ENHANCING THE SPECTRUM UTILIZATION IN THE CELLULAR BAND: A STUDY OF GSM – AD-HOC INTER-WORKING

S. Sankaranarayanan and A. Mishra
Wireless Internet Networking Laboratory
Bradley department of Electrical and Computer Engineering
Virginia Tech.
Blacksburg, USA
srivatsan@vt.edu, mishra@vt.edu

Abstract - There is a strong belief that the spectrum both in the public and private sectors in the United States is getting scarce. This has driven research into developing efficient protocols and mechanisms for improving wireless spectrum utilization. While using more efficient schemes can certainly help improve the spectrum utilization, most of these works ignore the fact that the actual amounts of spectrum utilized by deployed systems on an average, is quite low. In this paper, a network architecture and MAC protocol that enables adaptive spectrum sharing based on the concept of primary-secondary (PS) system is presented. In a PS system, there is a main user or the primary user to whom the spectrum is licensed. Nevertheless operation of secondary systems in conjunction with the primary is allowed and encouraged. The deployment of such systems can be regulated by a central body such as the FCC. Our study uses the well known Global System for Mobile Communications (GSM) for the primary system. The secondary consists of an ad-hoc overlay whose existence is transparent to GSM. The issues involved in achieving this inter-working between GSM and the ad-hoc network are exposed. A novel MAC protocol for the ad-hoc network which we call Ad-hoc Secondary MAC (AS-MAC) is proposed. We show that our AS -MAC improves the spectrum utilization by up to 70%.

Keywords: GSM, spectrum pooling, inter-working, multi-channel, dynamic channel selection, primary-secondary system.

1. INTRODUCTION

There is a strong belief that the spectrum both in the public as well as private sector in the United States is getting scarce. But the recent measurements of the spectrum utilizations taken in major metropolitan areas [1, 2] in the private as well as public bands (UHF & VHF) suggest that spectrum utilization in several frequency bands is very low for extended periods of time.

In order to motivate the reader we refer to some recent spectrum measurements. Figure 1 shows the general nature of spectrum occupancy in approximately 700 megahertz block of spectrum below 1 GHz in Atlanta, New Orleans, and San Diego. This data (reproduced from [1]) was taken by FCC's Enforcement Bureau in June 2002. It can be seen that there can be a large variation in spectrum use intensity within the spectrum below 1 GHz in Atlanta. The lower frequencies in the observed spectrum tend to have lower utilizations than the higher frequencies. The reader is referred to [2] for more measurement results indicating inefficient spectrum use.

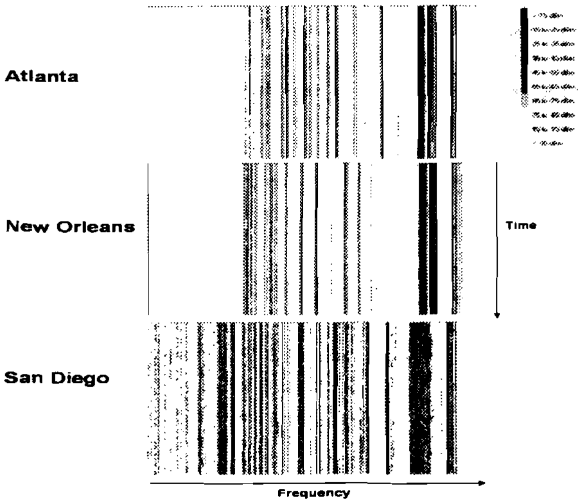

Figure 1. Occupancy of approx. 700 MHz of spectrum below 1 GHz (Ref. 5)

Thus a main reason for the currently perceived spectrum scarcity in the licensed band may be partially due to its low utilization rather than to inefficient spectrum access technologies. This motivates one to consider utilizing the "holes" that exist in the currently allocated spectrum blocks by others who need to access the spectrum at that time. Our goal in this paper is to propose an architecture for adaptive spectrum utilization based on the concept of primary-secondary (PS) system. The concept of a PS system is not entirely new. Related work is outlined in the next section.

In a PS system, there is a main user or the primary user to whom the spectrum is licensed. Nevertheless operation of secondary systems in conjunction with the primary is allowed and encouraged. An important characteristic of such a system is that the secondary in no way affects the performance of the primary. For example, the secondary can use only that portion of the spectrum which is not used by the primary at any given point in time. Thus the secondary system has to live with non-guarantee-able system resources such as bandwidth which may not enable the secondary to offer guaranteed services. Still the secondary can offer non-real-time and relatively time insensitive applications leading to appreciable use of the otherwise unused spectrum. The deployment of such systems may be regulated by a central body like the FCC. FCC is already working on policy regulations for such primary-secondary systems. We are currently investigating the issue of providing real-time services on the secondary network as well and will report our findings in the near future.

In this paper, a specific case of a PS system is considered. The well known Global System for Mobile Communications (GSM) is used as the primary system. The secondary consists of an ad-hoc overlay whose existence is transparent to GSM. The ad-hoc secondary network is referred to as ASN (Ad-hoc Secondary Network) and the nodes of the ad-hoc secondary are referred to as ANs (Ad-hoc Nodes). The issues involved in achieving this inter-working between GSM and ASN are exposed. A novel MAC protocol for ASN, which we call AS–MAC (Ad-hoc Secondary MAC), is proposed. AS-MAC is shown to improve the spectrum utilization by up to 70%.

This paper is organized as follows. Section 2 gives some background and sheds light on related work. Section 3, outlines the proposed operational mechanism for the secondary ad-hoc network and also explains in detail our new MAC called AS-MAC. Section 4, explains the simulation methodology and results. Section 5 concludes the paper and future work is outlined in Section 6.

2. BACKGROUND AND RELATED WORK

A brief outline of the GSM system is first given. Interested readers are referred to [4] for more details. A GSM system is composed of one or more cells. Every cell has a Base Station (BS) and several Mobile Stations (MSs) associated with the BS. GSM 900 uses the spectrum from 890-915 MHz for uplink transmissions (from MS to BS) and 935-960MHz for downlink transmissions (from BS to MS). Each of these two bands is divided into several smaller frequency bands of width equal to 200 KHz. One frequency from each of these two bands that are 45 MHz apart constitute one uplink-downlink frequency pair. We consider a GSM system employing FDMA/TDMA for its physical layer. Every frequency channel is divided into frames each being divided in turn into eight time-slots. A frequency-timeslot combination is the smallest quantum of communication in GSM.

There are several logical signaling channels defined in the GSM system for both uplink and downlink transmissions. In the uplink the only channel is the random access channel called RACH. This channel is used by the mobiles for sending channel request messages to the BS. In the downlink, there are several logical control channels. They are broadcast control channel (BCCH), frequency correction channel (FCH), synchronization channel (SCH), paging and access grant channel (PAGCH), and cell broadcast channel (CBCH). FCCH and SCH messages help the MS attain time synchronization with the BS. BCCH and CBCH are used for broadcasting general information, ex: cell Id, etc. PAGCH is used to page the mobiles in the case of an incoming call.

Little work has been done on spectrum pooling and primary-secondary systems. The concept of spectrum pooling was first mentioned in [8]. It proposes the use of cognitive radios as an extension of software radios for flexible pooling of radio spectrum using a new class of protocols called formal radio etiquettes. The mechanism proposed employs model based reasoning about users, multimedia content, and communications context.

Two different ways to implement spectrum pooling were discussed in [6]. One way is that the primary first tries to look for free channels among the channels that are not being used by the secondary. If no free channel is available, the primary may use any of the channels that the secondary is currently using. In the second method, the primary does not know about the existence of the secondary. The primary assumes that all the channels are available for its use. It is the secondary's responsibility to make sure that its usage pattern does not lead to unacceptable levels of interference. These two methods are evaluated in terms of spectrum utilization, blocking probability and forced termination probability in [6]. The physical layer detection mechanism of a renter is outlined in [5] assuming the owner is FDMA/TDMA based, like GSM. The paper points out that spectrum pooling seems to be technically feasible.

The model considered in this paper is the second method of [6] just mentioned above. That is, GSM does not know the existence of the secondary ad-hoc network. The ad-hoc network senses every GSM time-slot and uses only those slots that are currently not used by the GSM system.

The work in [7] concerns itself with the problem of physical layer sensing of spectrum by the renter to know the owner's channel usage information. This is needed by the renter so that it can avoid using the channels required by the owner. It uses the specific case of a GSM owner and an OFDM based WLAN for a renter.

To the best of our knowledge there exists little or no previous work on MAC layer protocols for spectrum pooling, which is the main goal of this paper.

The MAC protocol we propose in this paper is a essentially a multi-channel MAC protocol as it operates over the multiple channels of the GSM system. We first give an outline of the existing literature on multi-channel MAC.

A CSMA based multi-channel MAC protocol was proposed in [9, 10]. Every node builds a list of free channels by sensing all available channels. When a node has a packet to transmit, it checks to see if its free channel list is non-empty. If so, it tries to pick a channel (preference is given to that channel which the node used successfully in the most recent past), from the free channel list and transmits the packet after a delay period. There are channel reservations in this scheme.

A multi-channel MAC similar to 802.11 with channel reservations in the form of RTS and CTS was presented in [11]. Here again every node builds a list of free channels. When a node has a packet to transmit, it sends an RTS which contains the free channel information. On receipt of RTS, the receiver checks to see if there is a common entry (or channel) between its own free channel list and the list received from the sender. If there are no channels in common, the receiver refrains from sending a CTS. If there are common channels, the receiver picks the one that has the least received power according to its own sensing. On receipt of CTS, the sender transmits the data packet on the common channel. Finally, the receiver transmits ACK on the same channel. Other nodes receiving the sender's RTS defer using the control channel only until the time when CTS is expected. Other nodes receiving the CTS, defer using the selected data channel until the entire transmission is completed.

The above protocol has the following handshake mechanism: RTS-CTS-DATA-ACK which is similar to 802.11. A multi-channel MAC with a slightly different handshake mechanism is presented in [12] which uses RTS-CTS-RES-DATA-ACK handshake. The additional control message, viz. RES (reservation) is included. This RES is transmitted by the sender to indicate to its neighbors that it intends to use the selected data channel for a given period of time. As usual, CTS does this job too with the receiver's neighbors. Thus the superset of the neighbors of both the senders and receivers are informed about the impending transmission.

Our MAC is similar to the above MAC presented in [12] as we also use the RTS-CTS-RES-DATA-ACK handshake mechanism. But changes are needed to the MAC of [12] to cater to the needs of the unique GSM environment. The main driver for the new changes that are needed is the fact the GSM physical layer is also time-slotted with a slot size being too small to accommodate a single packet. A single GSM slot can accommodate only about 19 bytes at the GSM transmission rate, as a result data needs to be fragmented and transmitted over multiple time slots. This makes the error recovery procedures fairly complicated.

Thus, none of the existing multi-channel MAC protocols can be directly applied to our unique problem of GSM – Ad-hoc inter-working. Though we have borrowed the basic ideas from previous work, modifications are needed, and thus the birth of our new MAC.

3. SYSTEM ARCHITECTURE AND AS-MAC DESCRIPTION

3.1 OVERVIEW AND ASSUMPTIONS

We consider a single GSM cell having one base station and some GSM mobile stations (MSs). Figure 2 shows the overall system architecture. One BS, three MSs and three ANs are shown. The ANs constitute the secondary ad-hoc network in which the communication is peer-to-peer. It is seen that GSM communications and AN communications can coexist.

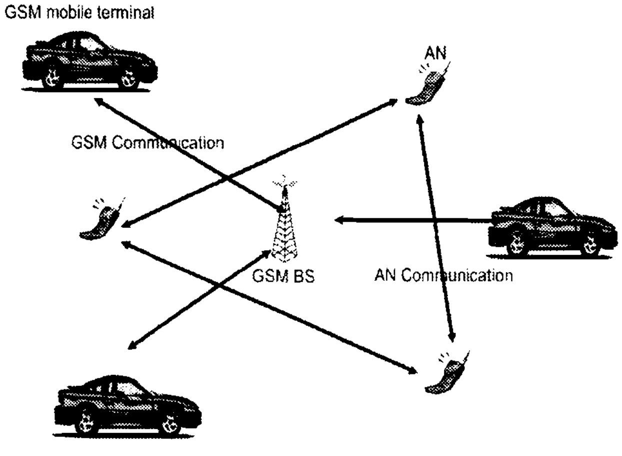

Figure 2. System architecture.

Only downlinks are utilized by ASN due to issues with using uplinks to be explained in section 3.2. Thus the uplink time-slots, even if they are free, are not used for transmission by AN in this paper. GSM power control is ignored.

ASN uses the FDMA/TDMA based GSM physical layer. It is assumed that the complete GSM slot availability is known to all AN nodes. That is, all the AN nodes know which frequency-timeslot combinations are being used at any given time by the GSM system. This can be achieved in reality by having the ANs sense carrier on all the downlink frequencies at the beginning of every time-slot for a short duration of say 5μs.

ASN is a packet oriented network. Peers in ASN communicate by stealing channels from GSM transparently. The existence of ASN is not known to GSM. Because ASN uses only the unused channels of GSM, ASN does not affect the network level performance of GSM.

It is also assumed that the ASN has a dedicated common control channel which is one of the GSM downlinks. Not all the slots on this channel need to be reserved for ASN. A few slots alone can be reserved for exclusive use by ASN. Also this common channel need not be part of GSM. It can be a separate out of band channel. Usage of a GSM downlink as the ASN control channel creates two issues. The first issue that the channel may not be available when needed because it may not be free at that time. The second issue is that even if the channel available, it may not have enough timeslots that are needed for signaling. Both these issues are resolved by assuming that this channel is available with sufficient time slots when needed by the ASN. For ASN to have better control over its signaling, which is very important for its proper functioning, we make this reasonable assumption that ASN has a reserved amount of control bandwidth exclusively available for its use.

It is also assumed that every ad hoc node (AN) has two half-duplex transceivers. One of the transceivers always operates on the fixed control channel while the other operates on the dynamically selected data channel.

3.2 SLOT BOUNDARY FOR AN

For the communication to take place among the ad hoc nodes (ANs), it is necessary that ANs be synchronized (sync) in time with the GSM base station (BS). Sync can be achieved in the same way as a GSM MS achieves sync, by decoding FCCH and SCH message transmitted by BS in the downlink signaling channels. Due to non-negligible propagation delays between BS and AN, the slot boundaries of an AN and BS can be different in several realistic cases. The slot boundaries are the same only if the AN is located very close to the BS. Figure 3 shows the relative timing of slot boundaries of an AN with reference to that of BS. From Figure 3, it is evident that the timing of AN is delayed by its propagation delay from BS. For uplinks, the time reference of AN will be advanced by its propagation delay.

ANs need to know their propagation delays from BS to use the uplinks. To understand this we draw the reader's attention to the concept of timing advance in GSM [3]. Remember that the GSM mobile stations (MSs) know the BS time reference by decoding the FCCH and SCH downlink signaling messages from BS. Moreover the propagation delay information is conveyed to the MSs by the BS through downlink signaling messages. In GSM, all the MSs advance their transmissions (with respect to BS time) by their propagation delays from BS. This ensures that the different signals transmitted by different MSs (in the same time-slots on different channels), reach the BS at the same time. This helps establish a common time frame among the MSs to use the uplinks channels properly.

Since the ANs have to adopt the same strategy to use the uplinks, it is important for ANs to know their propagation delays from BS. Since the inter-working between the two networks is assumed to be transparent there is no scope for BS-ASN signaling. Therefore, it is impossible for AN to know its propagation delay from BS. To be able to transmit on uplinks, AN has to know this delay so that it can properly schedule its transmissions. This is the primary reason that uplink usage by ASN is not considered.

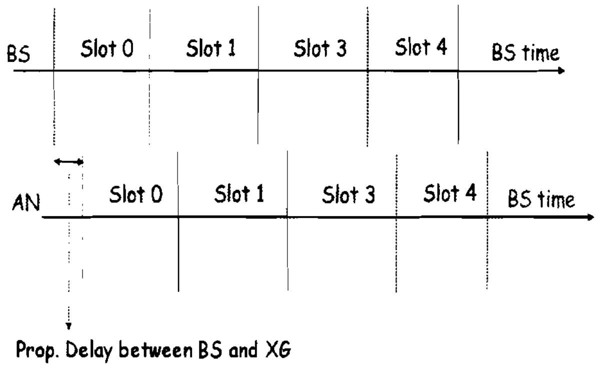

Figure 3. Slot boundary of AN with respect to BS

To transmit on downlinks, AN can begin transmitting as soon as the slot boundary begins. Since we assume that AN needs some time for channel/slot sensing, the first 5μs of every slot is reserved for this channel sensing. Thus AN can transmit any time after that in a given slot. Also it needs to end its transmission before the end of its slot boundary with a margin. This margin is mainly needed to reduce the amount of interference ASN can impose on GSM. Next we illustrate how AN transmissions can interfere with GSM.

3.3 INTERFERENCE SCENARIOS

To illustrate the interference caused by the transmission from an AN to a MS, three cases appear to be important. These are:

1. Both the AN and the MS are close to each other (either near the BS or elsewhere)
2. The MS is close to the BS but the AN is far away
3. Both the MS and the AN are far away from BS

These scenarios are shown in Figures 4, 5, and 6 respectively. The events indicated in the figures are self explanatory. The events are now explained only for the case 2, Figure 5. (1) The first slot is free in GSM, i.e. no GSM transmission takes place in it. (2) So AN can transmit in this slot to some other AN (which is not shown in Figure 5). (3) In the next slot an MS receives a signal from BS. (4) We see that the transmission from AN overlaps partially with the MS's reception from BS. In this example, the duration of overlap is shown as 200μs, which is twice the delay between AN and BS. For case 1, there occurs no overlap. For case 3, the amount of overlap is same as for case 2. Cases 2 and 3 represent the worst case overlap durations possible.

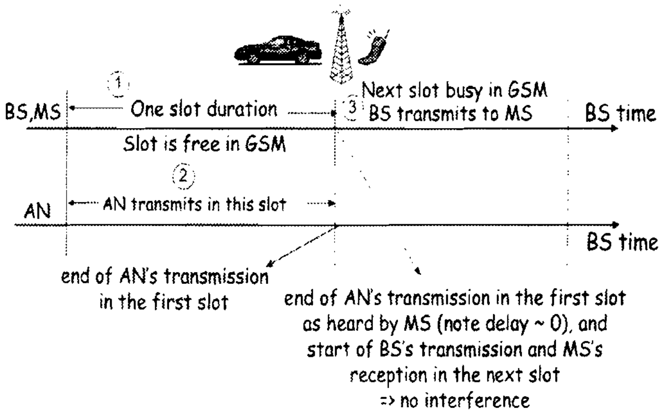

Figure 4: Interference between AN and MS (Case 1): Example assumes both AN and MS are close to the BS

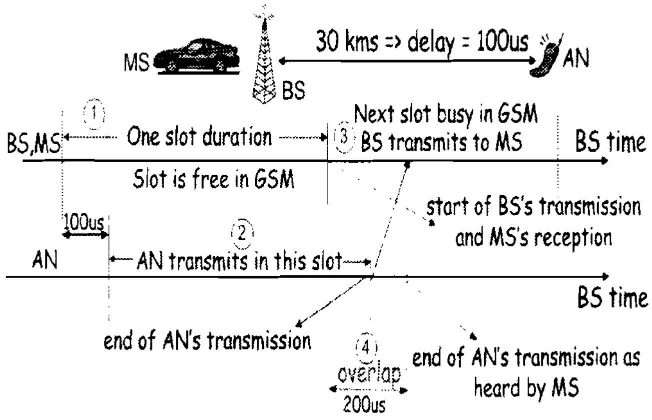

Figure 5: Interference between AN and MS (Case 2)

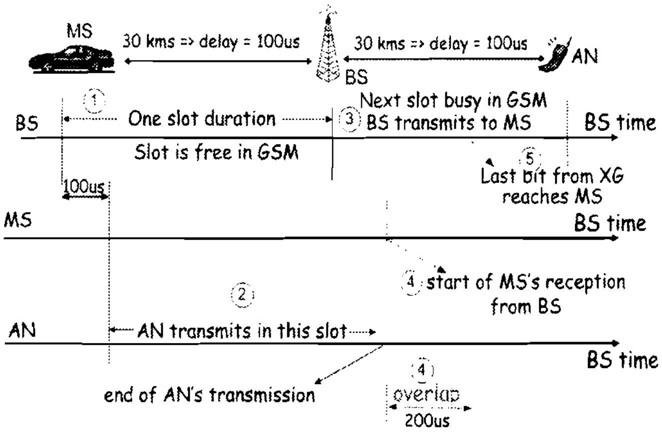

Figure 6: Interference between AN and MS (Case 3)

3.4 CONTROLLING DOWNLINK INTERFERENCE

For case 2, from Figure 5 it is seen that if AN finishes its transmission 200μs before its slot ends, it will cause no interference to the MS. This 200μs is nothing but twice the difference between the propagation delays from BS to itself and from BS to the concerned MS. Similar observation applies for case 3 as well.

This delay will make sure that the latest signal of AN's transmission will arrive at the concerned MS just before the

MS begins to receive a signal from the BS. But having a margin as high as 200μs is too wasteful of bandwidth, and can be avoided if MS is willing to tolerate a small amount of interference. All that needs to be done is to have a small margin say 25μs. This means that AN will not interfere with any MS about 12.5μs away, which translates to about 4 km. By the time the signal from AN travels 4 km it may get sufficiently attenuated so as not to cause appreciable interference to the MS. We use this argument as a means to reduce interference effects.

A quantification of downlink interference in now given. Consider that the ASN requires a transmission range of 500m and that the maximum cell radius of GSM is 30Km. Assume that the transmit powers of BS and XG are such that the received powers at their transmission ranges are the same. Further assume that the receive sensitivities and thresholds for successful reception of MS and XG are the same. These assumptions are made just for the sake of simplicity. The received power at a distance r from a transmitter is given in [13] is reproduced in Equation 1.

$$P_r = P_t\left(\frac{\lambda}{4\Pi r}\right)^n g_t g_r \qquad (1)$$

In the above, P_r is the received power, P_t is the transmitted power, λ is the carrier wavelength, g_t and g_r are the antenna gains at the transmitter and the receiver respectively. Since we consider a GSM system, say GSM 900, λ can be taken to correspond to a frequency of approximately 900Mhz. Assume n to be 3.

The attenuation in dB at a distance r can be easily obtained from the above equation. The attenuation at 500m (which is assumed to be the transmission range of XG) is 132dB. This means that the transmit power of XG is 132dB lower than that of BS. The received power at r = 500m for XG (which is the same the received power at r = 30Km for BS) is used as the base power over which attenuation (or signal to interference ratios are calculated). Note that any MS within a cell receives at least this much power from BS. As r increases, attenuation also increases, and the ratio of the base power to the received power at that point from XG increases. This means that the interference caused by XG to MS keeps going down as r. This is quantified in Figure 7 below. Note that the X-axis is in time units. The distance r, is translated to the equivalent time margin allowed, i.e., a distance of 1Km is translated to a time margin of 10μs (the propagation delay for traveling 1Km is 5μs). To understand why this translation is done refer to explanation given in the context of Figure 5. To brief it out, remember that there is no interference between XG and MS if twice the difference in their propagation delays from BS is less than the allowed margin. Interference is caused only after the signal transmitted by XG travels a distance corresponding to half of what is shown in the X-axis in Figure 7. The farther the signal travels the less would be the XG interference. Thus the figure shows the maximum amount of interference that XG can cause. It is assumed that XG, MS and BS are collinear. If that is not the case, the interference caused by XG would be even less than when they are collinear with the same set of parameters.

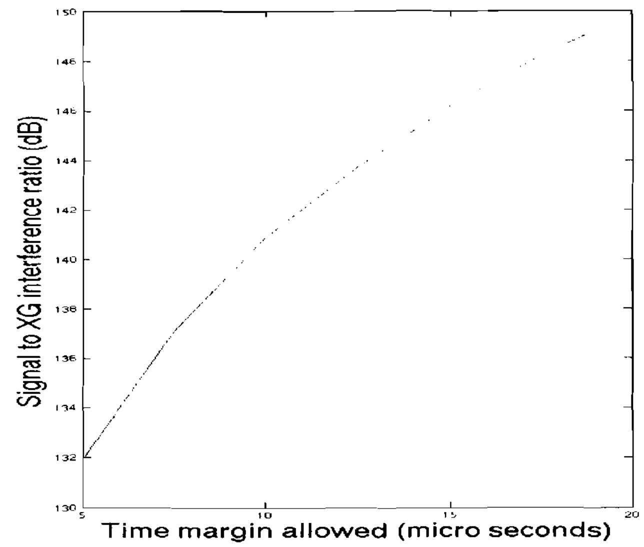

Figure 7. Interference caused by XG to MS Vs the allowed time margin by XG for its transmission.

Figure 7 shows the achieved signal to interference ratio (in dB) for various values of the time margin allowed by XG. Remember that XG stops its transmission before the end of the slot by this time margin. This helps in avoiding XG interference to a nearby MS that is receiving a signal from BS in the given time slot. Typically in cellular systems, MSs see a signal to (interference + noise) ratio of 8dB. If XG has to add only a negligible level of interference to the already existing interference, the ratio shown in Figure 7, has to well above 8dB, say 16dB. Thus a ratio of 148dB (132 + 16) is needed which is achieved by having a time margin of less than 20μs. This shows that the amount of interference added by XG can be easily controlled by having reasonable amounts of time margin.

3.5 AS-MAC DESCRIPTION

We are now ready to describe the operation of AS-MAC protocol beginning with the brief introduction of well known IEEE 802.11 MAC to help aid the reader in understanding the AS-MAC operation.

3.5.1 802.11 BASICS

The IEEE 802.11 MAC uses single-channel and has two modes of operation. The first mode is called the Distributed Co-ordination Function (DCF) and the second is the Point Co-ordination Function (PCF). DCF operation is mainly meant for the ad-hoc mode (where communication is peer-to-peer) while PCF operation is centralized and polling based and calls for the presence of a central entity called the Access Point (AP). Here we are concerned with the DCF functionality only.

In DCF, data transmission can follow an optional reservation mechanism. This reservation is meant to address the presence of hidden terminal problem. The hidden terminal problem is solved by RTS/CTS messages. The sender sends an RTS packet to the receiver. The RTS carries Network Allocation Vector (NAV) information. The receiver then sends back CTS packet which also carries the NAV information. The CTS packet essentially gives permission to the sender to transmit the data packet. Other nodes on hearing the RTS and CTS know that a communication is impending, hence defer their transmissions until the data transfer between the sender and the receiver is complete. Use of RTS and CTS messages increases the chances of successful data delivery once reservation succeeds. The sender on receipt of CTS sends the data (which can be fragmented if need be). Upon receiving the data, the receiver replies with an ACK. If the sender does not receive the ACK within a timeout period, it backs off and a subsequent retransmission of the data packet is scheduled.

3.5.2 AS-MAC BROADCAST PACKET TRANSFER

The AS-MAC protocol deals with the transfer of control traffic (broadcast packets) as well as unicast data packets. The MAC operation for the broadcast packet transfer at the AS-MAC layer is described first. Broadcast packet transfer takes place only on the control channel. This is done to avoid the dynamic channel selection (DCS) procedure since there is not one specific receiver in this case. Though the sender can select the channel, announce it in RTS and then transmit the data on the selected channel without expecting CTS, we do not recommend this approach for two reasons, 1) AN nodes that are already doing a data transfer with some other node cannot receive the broadcast packet. Since broadcast packets carry important control information for the higher layers, we would like the broadcast transfers to be heard by as many nodes as necessary; 2) the selected channel may become unavailable (the NAV may expire) before the transmission ends. Thus the packet will need to be retransmitted. Both the cases are not possible if the common control channel is used because it is fully reserved for exclusive use by ASN.

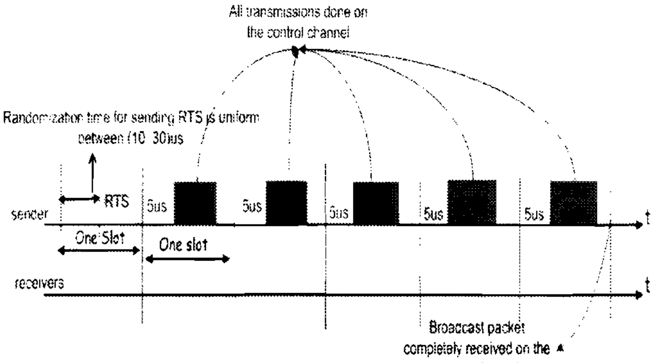

Figure 8. AS-MAC broadcast packet transfer.

Our MAC always gives higher priority to broadcast packets than data packets to ensure that the control information gets transferred as soon as possible. Broadcast packets carry crucial control information such as routing table updates, etc.

The broadcast packet transfer operation of AS-MAC is outlined in Figure 8. When a AN has a broadcast packet to transmit, it waits until a free control slot is available, then schedules a broadcast RTS transmission after a delay which is uniform in (10, 30)μs from the slot beginning. Since the RTS transmission window for broadcast and unicast do not overlap (unicast uses a different window), broadcast RTS is given higher priority over unicasting. When the scheduled time occurs for a broadcast RTS, the node senses carrier on the control channel to see if any other node has initiated transmission. If not, it transmits its RTS immediately. Otherwise, it defers until the next free control slot and tries to send RTS again. No back-off mechanism such as exponential back-off is used in broadcast RTS transmission. This is because broadcast traffic load is expected to be fairly low. No CTS is expected for a broadcast RTS. Once the RTS is transmitted, the source fragments the control data packet into several fragments and transmits them individually on the free slots of the control channel. No acknowledgements are expected for broadcast transfer either. When all the fragments corresponding to a packet are transmitted, the transmission ends. A receiver of a broadcast RTS prepares itself to receive data fragments (of the specified packet Id) on the control channel. The last data fragment from the source has a final flag set in the data header. This helps the receivers to know that the transmission has been completed.

The unicast packet transfer of AS-MAC is outlined next.

3.5.3 AS-MAC UNICAST PACKET TRANSFER

AS-MAC is a multi-channel MAC protocol that operates over the multiple channels of the GSM system. The unicast packet transfer is shown in Figure 9. The sender first waits for a free control slot and then sends RTS after a time interval chosen uniformly from the window (40μs – 140μs). Note that this window has absolutely no overlap with the broadcast RTS window. If CTS is not received by the sender or if the sender had to back off while waiting to transmit RTS, a binary exponential back off is initiated and RTS is tried again subject to a maximum number of retries. On receipt of RTS, receiver sends CTS which contains NAV info and also the channel selected for communication. The sender on receipt of CTS sends a reservation (RES) packet which also has NAV and channel info. The CTS and RES serve to inhibit other nodes from interfering with the dialog between the sender and receiver.

After the sender and receiver complete the RTS/CTS/RES handshake, the sender uses all the free slots on the selected channel for communicating with the receiver. The sender fragments the packet into as many fragments as necessary and transmits them one by one on the selected channel. To reduce the overhead due to acknowledgements, ACK from the receiver is expected only once for a given number of transmitted fragments called *"fragmentsPerBlock"*. It may be noted that in 802.11, an ACK is expected for every data fragment (when fragmentation is used). But it would be inefficient to apply the same technique to our case as it would lead to a significant bandwidth overhead for sending ACKs after every single data fragment. ACK is also expected when the sender transmits a fragment and does not have any more to transmit. An ACK flag is set in the fragment header when the sender wants an ACK after the receiver receives this fragment. The ACK from the receiver contains the fragment *Ids* of all the fragments received in the current cycle (cycle refers to the time period in which one train of fragments is sent by the sender and an ACK is sent by the receiver). On receipt of ACK, the sender updates its knowledge of what fragments have been received successfully by the receiver and sends those fragments which have not yet been received, beginning with the least numbered fragment (fragments are identified by a sequence number beginning from zero). This complicated error recovery scheme is adapted to enable better utilization of the precious wireless bandwidth. To the best of our knowledge this kind of recovery is not employed by earlier multi-channel MAC protocols.

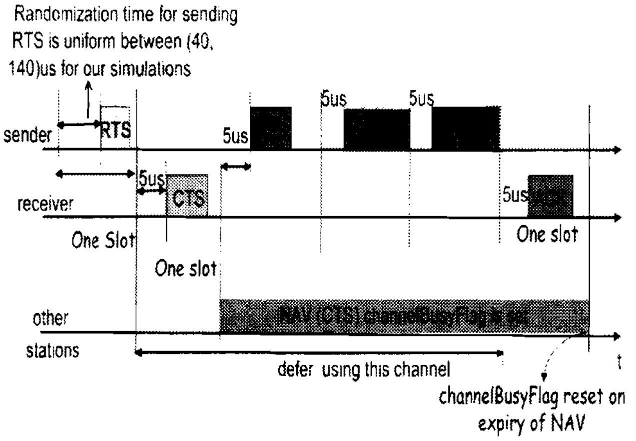

Figure 9. AS-MAC unicast packet transfer.

When the sender sends the last numbered fragment of the current packet, it sets a final flag in the fragment header. On receipt of this fragment, the receiver knows that the last fragment has been received. Thereafter on receipt of every fragment, the receiver checks to see if it has then received all the fragments. If so, the entire packet has been received and is passed on to the higher layer. Else, the receiver waits for more fragments from the sender.

On an ACK timeout by the sender, or if not all the fragments sent in the current cycle are ACKed by the receiver (indicating some loss), sender expects an ACK after transmitting just one data fragment in the next cycle, i.e. *fragmentsPerBlock* is set to 1. On receipt of an ACK that contains all the sequence numbers that the sender sent in the current cycle, *fragmentsPerBlock* is restored to its original value. This is done in order not to waste too much bandwidth due to ACK timeouts. Also requiring the receiver to send a quick ACK gets the sender and receiver synchronized again soon. After a given number of ACK timeouts, the transmission is aborted and the packet dropped by the sender.

We assume that all of these 19 bytes that can be accommodated in a single GSM time slot can be used by AS-MAC. In reality, there will be some overhead due to the physical layer header which we ignore. The maximum number of data bytes transported in a single slot (*dataBytesPerFragment*) is set to 15 bytes (plus 4 additional bytes for the header). Note that the CRC field has been ignored in the MAC PDU which can always be added at the expense of a few bytes. We point out here that it is possible to increase the number of bytes transmitted in a time-slot by adopting different physical modulation and/or coding techniques for signal transmission within a given time-slot.

3.5.4 NAV COMPUTATION

The computation of NAV is outlined now. NAV is sent in CTS and RES fragments. The number of free slots needed for transferring a data packet is calculated based on its size, *dataBytesPerFragment*, *fragmentsPerBlock* and some margin is allowed to take care of additional slots needed in case some fragments undergo transmission errors.

$$NAV = numDataSlots + numAckSlots + numM \arg inSlots \qquad (2)$$

$$numDataSlots = packetSize / dataBytesPerFragment \qquad (3)$$

$$numAckSlots = numDataSlots / fragmentsPerBlock \qquad (4)$$

$$numM \arg inSlots = \max \quad (1, \quad 12 * fragmentErrorRate * (numDataSlots + numAckSlots)) \qquad (5)$$

$$fragmentErrorRate = BER * 19 * 8 \qquad (6)$$

In Equation 6, *BER* is multiplied with 19*8 because a single slot is assumed to hold 19 bytes of data.

numDataSlots is the number of free slots needed for transmitting the data fragments alone. *numAckSlots* is the number of free slots needed for transmitting acknowledgements alone. A reasonable margin computed as *numMarginSlots* is added to the NAV to take care of additional free slot requirements due to retransmission caused by errors.

If the NAV expires at the sender or receiver before the transmission is completed, the transmission is aborted and the packet is reinserted into head of the queue (with priority next only to broadcast packets). This gives another quick chance for the packet to get transmitted. But if the transfer is aborted for other reasons such as ACK timeout, the packet is dropped.

4. SIMULATION AND RESULTS

A single GSM cell with one BS is considered in the simulation. There are 11 pairs of frequencies (uplinks and downlinks) that are available for use by ASN. Since only downlinks are used, this gives rise to 11 frequencies available. Of these one is set aside to act as the common control channel for ASN. Thus ASN has 10 frequencies for its use.

The usage of time-slots by GSM entities is assumed to have a uniform probability distribution with a given mean. Thus 50% usage means that, on an average every time-slot is busy for half of the total time duration considered. An on-off distribution characterized by exponential on-off times may be more suitable for an environment such as the GSM. But as we are primarily interested in quantifying the improvements in spectrum utilization, the distribution of GSM slot occupancy by primary is not of much concern. Note that the performance with respect to other metrics such as packet delay may be different when GSM slot occupancy is characterized by different distributions. The assumption of uniform slot occupancy also enables shorter simulation runs.

Twenty pairs of ANs are placed close to each other. Their placement is shown in Figure 10. Each pair corresponds to a flow or a session. Of each pair, one acts as the source and the other acts as the destination. Capture effect is not assumed. Noise is assumed to be negligible. We show results when AS-MAC operates by itself and also in conjunction with higher layers including the application layers such as FTP and CBR.

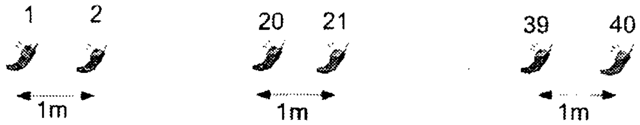

Figure 10.

AODV is used to find routes from the sources to the destinations. AODV uses the broadcast packet transfer service provided by AS-MAC for transmitting some of its control packets. The route timeout of AODV is set to a high value, therefore the routes once found during the beginning of simulation stay forever. No errors and collisions are assumed during this phase of route discovery. Our results focus only on the unicast packet transfer service of AS-MAC. Performance evaluation of broadcast packet transfer service is beyond the scope of this work. In our simulation we consider the nodes to be static.

Our results are organized for the following four cases. Cases 1 and 2 do not include the layers above MAC, such as network, transport and application layers. These two cases give a feel for the stand-alone performance of AS-MAC independent of any application characteristics. In Case 1, no channel errors are considered. This helps evaluate the best case performance of AS-MAC. In Case 2, channel errors are introduced to make the scenario more realistic. Cases 3 and 4 represent realistic scenarios and show results when CBR and FTP applications respectively are run over AS-MAC.

Spectrum utilization is now defined. GSM spectrum utilization is defined as the percentage of the number of slots used by GSM entities such as the BS and MS to the total number of slots in the period of measurement during the simulation. The period of measurement starts after a while after the beginning of simulation to let the system reach steady state and ends a while before the simulation ends. XG spectrum utilization is defined similarly as the percentage of the slots used by ANs to the total number of slots in the period of observation during the simulation.

Case 1: Only MAC Layer with no Errors

For this scenario, the protocol stack consists of only the GSM physical and the AS-MAC layers at all ANs. We assume that a single packet is always pending transmission at the MAC layer of every source AN. Remember that there are twenty such sources and twenty destinations. Thus there are 20 simultaneous MAC-level sessions competing for the ten available downlinks. Due to no errors, the margin required for packet transmission is set to just one slot. Packet size if set to 510 bytes. Table 1 shows the ASN utilization against GSM utilization as the number of free control slots is varied.

When 8 control slots are free, ASN utilization is about 41% when GSM uses 50% of the slots. This means that the AS-MAC achieves more than 80% efficiency in using the unused spectrum. When the number of control slots is reduced to four, still the usage is almost the same at 39%. This means that even 4 free control slots are enough to carry the control traffic generated by ASN. But when the free control slots are reduced to two, the utilization drops drastically to 28%. This is because though data slots are available, AS-MAC is not able to utilize them due to lack of enough control capacity to do the RTS/CTS/RES signaling. Similar effect can be observed when GSM uses 20% of the slots.

TABLE 1: ASN utilization Vs GSM utilization as the number of free control slots is varied

Number of free control slots	GSM utilization in %		
	20	50	80
8	64.23	41.16	16.84
4	60.48	39.81	16.67
2	28.32	28.34	15.78

When GSM uses 80% of the slots, the availability of free data slots becomes the bottleneck, so no significant degradation in utilization is seen when the control slots are reduced to two. Thus the amount of control bandwidth available can significantly affect the performance of AS-MAC. This needs careful design considerations.

Case 2: Only MAC layer with errors

This scenario is the same as the one for case 1, except that channel errors are introduced now. This error is modeled by a random number generated with a uniform probability of suitable mean. For every received packet, the value of the above random number is compared against the fragment error rate (which in turn is computed from BER). If the random number is smaller than the fragment error rate, an error is assumed and the packet is not delivered to AS-MAC. Else, the packet is assume to be free of errors and delivered to AS-MAC.

TABLE 2: ASN utilization Vs BER without and with control fragment errors.

BER	ASN utilization %	
	without control fragment errors	with control fragment errors
0.00001	34.33	34.27
0.0001	32.19	31.50
0.0002	29.87	26.17
0.0004	25.21	18.33
0.00066	18.81	12.24
0.00132	10.24	4.33

For this case the GSM utilization was set to 50%. ASN utilization is shown in Table 2 both when only data fragment errors are allowed and when both control and data fragment errors are allowed. RTS, CTS, RES, ACK and data fragments which have ACK and/or final flags set are considered as control fragments. Now margin is included when computing NAV as it is needed to handle cases where some fragments undergo errors and need to be retransmitted thereby requiring additional slots.

Note that the range of the error rate is very wide. Typical BER (Bit Error Rate) for wireless networks is of the order of 0.00001. Note that the protocol achieves a reasonable utilization of about 25/18% when the error rate is as high 0.0004. Note that the degradation due to the presence of errors in control packets is significant only for high error rates. The reason is that only a simultaneous loss of both CTS and RES (which is quite possible only at higher error rates) leads to a complete loss of NAV info by other nodes, thereby leading to collisions.

Case 3: CBR added with errors

Real time applications are modeled using CBR traffic sources. The purpose behind these experiments is to evaluate the suitability of AS-MAC for real-time applications such as VoIP. VoIP happens to be an interesting and one of the important applications supported over ASN.

TABLE 3: CBR statistics

Per session rate (kbps)	Application usage %	Avg. delay (ms)	Packet drop (%)
4	3.63	52.43	5.75
16	14.45	55.39	6.39
26.67	23.93	61.09	6.9
32	26.95	647.34	12.72

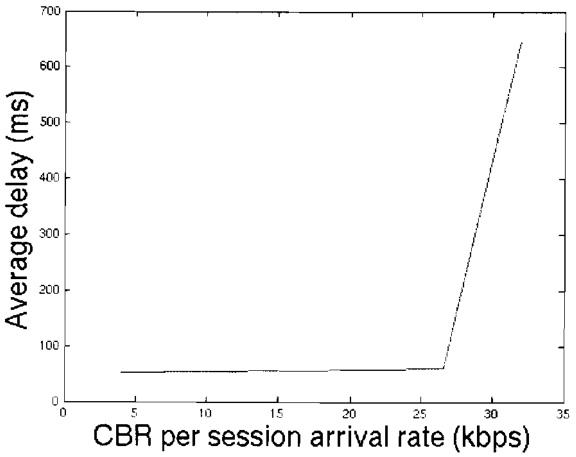

Figure 11: Average delay Vs arrival rate for CBR traffic.

A CBR session is established between each pair of ANs. Remember there are a total of twenty sources and twenty corresponding destinations. Thus there are a total of twenty CBR sessions competing for the ten available frequencies. At every AN, IP inserts packets ready for transmission into the MAC buffer whose size is set to five packets. This limit on the buffer size is placed to avoid excessive waiting times and hence delay for the CBR packets. BER is set to 0.0001. CBR sends packets of size 510 bytes at fixed intervals. Results are shown in Table 3. The average delay is plotted in Figure 11. Average delay is the mean application level delay over all successfully received packets. Packet drop rate is the percentage of the packets dropped at the application level normalized over the total number of packets sent by all the CBR sources.

The trends are quite intuitive. The spectrum utilization increases with increasing CBR rate and gradually saturates. This is expected as the AS-MAC is kept busier with increasing arrival rate. There is a sharp increase in delay after a rate of 26 kbps. This means that queue build up takes place around this rate and almost all the packets have to wait for the entire queue to become empty before they can be transmitted. The drop rate also increases dramatically after an arrival rate of 26% indicating that queue build up is taking place at that arrival rate. These results indicate that the network parameters such as the buffer size need to carefully set to optimize the performance of AS-MAC for a given situation.

Case 4: FTP added with errors

We now study the performance of AS-MAC in terms of spectrum utilization when FTP is used as the application. FTP is the most widely used data application, hence this scenarios exposes the performance of AS-MAC for a realistic scenario. There are twenty FTP sessions established between the twenty FTP sources and the corresponding twenty FTP destinations. FTP operates over TCP and always keeps it busy with packets. Therefore, this scenario also represents the best case scenario for the performance of XG-MAC in terms of spectrum utilization because packets are always available at the TCP layer for transmission. BER is set to 0.0001. Table 4 shows the results.

TABLE 4. ASN utilization Vs GSM utilization when FTP is used over AS-MAC.

GSM use %	Application use %
20	37.87
40	28.88
50	24.21
60	19.26
80	9.50

We see that for GSM use of 50%, Application level ASN use is 25%. Utilization at AS-MAC layer is about 28% (not shown in the table). This difference of 3% is mainly due to the additional overhead at the MAC layer caused by IP and TCP headers. Note that these results are for a BER of 0.0001 which is quite high for wireless networks under normal circumstances.

5. CONCLUSIONS

In this paper, we investigated the technique of spectrum pooling based on the concept of a primary-secondary system, to improve spectrum utilization. We studied the particular case of a GSM primary (using FDMA/TDMA physical layer) and an ad-hoc secondary network. We proposed a new MAC called AS-MAC for this particular scenario that enables the operation of the secondary network without affecting the primary. Our simulation results show that an appreciable amount of improvement in spectrum utilization can be achieved by our AS-MAC.

6. FUTURE WORK

In future we plan to investigate the following.

1. Issues involved in using the uplinks of the GSM system.

2. Performance of AS-MAC under multi-hop environments
3. Studying the performance and suitability of ad-hoc routing protocols to the GSM – Ad-hoc inter-working environment.
4. Study the inter-working of GSM primary with 802.11 secondary.

REFERENCES

[1] http://www.fcc.gov/sptf/files/SEWGFinalReport_1.pdf. FCC Spectrum policy task force, "Report of the spectrum efficiency group", November 15, 2002

[2] http://www.sharedspectrum.com/Measurements.htm

[3] M. Mouley and M. B. Pautet, "The GSM system for mobile communication", Cell & Systems, Palaiseau, France, 1992

[4] T.A. Weiss and F.K. Jondral, "Spectrum pooling: an innovative strategy for the enhancement of spectrum efficiency", IEEE Communications Magazine, Volume: 42, Issue: 3, March 2004, Page(s): 8–14

[5] F. Capar, I. Martoyo, T. Weiss, and F. Jondral, "Comparison of bandwidth utilization for controlled and uncontrolled channel assignment in a spectrum pooling system", IEEE 55th Vehicular Technology Conference, Volume: 3, Page(s): 1069–1073, May 2002

[6] M. Oner and F. Jondral, "Extracting the channel allocation information in a spectrum pooling system using a prefilter delay and multiply nonlinearity", IEEE Workshop on Statistical Signal Processing, Page(s): 46–49, Sept.28 - Oct.1, 2003

[7] J. Mitola, III, "Cognitive radio for flexible mobile multimedia communications", IEEE International Workshop on Mobile Multimedia Communications, (MoMuC '99) Nov. 1999, Page(s): 3–10

[8] C. E. Perkins, E. M. Royer, and S. R. Das, "Ad Hoc On-demand Distance Vector (AODV) Routing", IETF Internet draft, draft-ietf-manet-aodv-09.txt

[9] A. Nasipuri, J. Zhuang, and S. R. Das, "A multi-channel CSMA MAC protocol for multi-hop wireless networks", IEEE Wireless Communications and Networking Conference, Sept. 1999, Volume 3, Page(s): 1402 – 1406

[10] A. Nasipuri, S. R. Das, "Multi-channel CSMA with signal power-based channel selection for multi-hop wireless networks", 52nd IEEE Vehicular Technology Conference, Sept. 2000, Volume: 1, Page(s): 211 – 218

[11] N. Jain, S. R. Das, A. Nasipuri, "A multichannel CSMA MAC protocol with receiver-based channel selection for multi-hop wireless networks", Proc. Of Tenth International Conference on Computer Communications and Networks, 2001, Oct. 2001, Page(s): 432 – 439

[12] Shih-Lin Wu, Chih-Yu Lin, Yu-Chee Tseng, and Jang-Laing Sheu , "A new multi-channel MAC protocol with on-demand channel assignment for multi-hop mobile ad hoc networks", Proc. of the International Symposium on Parallel Architectures, Algorithms and Networks, Dec. 2000, Page(s): 232 – 237

[13] E. K. Wesel, "Wireless Multimedia Communications: Networking Video, Voice, and Data", Addison-Wesley, Reading, Massachussetts, USA, 1998

Amitabh Mishra received the M. Eng. degree in 1982 and the PhD degree in 1985 in electrical engineering, from McGill University in Montreal, and the MS degree in computer science in 1996 from the University of Illinois at Urbana-Champaign. He is currently an associate professor in the Bradley Department of Electrical & Computer Engineering at Virginia Tech. in Blacksburg, Virginia where he conducts research primarily in the wireless ad hoc and sensor networks, and mobile computing area. He is currently investigating scalability issues in wireless sensor networks, energy efficient MAC as well as cross-layer protocols for wireless sensor networks, intrusion detection and prevention algorithms for ad hoc networks, inter-working architectures for heterogeneous wireless networks, and quality of service issues for 3G/4G cellular networks. His other interests include wireless networking solutions for national critical infrastructure protection, fault tolerance, and software performance engineering. He has authored over 50 papers in journals and conferences and has seven patents granted and/or currently pending. From 1987 to 2000, he was a Member of Technical Staff at Lucent Technologies – Bell Laboratories in Naperville, Illinois where his focus was on application architecture performance. AIN, ISDN, ATM, GPRS, CDMA200 and UMTS were the major areas he worked on while with Bell Laboratories. He received his B. Eng. and M. Tech. degrees in electrical engineering from Jabalpur university and Indian Institute of Technology, Kharagpur in 1973, and 1975 respectively. He is a Chair of IEEE Communications Software, & Secretary of IEEE Internet Technical Committees, a Senior member of IEEE, a member of ACM and SIAM. He has served on the program committees of at least 20 conferences over the past three years.

Srivatsan Sankaranarayanan took his B.Tech in Electrical Engineering from the Indian Institute of Technology, Madras in 2000, and M.S degree in Electrical Engineering from the George Washington University, DC, USA. Currently he is a doctoral student at Virginia Polytechnic Institue and State University in the Wireless Internet Networking Group. His main research interests are in wireless networks architecture and protocols with emphasis on MAC.