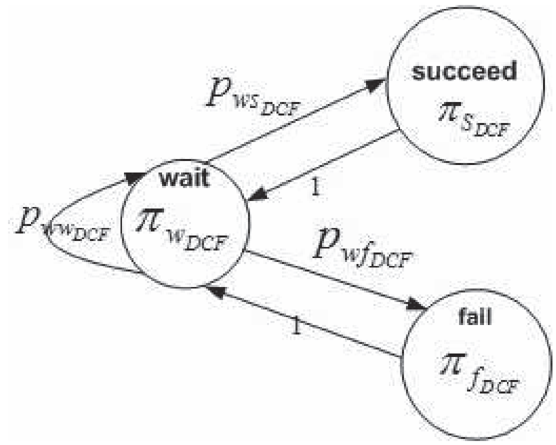

Figure 4: Markov chain model for a channel during the DCF period

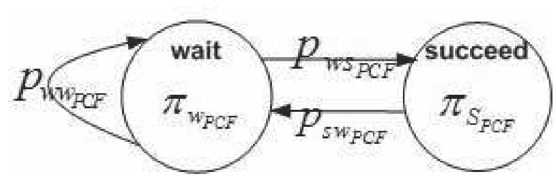

Figure 5: Markov chain model for a channel during the PCF period

where $\psi_1 = \pi_{i_{DCF}} T_{idle_{DCF}}$ and $\psi_2 = \pi_{i_{DCF}} T_{idle_{DCF}} + \pi_{l_{DCF}} T_{long_{DCF}} + \pi_{S1_{DCF}} T_{short1_{DCF}} + \pi_{S2_{DCF}} T_{short2_{DCF}}$. The limiting probability $\Pi_{I_{unique}}$ that the unique-channel of node X is found idle can be derived by the following equation, where we need to consider both the DCF and the PCF-based operations:

$$\Pi_{I_{unique}} = \frac{\psi_3}{\psi_4} + \frac{\psi_5}{\psi_6} \qquad (20)$$

where $\psi_3 = \frac{T_{DCF}}{T_{DCF}+T_{PCF}} \pi_{i_{DCF}} T_{idle_{DCF}}$, $\psi_4 = \psi_2$, $\psi_5 = \frac{T_{PCF}}{T_{DCF}+T_{PCF}} \pi_{i_{PCF}} T_{idle_{PCF}}$, and $\psi_6 = \pi_{i_{PCF}} T_{idle_{PCF}} + \pi_{l_{PCF}} T_{long_{PCF}}$. T_{DCF} of a unique channel refers to the contention-period while T_{PCF} refers to the contention-free-period. In equations (19) and (20), $T_{long_{DCF}}$ and $T_{long_{PCF}}$ can be approximated to $T_{long}(max)$ as given by equation (9). In order to calculate the probability p_s that a node X starts successfully a two-way handshake in a time-slot, the state of a node X is modeled by a three-state Markov chain in the case of DCF-based operation, and by a two-state Markov chain in the case of PCF-based operation, as shown in figures 4 and 5. "Wait" is the state when a node has no packets to transmit or defers for other nodes or backs off, "fail" is the state when a node initiates an unsuccessful handshake, and "succeed" is the state when a node completes a successful two-way handshake with other nodes. For analytical purposes, "succeed" and "fail" are regarded as the states in which two different kinds of virtual frames are transmitted and their durations are as follows:

$$T_{succeed} = \begin{cases} T_{succeed_{DCF}} \\ T_{succeed_{PCF}} \end{cases}$$

$$T_{succeed} = T_{long}(max) = l_{RTS} + l_{DATA} + 2\tau \quad (21)$$

$$T_{fail} = T_{fail_{DCF}} = T_{short2} = l_{RTS} + l_{CTS} + 2\tau \quad (22)$$

Since only one node is expected to transmit in a given time-slot, there is no phenomenon of "failure" during the PCF-based operation; hence, $T_{fail_{PCF}} = 0$. Also, it is obvious that the duration of a node in "Wait" state T_{wait} is either zero or τ depending on the intensity of high-priority traffic that it handles, and which is assumed to be equally probable when a node is subject to the PCF-based operation by a channel. During the contention-period, no node is allowed to transmit data packets continuously. Hence, the transition probabilities from state "succeed" to "Wait" and from state "fail" to "Wait" are both 1. However, this is not the case, when a channel uses the PCF-based operation, as any node can transmit continuously depending on the intensity of high-priority traffic that it handles. In [13] it has been shown that $P_{ws_{DCF}}$ can be derived from the following equation, for the case a node is subject to the DCF-based operation:

$$P_{ws_{DCF}} = [2p'(1-p')e^{-p'N}] \int_0^1 re^{-\xi} dr \quad (23)$$

In equation (23), $\xi = p'N[1-2q(r/2)/\pi][2l_{RTS}+1]$, r is the distance between two communicating nodes, and $q(t) = \arccos(t) - t(1-t^2)^{1/2}$. From the Markov chain shown in Fig. 4, the transition probability $P_{ww_{DCF}}$ that node X continues to stay in "Wait" state in a time-slot, when it is under the DCF-based operation, is equal to $(1-p')e^{-p'N}$. Here it is assumed that the node does not initiate any transmission, and there is no node around it initiating a transmission. Let $\pi_{S_{DCF}}$, $\pi_{w_{DCF}}$ and $\pi_{f_{DCF}}$ be the steady-state probabilities of states "succeed", "Wait" and "fail" respectively during the time the node is subject to the DCF-based operation by a channel. The parameter $\pi_{w_{DCF}}$ during the time any node is subject to the DCF-based operation can be determined from Fig. 4 as follows:

$$\begin{aligned} \pi_{w_{DCF}} P_{ww_{DCF}} + \pi_{S_{DCF}} + \pi_{f_{DCF}} &= \pi_{w_{DCF}} \\ \pi_{w_{DCF}} P_{ww_{DCF}} + 1 - \pi_{w_{DCF}} &= \pi_{w_{DCF}} \end{aligned}$$

$$\pi_{w_{DCF}} = \frac{1}{2 - P_{ww_{DCF}}} = \frac{1}{2-(1-p')e^{-p'N}} \quad (24)$$

In order to derive the saturation throughput of our MAC scheme, it is important in our analysis to determine the steady-state probability of state "succeed" π_S. During contention-period of a channel, the steady-state probability of state "succeed" $\pi_{S_{DCF}}$ can be calculated from Fig. 4 as:

$$\pi_{s_{DCF}} = \pi_{w_{DCF}} P_{ws_{DCF}} = \frac{P_{ws_{DCF}}}{2-(1-p')e^{-p'N}} \quad (25)$$

Now $\pi_{f_{DCF}}$ of Fig. 4 can be determined by deducting $\pi_{S_{DCF}}$ and $\pi_{w_{DCF}}$ from one. Any unique channel has a maximum of $(N-1)$ number of one-hop neighbor nodes that can use the channel. In other words, any node can use $(N-1)$ number of unique-channels. Therefore, the Markov chain model for a unique channel during the CFP can be approximated to the Markov chain model for a node during the PCF-based operation. Now let $\pi_{S_{PCF}}$ and $\pi_{w_{PCF}}$ be the steady-state probabilities of states "succeed" and "Wait" respectively during the PCF-based operation of a channel. Therefore, from equation (18) and Fig. 5,

$$\pi_{w_{PCF}} = \pi_{i_{PCF}} = \left[\frac{1-p}{N-1}\right]^{N-1} + [1-\varphi]^{N-1} \quad (26)$$

$$\pi_{S_{PCF}} = 1 - \pi_{w_{PCF}} \quad (27)$$

The overall steady-state probability "succeed" π_S of a node can be calculated from the following equation:

$$\pi_S = \sum_{n=1}^{N} \left[\Upsilon_{DCF_n} * \pi_{S_{DCF_n}} + \Upsilon_{PCF_n} * \pi_{S_{PCF_n}}\right] \quad (28)$$

N is the maximum number of channels that any node can use at a time with probability $\frac{e^{-\tilde{N}}\tilde{N}^N}{N!}$, where $\tilde{N}$ is given by equation (6), $\Upsilon_{DCF_n} = \frac{T_{DCF_n}}{T_{DCF_n}+T_{PCF_n}}$, $\Upsilon_{PCF_n} = \frac{T_{PCF_n}}{T_{DCF_n}+T_{PCF_n}}$, T_{DCF_n} is the time period during which the n^{th} channel uses the DCF-based operation,T_{PCF_n} is the contention-free-period of n^{th} channel, $\pi_{S_{DCF_n}}$ is the steady-state probability for state "succeed" of a node in the n^{th} channel when it uses the DCF-based operation, $\pi_{S_{PCF_n}}$ and is the steady-state probability for state "succeed" of a node in the n^{th} channel when it uses the PCF-based operation. The $\pi_{s_{DCF_n}}$ and $\pi_{S_{PCF_n}}$ for n^{th} channel can be calculated from equations (25) and (27) respectively. Similarly, we can determine the net value for π_w and π_f of a particular channel by considering both the DCF and PCF-based operations. It is important here to note that although any node can have a maximum number of N channels (i.e., $(N-1)$ unique-channels plus the common-channel) at a time for its use, it may not use all the N channels at a moment. Further in the case of common-channel $T_{PCF}=0$, as it operates only on the DCF-mode.

It can be noted here that the parameter π_s is in fact synonymous to the previous parameter P_s of equation (12), and both of them can be computed easily with numerical methods. Accordingly, the throughput,Th_{single}, on a single channel can be deduced using the following equation.

$$Th_{single}=\frac{\pi_S l_{DATA}}{\pi_w T_w+\pi_S T_S+\pi_f T_f} \qquad (29)$$

In equation (29), $T_w = T_{wait}$, $T_S = T_{succeed}$ and $T_f = T_{fail}$. Therefore, the total throughput, when considering a particular node and its $(N-1)$ neighbors, can be computed from the following equation:

$$Th_{Total}=\sum_{k=1}^{N-1} Th_{single_k}+Th_{common} \qquad (30)$$

Th_{common} denotes the throughput on the common-channel, and which can be approximated to that of equation (29) after considering the fact that the common-channel does not use PCF-based operation. With this approximation, the total throughput in the neighborhood of a particular node is given by:

$$Th_{Total}=\sum_{k=1}^{N} Th_{single_k} \qquad (31)$$

Our mathematical analysis demonstrates the increase in the saturation throughput/capacity of our proposed MAC scheme as given by equation (31), when compared to that of the single radio DCF-based operation of legacy IEEE 802.11 that is approximately given by equation (29) [13]. This increase is nearly N times as high as the throughput of the single radio DCF-based operation of the legacy IEEE 802.11 MAC, where the N is taken to be the number of neighbor any node X has within its range R. Although this increase is shown to be nearly N-times higher, it would be even substantially higher due to the following two reasons: i) in the case of state "long" in our approach, we consider two-way handshake as opposed to the four-way handshake of the typical single radio DCF-based operation of the legacy IEEE 802.11 MAC scheme. As a result, the T_S of equation (29) for the original IEEE 802.11 MAC would be higher than (nearly twice) that of our scheme, and ii) the longer each unique channel uses the PCF-based operation, the greater the utilization of channel bandwidth would be. As a result the net value (considering both PCF and DCF-based operations) for the T_w and T_f of equation (29) for our scheme would be much lower than those for the DCF-based operation of the original IEEE 802.11 scheme [13]. Also, due to less likelihood for contention, the steady-state probability for state "succeed" (π_S of equation (28)) in our scheme would take a higher value than that of the original IEEE 802.11. As a result, the Th_{single} as given by equation (29) for the legacy IEEE 802.11 would be much lower than that of our scheme for the single-channel case.

V. Conclusion and Future Work

In this paper we presented a QoS-aware MAC protocol for multimedia traffic in MANETs, and evaluated its performance through mathematical analysis. The proposed protocol makes use of the packet-switching concept based on the PCF for the first time in multi-hop MANET. This mechanism enables cross-layer optimizations by working in conjunction with location-based forwarding strategy (i.e., constant network- (especially routing) and MAC-level interactions). In addition, our strategy allows two-way admission control, which is partly enabled through our forwarder-node selection algorithm. The mathematical analysis confirm the performance improvements of our scheme. In addition, our proposed approach leads to fewer collisions and hence minimizes the need for re-transmissions. This fact will in turn conserve scarce resources such as battery power and bandwidth – although this feature is not assessed in the current version, we intend to deal with this in the future. As explained before, the MAC functionality of a node is adaptive and network-aware depending on the type of traffics and relative mobility patterns of mobile nodes. Since this work is mainly based on the IEEE 802.11 standard, it can be relatively easily integrated into existing systems. We also plan to extend this work in the future with the use of the recent IEEE 802.11e standard in order to support multiple traffic classes.

VI. Acknowledgement

This work was undertaken in the context of the UK Engineering and Physical Sciences Research Council (EPSRC) Programmable Ad-hoc Networks (PAN - GR/S02129/01) research project.

References

[1] S. Sivavakeesar, and G. Pavlou, *Quality of Service Aware MAC Based on IEEE 802.11 for Multihop Ad-Hoc Networks*, Proc. of IEEE WCNC, Atlanta, GA, Mar. 2004, pp. 1488 - 1487.

[2] P. Kyasanur, and N. Vaidya, *Routing and Interface Assignment in Multi-Channel Multi-Interface Wireless Networks*, Proc. of IEEE WCNC, vol. 4, Mar. 2005, pp. 2051 - 2056.

[3] S. Sivavakeesar, and G. Pavlou, *Scalable Location Services for Hierarchically Organized Mobile Ad hoc Networks*, Proc. of ACM MobiHoc, Urbana-Champaign, Illinois, USA, May 2005.

[4] M. Joa-Ng, and I-T. Lu, *Spread Spectrum Medium Access Protocol with Collision Avoidance in Mobile Ad-hoc Wireless Networks*, Proc. of IEEE INFOCOM, vol. 2, Mar. 1999, pp. 776 - 783.

[5] I. Joe, and S.G. Batsell, *Reservation CSMA/CA for Multimedia Traffic over Mobile Ad-hoc Networks*, Proc. of IEEE ICC, vol. 3, Jun. 2000, pp.1714 - 1718.

[6] M. Veeraraghavan, N. Cocker, and T. Moors, *Support of Voice Services in IEEE 802.11 Wireless LANs*, Proc. of IEEE INFOCOM, vol. 1, Apr. 2001, pp. 488 - 497.

[7] IEEE Draft International Standards, *Part 11:Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, ISO/IEC 8802-11, IEEE P802.11/ D10, Jan. 1999.

[8] C.W. Ahn, C.G. Kang, and Y.Z. Cho, *Soft reservation multiple access with priority assignment (SRMA/PA): a novel MAC protocol for QoS-guaranteed integrated services in mobile ad-hoc networks*, Proc. of IEEE VTC-Fall, vol. 2, Sep. 2000, pp. 942 - 947.

[9] I. Chlamtac, A. Farago, A.D. Myers, V.R. Syrotiuk, and G. Zaruba, *ADAPT: A Dynamically Self-Adjusting Media Access Control Protocol for Ad Hoc-Networks*, Proc. of IEEE GLOBECOM, 1999, vol. 1A, pp. 11 - 15.

[10] Z. Cai, and M. Lu, *SNDR: a new medium access control for multi-channel ad hoc networks*, Proc. of IEEE VTC-Spring, vol. 2, May. 2000, pp. 966 - 971.

[11] I. Aad, and C. Castelluccia, *Differentiation Mechanisms for IEEE 802.11*, Proc. of IEEE INFOCOM, vol. 1, Apr. 2001, pp. 209 - 218.

[12] B. Bensaou, Y. Wang, and C.C. Ko, *Fair Medium Access in 802.11 Based Wireless Ad Hoc Networks*, Proc. of ACM MobiHOC, Aug. 2000, pp. 99 - 106.

[13] Y. Wang, and J.J. Garcia-Luna-Aceves, *Performance of Collision Avoidance Protocols in Single-Channel Ad Hoc Networks*, Proc. of IEEE ICNP,Paris, France, Nov. 2002.

[14] A. Chandra, V. Gummalla, and J.O. Limb, *Wireless Medium Access Control Protocols*, IEEE Communications Surveys and Tutorials, Second Quarter 2000, vol. 3, no. 2.

[15] W. Su, S-J. Lee, and M. Gerla, *Mobility Prediction and Routing in Ad Hoc Wireless Networks*, Int'l Journal of Network Management, vol. 11, no. 1, John Wiley and Sons, Jan.- Feb. 2001, pp. 3 - 30.