

Performance of VoIP over Multiple Co-Located IEEE 802.11 Wireless LANs

An Chan, *Student Member, IEEE*, and Soung Chang Liew, *Senior Member, IEEE*

Abstract—IEEE 802.11 WLAN has high data rates (e.g., 11 Mbps for 802.11b and 54 Mbps for 802.11g), while voice streams of VoIP typically have low-data-rate requirements (e.g., 29.2 Kbps). One may, therefore, expect WLAN to be able to support a large number of VoIP sessions (e.g., 200 and 900 sessions in 802.11b and 802.11g, respectively). Prior work by one of the authors, however, indicated that 802.11 is extremely inefficient for VoIP transport. Only 12 and 60 VoIP sessions can be supported in an 802.11b and an 802.11g WLAN, respectively. This paper shows that the bad news does not stop there. When there are multiple WLANs in the vicinity of each other—a common situation these days—the already low VoIP capacity can be further eroded in a significant manner. For example, in a 5×5 , 25-cell multi-WLAN network, the VoIP capacities for 802.11b and 802.11g are only 1.63 and 10.34 sessions per AP, respectively. This paper investigates several solutions to improve the VoIP capacity. Based on a conflict graph model, we propose a clique-analytical call admission scheme, which increases the VoIP capacity by 52 percent from 1.63 to 2.48 sessions per AP in 802.11b. For 11g, the call admission scheme can also increase the capacity by 37 percent from 10.34 to 14.14 sessions per AP. If all the three orthogonal frequency channels available in 11b and 11g are used to reduce interferences among adjacent WLANs, clique-analytical call admission scheme can boost the capacity to 7.39 VoIP sessions per AP in 11b and 44.91 sessions per AP in 11g. Last but not least, this paper expounds for the first time the use of coarse-grained time-division multiple access (CoTDMA) in conjunction with the basic 802.11 CSMA to eliminate the performance-degrading exposed-node and hidden-node problems in 802.11. A two-layer coloring problem (which is distinct from the classical graph coloring problem) is formulated to assign coarse time slots and frequency channels to VoIP sessions, taking into account the intricacies of the carrier-sensing operation of 802.11. We find that CoTDMA can further increase the VoIP capacity in the multi-WLAN scenario by an additional 35 percent, so that 10 and 58 sessions per AP can be supported in 802.11b and 802.11g, respectively.

Index Terms—VoIP, multiple WLANs, CSMA, coarse-grained time-division multiple access, clique-analytical call admission control.

1 INTRODUCTION

VOICE-OVER-IP (VoIP) is one of the fastest growing applications for the Internet today. At the same time, driven by huge demands for portable access, the market for wireless Local Area Network (WLAN) based on the IEEE 802.11 standard is taking off quickly. Many city-wide deployments of WLAN are under plan. An important application over these networks will be VoIP over WLAN. A hurdle, however, is the low number of voice conversations that can be supported. As shown in previous investigations [1], [2], although in theory, many voice sessions can be supported in an 802.11b WLAN based on simplistic raw bandwidth calculation, in reality, only less than 12 can be accommodated.

Most of the prior investigative efforts [1], [2], [3], [4], [5] have been focused on the single isolated WLAN scenario. In practice, with the proliferation of WLAN these days, it is common to find numerous WLANs within a small geographical area—one only needs to do a cursory scan with a Wi-Fi-equipped personal computer to see the

considerable number of WLANs within a building. Recently, there has been much attention paid to multihop wireless mesh networks and VoIP over such networks [6], [7], [8]. In multihop wireless mesh networks, solutions to QoS of VoIP tend to be complicated as several higher layer issues, such as routing, have to be addressed. At the same time, most wireless networks currently deployed in the field have an infrastructure architecture in which terminals are connected to a base station in one hop. This paper is a first attempt to examine the VoIP capacity in the “multicell” environment in which many infrastructure WLANs are deployed in the same geographical area.

We find that the VoIP capacity is further eroded in the multicell scenario, and substantially so. For example, our NS2 [9] simulations show that the capacity of a 5×5 , 25-cell WLAN is only 1.63 VoIP sessions per access point (AP) in 802.11b and 10.34 sessions per AP in 802.11g. This dismal performance has important implications that deserve further attention in view of the accelerating productization of VoIP-over-WLAN technologies.

We identify the mutual interferences of the CSMA operation of adjacent cells as the major culprit for this dismal performance and provide solutions to alleviate the problem. Based on a conflict graph model, we set up a framework for call admission control to better manage the mutual interferences.

A major contribution of this paper is the proposal of a coarse-grained time-division multiple-access (CoTDMA) approach to alleviate multicell mutual interferences. In

- A. Chan is with the Department of Computer Science, University of California, Davis, 2231 Kemper Hall, 1 Shields Avenue, Davis, CA 95616. E-mail: anch@ucdavis.edu.
- S.C. Liew is with the Department of Information Engineering, The Chinese University of Hong Kong, 834 Ho Sin Hang Building, Shatin, Hong Kong. E-mail: soung@ie.cuhk.edu.hk.

Manuscript received 12 Dec. 2007; revised 7 Aug. 2008; accepted 9 Dec. 2008; published online 18 Dec. 2008.

For information on obtaining reprints of this article, please send e-mail to: tmc@computer.org, and reference IEEECS Log Number TMC-2007-12-0379. Digital Object Identifier no.10.1109/TMC.2008.176.

TABLE 1
Attributes of Commonly Used VoIP Codes

Codec	GSM 6.10	G.711	G.722	G.723.1	G.726-32
Bit rate (Kbps)	13.2	64	48/56/64	5.3/6.3	32
Packetization Interval (ms)	20	20	20	30	20
Payload (Bytes)	33	160	120/140/160	20/24	80

CoTDMA, the time dimension is divided into multiple coarse time slots. Each VoIP session is assigned a coarse time slot, and it makes use of the basic 802.11 CSMA protocol to contend for channel access with other VoIP sessions assigned to the same time slot. To isolate intercell interference, VoIP sessions of adjacent WLANs that interfere with each other are assigned different time slots. Coarse-grained time slots could be implemented using the sleep mode of 802.11, originally intended for power conservation purposes.

The theoretical call admission control framework of CoTDMA corresponds to a new class of graph-coloring problem that is distinct from that of the classical graph-coloring problem. With only three coarse-grained time slots, VoIP capacity per AP can be boosted to 10 and 58 sessions in 802.11b and 802.11g, respectively.

The remainder of this paper is organized as follows: Section 2 shows that the VoIP capacity can be eroded further in a significant way in the multicell setting. Section 3 explains how the CSMA protocol and co-channel interference would affect the VoIP capacity. We present a call admission strategy based on clique analysis of a graph-theoretic formulation to confine intercell interferences. Section 4 considers using a time-dimension approach in conjunction with the basic 802.11 CSMA to further improve the VoIP capacity. Section 5 concludes the paper.

2 LOW VoIP CAPACITY OVER MULTIPLE WLANS

2.1 VoIP Attributes

VoIP packets are streams of packets containing encoded voice signals. There are different codecs for encoding voice signals (Table 1). Take GSM 6.10 as an example. The voice payload is 33 bytes, and 50 packets are generated in each second. After adding the 40-byte IP/UDP/RTP header, the minimum channel capacity required to support a GSM 6.10 voice stream in one direction (uplink or downlink) is 29.2 Kbps.

An 802.11b WLAN in theory can support nearly 200 VoIP sessions (divide 11 Mbps by two times of 29.2 Kbps) and an 802.11g WLAN, more than 900 sessions (divide 54 Mbps by two times 29.2 Kbps). However, prior investigations have shown that the actual VoIP capacity is severely limited due to various inherent header and protocol overheads. With the GSM 6.10 codec, for example, only 12 (60) VoIP sessions can be supported in an 802.11b (802.11g) WLAN [1], [2].

Besides GSM 6.10, other advanced VoIP codecs may help to increase the VoIP capacity over single isolated WLAN with techniques such as silence compression and

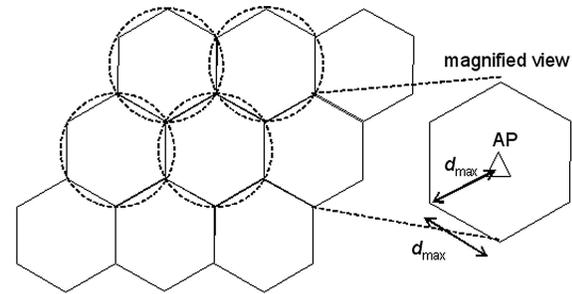


Fig. 1. A 3×3 multicell topology.

packet loss concealment. However, these codecs do not tackle the issue of co-channel interference, which arises when multiple WLANs are in the vicinity of each other. The large VoIP capacity penalty when VoIP operates over multiple WLANs, as mentioned in Section 1, remains. We will demonstrate the presence of this large capacity penalty in various codecs in the next section. For concreteness and direct comparison with the results of single isolated WLAN in [1], [2], we assume GSM 6.10 in the remaining discussion.

Voice signals could also be encoded with either constant bit rate (CBR) or variable bit rate (VBR) at the source. VBR takes advantage of silence periods in conversation to save bandwidth. Although VBR can support more VoIP sessions, the large capacity penalty when VoIP operates over multiple WLANs remains. We will investigate this penalty for both VBR and CBR in this paper. For explanation of concepts and solutions proposed, however, we mainly focus on CBR. In the case of VBR, the main ideas of our solutions do not change except that we have to add a probabilistic model into our solutions to take advantage of bandwidth conservation in VBR.

In this paper, we assume the allowance of 3 percent packet loss rate. The exact tolerable packet loss rate depends much on the codec used. Commercial products typically claim that their VoIP packet loss rate can be kept within 1-5 percent, and even less than 1 percent [10]. Codecs with packet loss concealment can tolerate larger packet loss rates. Loss rate of 3 percent is a common requirement for voice transmission [11]. Given the 3 percent loss allowance, the minimum channel capacity requirement is 28.32 Kbps for GSM 6.10 codec. If both the uplink and downlink of a VoIP session can have throughputs exceeding this benchmark, we say that the VoIP session can be supported in the WLAN.

2.2 Low VoIP Capacity over Multiple WLANs

To evaluate the VoIP capacity over multiple WLANs, we model a "WLAN cell" with a regular hexagonal area with side length 250 m. An AP is placed at the center of the cell. Any wireless client station inside the cell will be associated with the AP. The longest link distance (d_{max}) is therefore 250 m, which is also the data transmission range ($TXRange$) for 802.11 assumed in NS2. The transmission ranges of APs in WLANs partially overlaps. This is also the case in practice. The circles in Fig. 1 represent the coverage of

TABLE 2
VoIP Capacity over $D \times D$ Multicell Topologies

D	1	2	3	4	5
Avg. $C_{D \times D}$ in 11b	12.0	12.3	20.0	30.0	40.8
Avg. $C_{D \times D}$ in 11g	55.0	79.6	131.8	207.6	258.4

certain WLANs. By placing the cells side by side, we form a $D \times D$ multicell topology, where D is the number of cells on each side. Fig. 1 shows a 3×3 multicell topology.

We consider the use of the basic mode [12] of 802.11 in this paper since the short VoIP payload does not warrant the use of RTS/CTS. We assume carrier-sensing range (CSR_{range}) of 550 m, the default value in NS2.

We ran simulation experiments on the $D \times D$ multicell topology for $D = 1, 2, 3, 4,$ and 5 . In each run, wireless client stations (VoIP sessions) are added one by one randomly assuming uniform distribution. With each additional VoIP session, NS2 simulation is run and the throughput of each link recorded. When the next newly added VoIP session causes violation of the packet loss rate requirement (3 percent) by at least one of the sessions, we say that the capacity limit has been exceeded. This corresponds to a simplistic call admission scheme in which upon unacceptable performance caused by the newly added session, the newly added session will be dropped, and no more future sessions will be accepted. We will later consider a cleverer call admission scheme based on a clique analysis of a conflict graph so that we can “predict” the performance before deciding whether to admit a call. Table 2 summarizes our simulation results under GSM 6.10.

In Table 2, $C_{D \times D}$ is the total number of VoIP sessions that can be supported in a $D \times D$ multicell topology. Obviously, as D increases, more VoIP sessions can be supported. We further calculate $C_{AP,D}$, the per-AP capacity in a $D \times D$ multicell WLAN, defined as follows:

$$C_{AP,D} = C_{D \times D} / D^2. \tag{1}$$

We plot $C_{AP,D}$ against number of cells D^2 in Fig. 2. We find that as number of cells increases, per-AP capacity decreases. When the number of cells is 25, i.e., $D = 5$, only 1.63 VoIP sessions can be supported by each AP in 802.11b. Compared with the single-cell scenario, where each AP can support 12 VoIP sessions in 802.11b, this is a rather large penalty! Similar capacity penalty is also observed in 802.11g, where only 10.34 sessions per AP can be supported when $D = 5$, as opposed to 55 in the single isolated WLAN case when $D = 1$.¹

Similar large capacity penalties are also found in other VoIP codecs and VBR encoding. Table 3 shows the per-AP VoIP capacity in a 1×1 (single isolated) WLAN and that in a 5×5 multicell WLAN when different VoIP codecs and bit

1. VoIP capacity of a single-cell 802.11g WLAN measured in simulation is 55, not 60 as calculated in [1], [2]. This is due to the smaller minimum contention window size in 802.11g. But for consistency of comparison, we still take “60 VoIP sessions” as the theoretical upper bound for VoIP capacity over single isolated 802.11g WLAN in our following discussions.

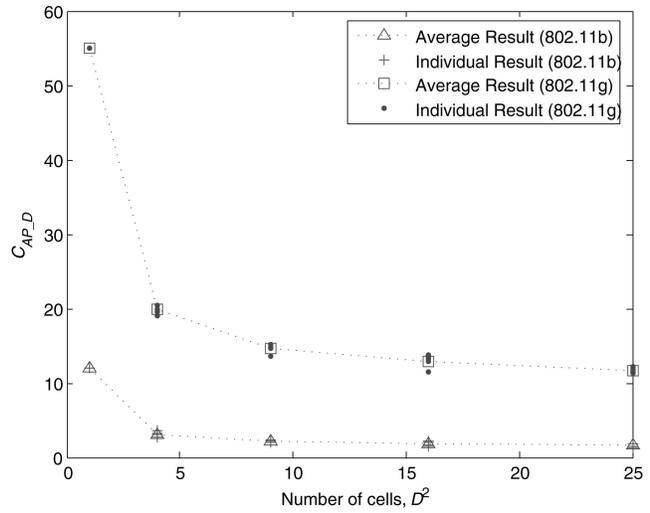


Fig. 2. Per-AP capacity of $D \times D$ multicell WLAN.

rates (CBR and VBR) are used. For VBR, Brady’s VBR model is assumed. For all codecs, regardless of whether CBR or VBR is used, large capacity penalties are incurred in the multiple-WLAN scenario.

We also note in passing that for networks larger than 5×5 , most cells will be surrounded by six adjacent cells and there will be proportionately fewer cells at the boundary, where there is less interference from other cells. One can therefore expect VoIP capacity per AP to drop even further. Indeed, for GSM 6.10 CBR VoIP over a 10×10 802.11b network, our simulation shows that the per-AP capacity is only 1.22 sessions.

2.3 Applying Frequency-Channel Assignment

To reduce mutual interference, a quick solution is to assign different frequency channels to different cells. If there are enough frequency channels, the neighbor cells could be assigned different frequency channels. This boils down to the same situation as in the single-cell case so that the per-AP VoIP capacity in the multicell case is the same as that in the single-cell case.

IEEE 802.11b/g has only three orthogonal frequency channels, and this is not sufficient to completely isolate co-channel interference between cells. Fig. 3 shows that if we have only three frequency channels, then the nearest

TABLE 3
Per-AP Capacity When VoIP in Different Codecs Operates over 1×1 and 5×5 802.11b WLANs

Codec	GSM 6.10		G. 711		G. 722*		G. 723.1**		G. 726-32	
	CBR	VBR	CBR	VBR	CBR	VBR	CBR	VBR	CBR	VBR
1-by-1 WLAN	12.0	18.0	11.0	17.0	11.0	18.0	18.0	27.0	12.0	17.5
5-by-5 WLAN	1.63	2.98	1.48	2.72	1.56	2.80	2.76	4.25	1.64	2.80

* Bit rate of 48 Kbps of G. 722 is simulated

** Bit rate of 5.3 Kbps of G. 723.1 is simulated

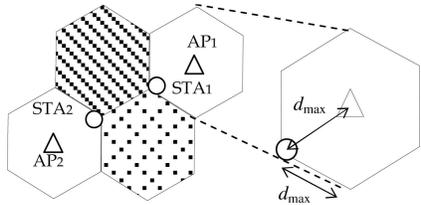


Fig. 3. Three-channel assignment in multicell WLAN.

distance between two cells using the same channel (e.g., the two unshaded hexagons in Fig. 3) is the same as the maximum link length within a cell, d_{\max} . Since these two cells may interfere with each other, the carrier-sensing range ($CSRange$) should be larger than $3d_{\max}$ to avoid hidden-node [13] collisions between the two cells. As defined in [13], hidden nodes (HN) occur when transmitters are outside of the $CSRange$ of each other but their receivers are close enough for packet collisions. That is, the carrier-sensing mechanism fails to prevent packet collisions because the transmitters cannot “hear” each other. To see this, consider link (AP1, STA1) and link (AP2, STA2) of the two cells in Fig. 3, where the distance between STA1 and STA2 is d_{\max} . The transmission of an ACK by STA1 to AP1 will interfere with the reception of DATA from AP2 to STA2. If $CSRange$ is set to $2d_{\max}$ for preventing packet collision within the same cell only, AP1 and AP2 become HN to each other and they may transmit at the same time. This may cause collisions in STA1 or STA2. To prevent the HN collision, the two APs should be able to carrier sense each other so that one would not start a transmission while the other is transmitting. Hence, $CSRange$ of $3d_{\max}$ is needed between these two links. Since a cell must now share airtime with other cells, the VoIP capacity per AP cannot be the same as that in the single-cell case.

IEEE 802.11a, on the other hand, has 12 orthogonal channels. Fig. 4 shows that a seven-channel assignment is sufficient for complete isolation of co-channel interference. The nearest distance between two cells using the same channel is $2.65d_{\max}$, which is larger than the minimum $CSRange$ $2d_{\max}$ used to prevent collisions *within a cell*. Thus, with a seven-channel assignment, co-channel interference can be completely isolated. However, if we simply use seven overlay networks in each cell (put seven APs inside each cell and operate in different channels), the number of VoIP sessions supported in each cell can be increased by seven times compared with that in the single-channel multicell topology. Therefore, we find that the channel assignment in Fig. 4 actually may not improve the VoIP capacity on a per-frequency channel basis, although on a per-AP basis, it does. Furthermore, not all countries permit unlicensed use of the spectrum of IEEE 802.11a and the available spectrum varies widely [14]. As a result, IEEE 802.11a is not commonly deployed. In short, the case of 802.11b/g, in which there are not sufficient frequency channels for complete elimination of co-channel interference, will remain to be of much practical interest.

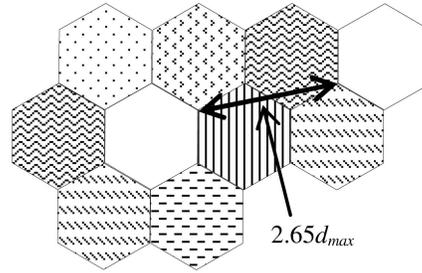


Fig. 4. Seven-channel assignment in multicell WLAN.

3 CLIQUE ANALYSIS AND CALL ADMISSION

To understand the cause for the heavy performance penalty in the multicell scenario, we consider here a clique analysis based on a graph model that captures the conflict and interference among the nodes. The clique analysis also suggests a call admission methodology. With this call admission scheme, the VoIP capacity can be increased to 2.48 sessions per AP from 1.63 in the case of 5×5 , 25-cell 802.11b WLAN. This constitutes a 52.1 percent improvement. In the 5×5 802.11g WLAN, this call admission scheme increases the per-AP capacity by 36.75 percent from 10.34 to 14.14 sessions.

3.1 Mapping Power Threshold to Virtual Distance

Before going into clique analysis, we would like to explain the relationship between “received signal power” and the concept of “distance” used in the algorithms of this paper.

In real equipment, the operating parameters are based on “power thresholds” rather than “distances.” For example, for carrier sensing and detection, it is the power received that matters rather than the distance. The implementations of the schemes in this paper are compatible with the power-threshold interpretation. For convenience, however, we will use distances (e.g., $TXRange$ and $CSRange$) to describe the system operation. Thus, the term “distance” is to be interpreted in a virtual sense, as explained below.

Two wireless stations i and j are said to be separated by a virtual distance $d_{i,j}$ if the power they receive from each other is $P_{i,j} = P_{j,i} = k/d_{i,j}^\alpha$, where k is a constant and α is a “reference” (not the actual) path loss exponent, assuming that all stations use the same transmit power. Applying common k and α to the whole network, we can then derive the virtual distance $d_{i,j}$ from the measured power transferred $P_{i,j}$. Given a power threshold requirement, there is then a corresponding virtual distance requirement. When we say that the $CSRange$ is set to d_{CS} , we mean that the carrier-sensing threshold power is set to $P_{CS} = k/d_{CS}^\alpha$, where α is the nominal constant adopted (e.g., $\alpha = 4$).

In the subsequent discussions, we assume that there is an underlying scheme to find out $P_{i,j}$ (hence, $d_{i,j}$) so that we can assign system resources (e.g., AP association and time-slot assignment) according to $d_{i,j}$. To limit our scope,

however, we will not discuss the $d_{i,j}$ discovery algorithms here. The reader is referred to [15] for possible schemes. In short, the logical correctness of the implementation of our schemes in this paper does not depend on spatial homogeneity. The performance, however, does depend on the α being assumed.

3.2 Conflict Graphs and Cliques

In our conflict graph model, vertices represent VoIP sessions (wireless links). An edge between two vertices means that the two VoIP sessions compete for the airtime. In other words, they cannot transmit packets at the same time. There are two scenarios under which they cannot transmit together: 1) first, nodes of the two sessions that are within the CSR_{Range} of each other will be prevented by the 802.11 protocol from transmitting together and 2) even if the two sessions are not within each other's CSR_{Range} , there may be mutual interference between them so that either one or both of their transmissions will fail if they transmit together: the well-known hidden-node problem [13] mentioned in Section 2.3. In either case, an edge is drawn between the two vertices. A clique is a subset of vertices in which there is an edge between any pair of vertices [16]. The vertices in a clique compete for common airtime. In particular, the sum of the fractions of airtimes used by the vertices should not exceed one.

In the single-cell scenario, all client stations are associated with the same AP. So, edges should be drawn among all vertices of the same cell. In an 802.11b single-cell WLAN, to ensure that all existing VoIP sessions have acceptable performance, the maximum clique size is 12. This is because 12 VoIP sessions will fill up all airtimes [1], [2]. In an 802.11g single-cell WLAN, the maximum clique size is 60.

In a multicell WLAN, multiple cliques can be formed because links in different cells may be able to transmit together. Let us examine cases 1 and 2 mentioned above. For case 1, CSR_{Range} is usually a fixed value (assuming no power control). By default, it is 550 m in NS2. For case 2, we consider the *Interference Range (IR)* defined as follows:

$$IR_k = (1 + \Delta)d_k, \quad (2)$$

where IR_k is the *Interference Range* of a node k (it can be a client station or an AP), d_k is the length of the link associated with the node k , and Δ is a distance margin for interference-free reception with typical value of 0.78 [9], [13]. Within a radius of IR_k , any other transmission will interfere with the node k 's reception of the signal. The maximum link length within a cell is $d_{max} = 250$ m (see Fig. 1). The corresponding maximum IR is therefore $IR_{max} = 1.78 \times 250$ m = 445 m. In a multicell WLAN, two cells could be separated by a distance larger than both CSR_{Range} and IR_{max} . So, there is no edge between vertices of these two cells.

From the discussion above, we know that there is a maximum clique size which cannot be exceeded if satisfactory performance of VoIP sessions is to be attained. To increase the VoIP capacity over multiple WLANs, we

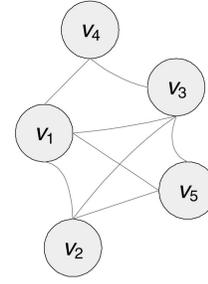


Fig. 5. An example of a conflict graph.

have to pack the VoIP sessions (vertices) efficiently by applying a call admission control. We discuss this call admission control in the next section.

3.3 Clique-Analytical Call Admission Control

We consider a call admission control mechanism based on the clique size of the conflict graph. Let E_{v_i} be the set of neighbor vertices with which vertex v_i has an edge. Let K_{v_i} be the set of all maximal cliques $C_{v_i,x}$ to which v_i belongs, where $x = 1, 2, \dots, |K_{v_i}|$ is the index of the cliques. Any clique in K_{v_i} must satisfy (3) below for it to be "maximal" and not contained in another clique [17]:

$$C_{v_i,x} \not\subset C_{v_i,y}, \quad x \neq y, \quad 1 \leq x, y \leq |K_{v_i}|. \quad (3)$$

Let m_{v_i} be the size of the largest clique in K_{v_i} :

$$m_{v_i} = \max_x |C_{v_i,x}|. \quad (4)$$

Fig. 5 gives an example of a conflict graph, where vertex v_1 has the following parameters:

$$\begin{aligned} E_{v_1} &= \{v_2, v_3, v_4, v_5\}, \\ K_{v_1} &= \{\{v_1, v_2, v_3, v_5\}, \{v_1, v_3, v_4\}\}, \quad |K_{v_1}| = 2, \\ \text{i.e., } C_{v_1,1} &= \{v_1, v_2, v_3, v_5\}, \quad C_{v_1,2} = \{v_1, v_3, v_4\}, \\ m_{v_1} &= 4. \end{aligned}$$

The pseudocode of the admission control algorithm is given in Algorithm I. There are three procedures in the algorithm. When there is a new call request (i.e., a new vertex v_i (VoIP session) wants to join), Procedure A is first executed, wherein a copy of the state $(K_{v_j}, m_{v_j}) \forall v_j$ is first saved in case the admission of v_i fails and we have to revert to the original state. After that, Procedure B is executed. Procedure B updates the state (K_{v_j}, m_{v_j}) assuming the addition of v_i . To satisfy (3), a function $NO_REDUNDANCY(K_{v_j})$ is called. Algorithm II gives the pseudocode of the function $NO_REDUNDANCY(K_{v_k})$. During the updating, Procedure B continually checks to see if a predetermined maximum clique size C_{max} is exceeded so as not to violate the loss rate requirement. If so, the algorithm is terminated and v_i is rejected; in which case, the state saved in Procedure A is restored. If Procedure B successfully runs to the end without C_{max} being exceeded, Procedure C is executed. Procedure C admits the new vertex v_i and calculates (K_{v_i}, m_{v_i}) .

Algorithm I**Procedure A**

Keep a copy of the state (K_{v_j}, m_{v_j}) for all existing v_j ;

Perform Procedure B;

Procedure B

```

for each  $v_j \in E_{v_j}$  {
  for each  $C_{v_j,x} \in K_{v_j}$  {
    if  $C_{v_j,x} \subset E_{v_i}$ 
      then add  $v_i$  to  $C_{v_j,x}$ ;
    else
      add  $\{C_{v_j,x} \cap E_{v_i}, v_i\}$  to  $K_{v_j}$ ;
  }
   $K_{v_j} = NO\_REDUNDANCY(K_{v_j})$ ;
  update  $m_{v_j}$ ;
  if  $m_{v_j} > C_{max}$ 
    then reject  $v_i$ ;
    revert the state using the copy stored in Procedure A;
    break out of Procedure B, algorithm is terminated;
} // All  $v_j \in E_{v_j}$  have been looked at if the algorithm comes to this point
Perform Procedure C;
```

Procedure C

```

 $v_i$  is admitted;
 $K_{v_i} = \emptyset$ ;
for each  $v_j \in E_{v_i}$  {
  if  $v_i \in C_{v_j,x}$ 
    then  $C_{v_j,x} = C_{v_j,x} \setminus v_i$  and  $K_{v_j} = K_{v_j} \cup C_{v_j,x}$ ;
}
 $K_{v_i} = NO\_REDUNDANCY(K_{v_i})$ ;
compute  $m_{v_i}$ ;
```

TABLE 4

Experiment Results of Applying the Admission Control Algorithm on 5×5 Multicell WLAN

C_{max}	Total Run-time (s)	Average No. of Sessions Admitted	Average Runtime (s)
8	48.6	62.0	0.160
12	112.1	70.4	0.374

vertices within a clique should not exceed the total allocated airtime. The details of the generalized clique-analytical call admission control scheme can be found in the Appendix. For the rest of this paper, we focus on the fixed C_{max} case.

We have performed an experiment in MATLAB to measure the execution time of the call admission control algorithm. The experiment assumes the 5×5 multicell topology setting in Section 2.2. For simplicity, we first consider the single-frequency channel case in which all cells are assigned the same frequency. The next section deals with the multifrequency case.

We applied the algorithm to identify which VoIP sessions out of a total of 300 randomly placed (with uniform distribution) potential sessions could be admitted in a 5×5 802.11b multicell WLAN with the C_{max} clique-size constraint. Unlike the simplistic call admission scheme earlier, here when a session is rejected, the call admission scheme continues to consider a next session if the 300 sessions have not been exhausted. We ran several sets of experiments for different random node distributions in a 5×5 WLAN. We used an ordinary personal computer with 3.2 GHz CPU and 1 GB RAM to perform the experiment. The results for $C_{max} = 8$ and 12 are shown in Table 4. The total runtime is the total time needed for the algorithm to go through all the 300 VoIP sessions. The average runtime is the time needed to admit or reject a call (Total runtime/300). We see that although the general clique problem is NP-complete, the execution time of our algorithm on the conflict graph that models 802.11 networks is acceptable.

Based on the call admission results in MATLAB, we then used NS2 to verify whether the admitted calls meet the maximum 3 percent packet loss rate requirement in the simulation. From Table 4, we find that when C_{max} is 8, the average number of VoIP sessions admitted by our admission control algorithm is 62.0. This is the average number of five runs of the MATLAB experiments. In each run, all the admitted VoIP sessions can meet our packet loss rate requirement.

However, if instead of setting C_{max} to 8, we set it to 12, then nearly one-third of VoIP sessions cannot meet the loss requirement. It is interesting that for the large-scale multicell case, the maximum clique size that should be imposed is 8 rather than 12 (recall that 12 is the maximum clique size in the single-cell topology if 802.11b is assumed) if loss rate requirement is to be satisfied. This is perhaps due to the interaction and "coupling" among different cliques caused by the 802.11 MAC protocol. In other words, 802.11 MAC may not achieve perfect scheduling in which the airtime usage within each clique is 100 percent tightly packed. This motivates us to explore the use of time-slot scheduling for performance improvement in Section 4.

Algorithm II

```

NO_REDUNDANCY( $K_{v_k}$ ) {
  for each pair  $(C_{v_k,x} \in K_{v_k}, C_{v_k,y} \in K_{v_k}), x \neq y$  {
    if  $C_{v_k,x} \subset C_{v_k,y}$ 
      then delete  $C_{v_k,x}$ ;
  }
  return ( $K_{v_k}$ );
}
```

It is worth noting that if all VoIP sessions use the same VoIP codec and do not use Auto Rate Fallback (ARF), which decreases the data transmission rate if the channel quality is not good, C_{max} can be a predetermined fixed value. However, if ARF is assumed, or different VoIP sessions in different cells use different codecs, then they have different packetization intervals, different transmission rates, etc. In that case, C_{max} is no longer a fixed value. The generalized version of this clique-analytical call admission control is to accumulate the airtime of vertices (VoIP sessions) in the same clique. To ensure that all the admitted VoIP sessions have acceptable performance, the sum of all the airtimes of

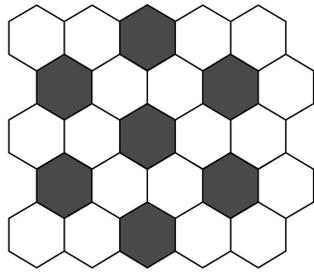


Fig. 6. Three-frequency-channel assignment is applied on multiple WLANs.

With $C_{\max} = 8$, the 5×5 multicell WLAN can support 2.48 sessions (62.0/25) per AP, yielding a 52.1 percent improvement over the simplistic call admission scheme in Section 2. The clique-analytical call admission control works similarly well in 802.11g WLANs. For the 5×5 multicell 802.11g WLAN, with $C_{\max} = 44$, the per-AP capacity can be increased from 10.34 to 14.14 VoIP sessions, yielding a 36.75 percent improvement.

The call admission control scheme above can be implemented by a central entity which connects all the APs through a backbone network. Clients report the signal strength from the new VoIP session they sensed to the APs to which they associate. The APs then report the power information to the central entity. Based on the power information, the central entity constructs the corresponding conflict graph and carries out the computation for the purpose of call admission control. Even if there are multiple vendors, we could still have a “ClearingHouse” (third party) to be the central entity as it has the attractive feature that helps service providers to admit new incoming calls without disrupting existing ongoing calls. A decentralized version of our call admission scheme will be an interesting direction for future work.

3.4 Clique-Analytical Admission Control in Three-Frequency-Channel WLANs

We now explore the impact of multiple frequency channels on VoIP capacity. The availability of multiple frequency channels allows us to separate the cells using the same frequency by a longer distance. Farther separation of cells leads to less conflict among transmissions of different VoIP sessions (although not eliminating conflicts entirely). Consequently, fewer edges are formed in the corresponding conflict graph.

Consider the multicell topology in Fig. 6, where we apply the three-frequency-channel assignment (as in Fig. 3). In Fig. 6, the shaded cells use the same frequency channel, while the unshaded cells use the other two frequency channels. Although the size of the topology in Fig. 6 is comparable to the 5×5 , 25-cell WLAN described in previous sections, the three frequency channels help to reduce conflicts and increase the number of VoIP sessions that can be supported by each AP. NS2 simulations show that applying the clique-analytical admission control to the three-frequency-channel layout can boost the per-AP capacity to 7.39 VoIP sessions in 802.11b and 44.91 VoIP sessions in 802.11g. These are, respectively, 2.98 and 3.17 times of the per-AP capacity in

802.11b and 802.11g 5×5 , 25-cell WLAN where single-frequency channel is used.

In the next section, we explore time division on 802.11 MAC which can eliminate all hidden nodes (HN) and alleviate the performance-degrading exposed-node (EN) problem. The VoIP capacity over multiple WLANs can be further improved.

4 TIME-DIVISION CSMA MAC

The low VoIP capacity over multiple WLANs is due to mutual interferences among neighboring cells, and such mutual interferences cannot be completely isolated even with careful frequency channel assignment. The detrimental effects of such mutual interferences on Quality of Service (QoS) in the multicell WLAN have often been overlooked in previous work. For example, IEEE 802.11e has been standardized recently to support QoS in WLAN [18], but it only focuses on the single-cell situation and does not take mutual interferences of cells into account.

In this section, we explore adding the time-division approach to the basic 802.11 CSMA protocol. We show that the integrated time-division CSMA approach can potentially increase the VoIP capacity over multiple WLANs quite significantly. Selected previous work has considered Time-Division Multiple-Access (TDMA) MAC. However, their focus is on the single-WLAN case [19], [20], [21], [22]. In addition, CSMA is proposed to be replaced entirely by TDMA [23], [24], [25], i.e., the motivation was not to explore a solution workable within the context of the widely deployed 802.11 technology.

For VoIP applications within 802.11, each transmission consists of a very small packet (relative to the raw data rate and the various packet headers). “Fine” TDMA, in which each time slot corresponds to a packet transmission time, also requires tight synchronization among the stations. Both factors may cause throughput degradation. In this section, our primary focus is on the principle of “coarse” time division, in which a relatively large time slot is allocated to a group of stations. The stations of the same time slot then contend for channel access using the original 802.11 CSMA scheme. We 1) lay out and investigate a graph-theoretic problem formulation that captures the principle of integrating coarse time division with CSMA in Section 4.1 and 2) provide a feasibility investigation of the approach within the context of 802.11 in Section 4.2.

4.1 Coarse-Grained Time-Division Multiple Access

In Section 2.3, we have discussed three-frequency-channel assignment in 802.11b/g WLAN (see Fig. 3) and argued that frequency channel assignment alone cannot completely isolate co-channel interference. Section 3.4 applies the clique-analytical call admission control to boost per-AP VoIP capacity in three-frequency-channel WLAN; however, the co-channel interference from different cells is still not isolated entirely. The goal of Coarse-grained Time-Division Multiple Access (CoTDMA) is to remove such co-channel interference. In particular, we impose a restriction such that two stations in different cells that interfere with each other would be allocated different time slots or different frequency channels.

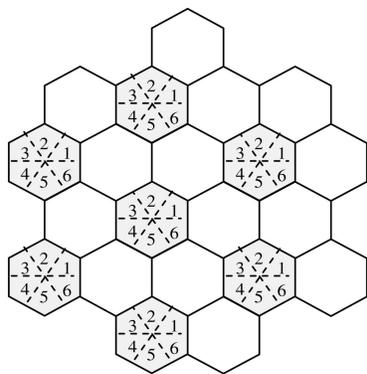


Fig. 7. Six-time-slot assignment in addition to three-frequency-channel assignment in multiple WLANs.

4.1.1 Basic Ideas of CoTDMA

We first explain the concept using the *simplified* scenario depicted in Fig. 7. In Fig. 7, in addition to the three-frequency-channel assignment, we divide each cell into six sectors and assign a distinct time slot to each sector. The shaded cells use the same frequency channel and the stations within each sector use the same time slot; the number in each sector indexes the time slot assigned to that particular sector. The frequency and time-slot assignments in Fig. 7 are such that different sectors do not interfere with each other because they are either active in different time slots, in cells of different frequencies, or sufficiently far apart from each other. Note that within each sector (time slot), there may still be multiple stations and the original 802.11 CSMA scheme is used to coordinate transmissions among these stations. In the example, we can shorten the $CSRange$ to d_{max} , the distance from the AP located at the center of each cell to the farthest corner in the cell. The nearest station with the same frequency and time-slot assignment in neighboring cells is at least $2d_{max}$ away, which is larger than both the $CSRange$ and IR_{max} defined in Section 3.2. Therefore, the co-channel interference from neighboring cells is completely isolated. In the best-case scenario, the *VoIP capacity per AP* in multiple WLANs can be the same as that in the single isolated WLAN case. To see this, consider 802.11b: if each sector has exactly two stations which are the maximum stations for each time slot (we will explain it later), then we can have a total of 12 VoIP sessions per AP. This example is merely for illustration purposes. In real implementation, no sectors would be divided in the cell and the time-slot assignment is not based on the location of the VoIP sessions but is based on a conflict graph with two-layer colors, as discussed below.

Although the sectorized cells in Fig. 7 illustrate the concept of CoTDMA, it has two implementation difficulties: 1) time-slot assignment requires the knowledge of the locations of the individual stations and 2) if the stations are not evenly distributed across the sectors, then it will not be as effective as the best-case scenario mentioned above. In the following, we present a graph model of CoTDMA to solve these problems. The graph model presented also gives a framework for performance analysis of CoTDMA.

For convenience, we will continue to describe the system operation in terms of distances such as $CSRange$ and IR . As stated in Section 3.1, implementation based strictly on the

“geometric” distance interpretation is unnecessary once we move on to the graph-theoretic formulation here. A mapping of distances to power thresholds is all that is needed. The construction of the conflict graph described below is compatible with the power threshold interpretation in real implementation.

Definition 1. In CoTDMA, m frequency channels and n time slots are assigned to the VoIP sessions. In each cell, at most k VoIP sessions are active in each time slot, where $k = \lfloor C_{AP-1}/n \rfloor$, and C_{AP-1} is the per-AP capacity in a single isolated WLAN.

4.1.2 Conflict Graph Modeling of CoTDMA

According to Definition 1, the parameter m is the number of frequency channels available. The purpose of CoTDMA is to completely isolate co-channel interference, so only orthogonal (nonoverlapping) frequency channels will be assigned to the multicell WLAN. For example, $m = 3$ in 802.11b/g and $m = 12$ in 802.11a. The variable n is the number of time slots we have. If $n = 6$, then $k = 2$ in 802.11b, and $k = 10$ in 802.11g, since the respective C_{AP-1} are 12 and 60.

We will look at the system performance as a function of n shortly. First, we formulate the corresponding graph-theoretic coloring problem. Coloring is a well-known problem in graph theory. However, the assignment problem in CoTDMA does not directly map to the classical coloring problem. For CoTDMA, a *modified* construction of the conflict graph as well as a *modified coloring problem* are needed to reflect the specifics of 802.11 CSMA scheme, as detailed below.

Instead of preassigning the three orthogonal frequency channels, as shown in Fig. 7, let us first set up a general framework which integrates the frequency channel assignment and time-slot assignment. In the CoTDMA conflict graph model, we use two layers of coloring. The first-layer colors represent frequency channels and the second-layer colors represent time slots. The first-layer coloring is applied at the cell level (assuming that all nodes within the cell use the same frequency channel in a static manner), while the second layer coloring is applied at the station (vertex) level. In the following, we first state the constraints of our coloring problem under the context of 802.11, and then describe how to capture the constraints in the conflict graph. Vertices are associated with the client stations in the following:

Constraint 1. The number of available first-layer colors is m and the number of available second-layer colors is n (see Definition 1).

Constraint 2. All vertices associated with the same AP (within the same cell) must have the same first-layer color.

Constraint 3. Within a cell, there can be at most k vertices assigned with the same second-layer color. Furthermore, the vertices which assigned the same second-layer color in a cell must be within the $CSRange$ of each other.

For vertices in the same cell, it is obvious that the client stations cannot transmit together since one end of the links is always the AP. The issue for vertices within the same cell is that whether the $CSRange$ of vertices (clients) can cover each other; hence, Constraint 3, i.e., if two vertices are not within each other’s $CSRange$, then carrier sensing

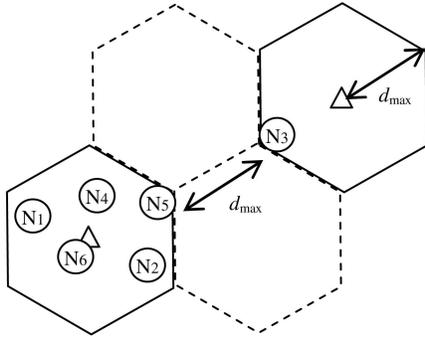


Fig. 8. A topology of VoIP over multiple WLANs with three-channel assignment.

between them does not work, and therefore, they should be assigned different time slots to avoid the hidden-node phenomenon.

Constraint 4. Consider two vertices of *different* cells, v_i and v_j . They conflict with each other and must be assigned different (first-layer color, second-layer color) combinations if one or more of the following inequalities below are satisfied:

$$CSRange \geq \min(d_{v_i, v_j}, d_{v_i, v'_j}, d_{v'_i, v_j}, d_{v'_i, v'_j}), \quad (5)$$

$$\begin{aligned} IR_{v_i} &> \min(d_{v_i, v_j}, d_{v_i, v'_j}), \\ IR_{v'_i} &> \min(d_{v'_i, v_j}, d_{v'_i, v'_j}), \\ IR_{v_j} &> \min(d_{v_i, v_j}, d_{v'_i, v_j}), \\ IR_{v'_j} &> \min(d_{v_i, v'_j}, d_{v'_i, v'_j}), \end{aligned} \quad (6)$$

where v'_i and v'_j are the corresponding APs that v_i and v_j associated, respectively. Note that both (5) and (6) describe the conditions under which simultaneous transmissions are not possible (see Section 3.2). However, there is a subtlety. Inequalities (6) capture the conditions that will lead to collisions. Inequality (5), on the other hand, only says that the CSMA mechanism will prevent the stations from transmitting together—that is, strictly speaking, if (5) is satisfied, the stations could still be assigned the same color combination, and the CSMA mechanism simply prevents simultaneous transmissions (if we did that, Constraint 3 would need to be modified to encompass the overall network). Constraint 4, however, disallows that as a design choice to simplify things. The reasons are as follows: 1) If different color combinations are used whenever (5) is satisfied for two vertices in different cells, then CSMA in different cells will then be decoupled in each of the time slots under CoTDMA, obviating the need for intercell CSMA. 2) When inequalities (5) and (6) are imposed on CoTDMA coloring, we may decrease $CSRange$ to only meet the need of intracell CSMA and there is no need for $CSRange$ to be large enough to ensure proper CSMA operation across cells to prevent HN (as in Section 2.3, where $CSRange$ has to span across cells). Short $CSRange$ has the advantage of reducing EN across cells [15]. EN hinders spatial reuse and arises when $CSRange$ is too large so that links that could potentially transmit together without collisions are prevented from doing so because their transmitters could sense each other (i.e., the false alarm

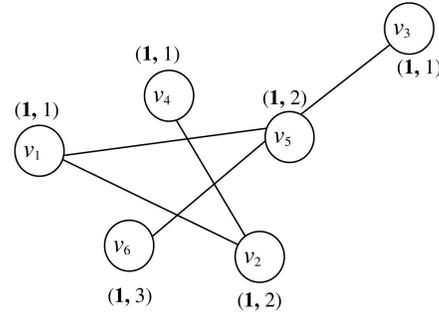


Fig. 9. A conflict graph with first-layer and second-layer colors for the network in Fig. 8.

scenario). Later in this section, we will explore the “optimal” value for $CSRange$ through simulations.

Capturing Constraints 2-4 in conflict graph. To capture Constraint 2, we could assign an AP_ID to the vertices in accordance with the APs to which they associate. Vertices with the same AP_ID must be given the same first-layer color.

For Constraints 3 and 4, an edge between two vertices means that they must be assigned different (first-layer color, second-layer color) combination.

To capture Constraint 3, there is an edge between two vertices v_i and v_j of the same cell if

$$d_{v_i, v_j} > CSRange. \quad (7)$$

Two vertices that are within the $CSRange$ of each other could be assigned the same or different second-layer color. However, there can be at most k vertices with the same second-layer color within a cell.

Constraint 4 can be captured by drawing an edge between two vertices of different AP_ID if there is a conflict relationship under inequalities (5) and (6). To avoid confusion, it is worth emphasizing again that between vertices of different AP_ID (from different cells), we use (5) and (6) for the edge formation criteria. For vertices of the same cell, we use (7) for the edge formation criteria.

A formulation of the CoTDMA problem is as follows:

Two-color assignment problem. Assign (first-layer color, second-layer color) to the vertices subject to Constraints 1-4 such that the total number of vertices successfully colored is maximized.

Fig. 8 illustrates the idea of CoTDMA. APs (triangles) are at the center of the cells, client stations (circles) in the same cell are associated with the same AP. The solid-line cells use the same frequency channel, while the dotted-line cells use the other two frequency channels. For simplicity, we assume the standard three-frequency-channel assignment here for the frequency channel assignment in CoTDMA. It is worth noting that this standard three-frequency-channel assignment is not a must for CoTDMA. In this case, we only draw a conflict graph for second-layer coloring. Due to the standard three-frequency-channel assignment, the vertices representing clients N_1 - N_6 all have the same first-layer color. Here, we assume 802.11b and $n = 6$ (which implies $k = 2$ given a capacity of 12 VoIP sessions per cell), and $CSRange = d_{max}$.

Fig. 9 shows the corresponding conflict graph together with the coloring. The bold numbers in parentheses beside

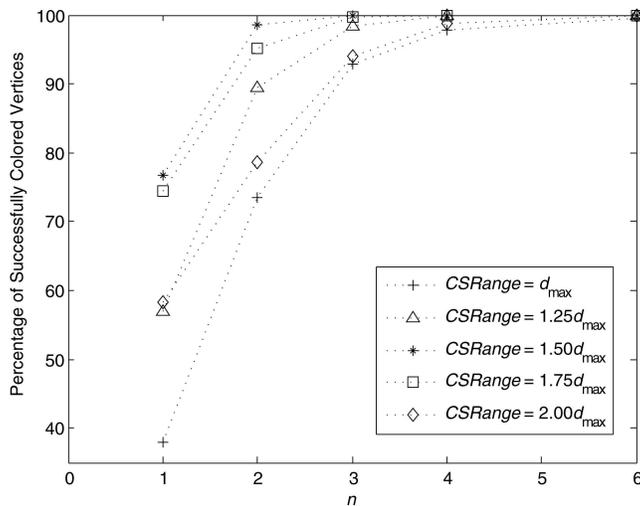


Fig. 10. Average percentage of colored vertices as $CSRange$ changes when $C_{AP,1} = 12$ (802.11b).

the vertices are the first-layer colors, while the other numbers are the second-layer colors. According to Constraint 3, for second-layer coloring, v_1 and v_4 are assigned COLOR1 (i.e., N_1 and N_4 are assigned time slot 1), v_2 and v_5 are assigned COLOR2, and v_6 is assigned COLOR3. v_3 and v_5 may interfere with each other under (5) and (6), so an edge is drawn between them. According to Constraint 4, the two vertices must be assigned with different color (first-layer color, second-layer color): in Fig. 9, COLOR1 is assigned to v_3 and COLOR2 is assigned to v_5 . Like the implementation of clique-analytical call admission control in Section 3, we also assume a centralized entity to construct the two-layer-colored conflict graph and to do (first-layer color, second-layer color) assignment by using some coloring algorithms (e.g., the Welsh-and-Powell algorithm [26]).

4.1.3 Parameter Values in CoTDMA

An important parameter in CoTDMA is n , the number of time slots (second-layer colors) available. Network designers are free to set different positive integral values to n . However, from Definition 1, k , the number of VoIP sessions that can be active in the same time slot in a cell, is set accordingly to n . In the above example, we set $n = 6$, so $k = 2$. The value of n directly affects the number of vertices that can be successfully colored. A larger n (i.e., a smaller k) means more finely divided time slots. In the extreme case, $k = 1$ ($n = C_{AP,1}$, i.e., $n = 12$ for 802.11b and $n = 60$ for 11g), which is a pure TDMA scheme. In this case, every VoIP session in a cell is assigned a distinct time slot. Hence, no carrier sensing is required for accessing the medium. The per-AP VoIP capacity is the same as that in the scenario of isolated single-cell WLAN. From the graph-theoretic coloring viewpoint, fine TDMA as such (i.e., small k , large n) would allow us to increase the number of vertices successfully colored in the two-layer-color assignment problem defined above. However, fine TDMA has an implementation cost not captured in the coloring problem—namely, there is the need for a “guard time” as we switch from slot to slot. This implementation issue will be further discussed in the next section. For the time being, it

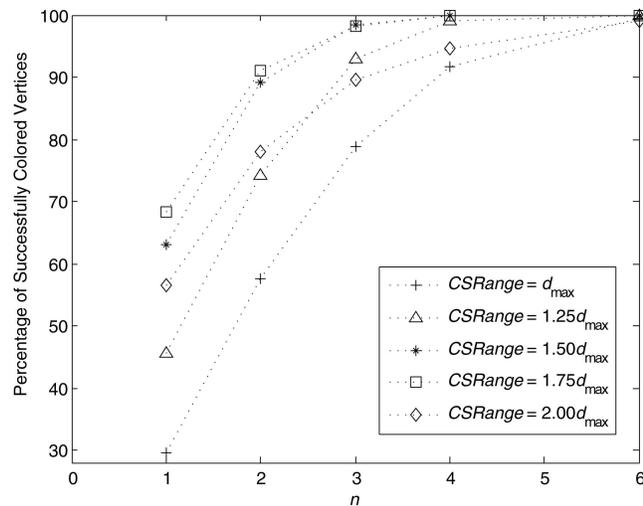


Fig. 11. Average percentage of colored vertices as $CSRange$ changes when $C_{AP,1} = 60$ (802.11g).

suffices to say that we are interested in making k as large as possible (i.e., making n as small as possible) while retaining the performance results of the case where k is set to 1.

Another important parameter in CoTDMA is $CSRange$. As mentioned in the explanation of Constraint 4 in Section 4.1.2, CoTDMA allows us to decrease $CSRange$ and the carrier-sensing mechanism needs only to work properly within a cell. Hence, it reduces EN and increases spatial reuse. However, if $CSRange$ is too small, a client may only carrier sense few other client stations within the cell. According to constraint 3 (and (7)), small $CSRange$ may cause more intracell edges in the conflict graph that restrict coloring freedom. It would therefore be of interest to explore tuning the $CSRange$ to maximize the number of vertices that can be successfully colored in the conflict graph.

We have performed MATLAB experiments assuming 802.11b ($C_{AP,1} = 12$) and 802.11g ($C_{AP,1} = 60$) to investigate the impact of n and $CSRange$ on the performance of CoTDMA. For simplicity, the standard three-frequency-channel assignment is assumed. Since the first-layer colors are prefixed in this experiment, only second-layer coloring (time-slot assignment) is considered. We use hexagonal cells to model WLAN, and 12 (for 802.11b) or 60 (for 802.11g) wireless client stations are randomly placed inside each cell with uniform distribution.

We use a heuristic algorithm of Welsh and Powell [26] to color the conflict graph. We add our coloring constraints 1-4 to tailor the algorithm to CoTDMA. The algorithm of Welsh and Powell does not give optimal graph coloring in general, but the effect of n and $CSRange$ is already quite pronounced even with the simple heuristic.

In the experiments, for each run, we set a fixed $CSRange$ for all values of n . For each n , we ran five experiments with different node distributions. We recorded the average percentage of successfully colored vertices in Figs. 10 and 11. In general, as n increases, more vertices can be successfully colored. Although not shown in the figures, the two cases for $n = 12$ in 802.11b and $n = 60$ in 802.11g have 100 percent of their vertices successfully colored in all runs of our experiments.

TABLE 5
System Performance for $CSRange = 1.637d_{max}$

n	1	2	3	4	6	$C_{AP,1}$
% vertices colored in 802.11b	80.0	98.6	100	100	100	100
% vertices colored in 802.11g	69.8	93.0	99.6	100	100	100

Across different runs, we also vary $CSRange$. We find that when $CSRange$ is d_{max} , the overall system performance is poor because many edges are formed within a cell (according to (7)). As $CSRange$ increases, the overall system performance improves. But beyond certain point (e.g., $1.5d_{max}$ in the figures), the overall system performance drops again. This phenomenon reveals the trade-off between intracell optimality and intercell optimality. When $CSRange$ is too large, say $2d_{max}$, many edges are formed between vertices of different cells (according to (5)), leading to an increase of EN. Through experimentation, we find that the “optimal” $CSRange$ which yields the “best” system performance is around $1.637d_{max}$. Indeed, we could use this setting for different n values with reasonably good results (see Table 5). With smaller n , the overhead of guard time for switching between time slots can be reduced (to be elaborated shortly), so from Table 5, $n = 3$ or 4 may offer the best design trade-off.

Our experiment results show that even though the freedom of the two dimensions of frequency and time in CoTDMA has not been fully utilized (due to our assumption of the fixed three-frequency-channel assignment), CoTDMA can generally color large portions (over 90 percent) of the vertices of the conflict graph when $n \geq 4$ for both 802.11b and 802.11g (see Figs. 10 and 11). The performance can be even better when an appropriate (“optimal”) $CSRange$ is set (see Table 5). By contrast, without CoTDMA and with three-frequency-channel assignment alone (i.e., $n = 1$), we find that only around 60 percent of the total capacity can be utilized.

4.2 Possible Realizations of Time-Division Multiple Access within Existing IEEE 802.11 Standards

Most previous work [20], [21], [22], [23], [24], [25] considered proprietary protocols for implementing TDMA on wireless networks. Our focus here is to implement time-slot assignment within the framework of the IEEE 802.11 standard. The most critical issue is how to realize the concept of “time slot” within the 802.11 CSMA structure. A possibility is to make use of the “sleep mode” in 802.11, which was originally designed for power conservation purposes. In the sleep mode, beacon frames are used for synchronization. Accordingly, in CoTDMA, all stations could wake up around the beacon time for synchronization. As illustrated in Fig. 12, in CoTDMA, within each beacon interval (BI) between the *end* of a beacon and the *beginning* of the next beacon, the time is divided into C frames (cycles), each of duration ΔT . Within each frame, there are n time slots, each of duration Δt . The offset from the end of the beacon to the beginning of the i th frame is $(i - 1)\Delta T$. A

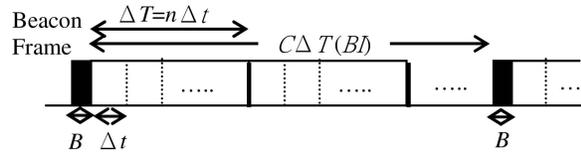


Fig. 12. Frame structure of CoTDMA when it is implemented using 802.11 sleeping mode.

station that has been assigned time slot j is to be awake within a BI only during the time intervals $[(i - 1)\Delta T + (j - 1)\Delta t, (i - 1)\Delta T + j\Delta t]$, $i = 1, 2, \dots, C$. Other than these time intervals and the beacon time, the station sleeps.

Guard-time overhead in CoTDMA. The guard time should be set to the duration of one VoIP packet. That is, no packet transmission should be initiated within the current time slot when the beginning of the next time slot is only a guard time away. This is to ensure that packet transmission will not straddle across two time slots. Let $r_{\Delta t}$ be the *maximum* number of VoIP packets (including the CSMA overhead) that could be transmitted and received within each time slot by all VoIP sessions which are active in that particular time slot. Then,

$$r_{\Delta t} = 2 \times \frac{R_{VoIP}}{10nC} \times C_{AP,1}, \quad (8)$$

where R_{VoIP} is the number of VoIP packets generated per second in a particular VoIP codec. The factor of 2 is due to each VoIP session having a downstream and an upstream flow. The factor of $1/10$ is due to the default 0.1 s separation time between two beacons. The guard-time overhead is a constant of one VoIP packet duration so that the time-slot efficiency is $(r_{\Delta t} - 1)/r_{\Delta t}$. Thus, smaller $r_{\Delta t}$ gives rise to lower efficiency, which, in turn, results in lower capacity. From (8), we can see that it is desirable to set n and C as small as possible. From the simulation experiments in the previous section, however, we need to make sure that $n \geq 3$ (see Table 5) so as to make sure most sessions within the system capacity limit can be admitted. While call admission consideration imposes a limit on how small n can be, the delay budget consideration imposes a limit on how small C can be, as explained in the next few paragraphs. In other words, the factors that bound on the size of n and C are different.

In CoTDMA, we allow $\Delta T + B$ as the maximum delay for a VoIP packet, where B is the duration of the beacon (see Fig. 12). To see this maximum delay, let us consider the station being assigned time slot j . In the worst case, it could generate a packet in the last frame within a BI just slightly after time slot j ends, thus missing it. The earliest time for the next time slot j (in the next BI) is $\Delta T + B - \Delta t$ after that. Within this next time slot j , in the worst case, the packet is sent just before the end of the time slot (due to the CSMA contention with other stations assigned the same time slot). We assume that the packet will be discarded if it fails to be sent out by this time so as to make way for a newly generated packet from the same VoIP session. The number of clients operating in the same time slot is limited by the parameter k according to Definition 1 of CoTDMA. The value of k is set so that the maximum VoIP capacity of a

WLAN will not be exceeded (i.e., the network is unsaturated). So, in an unsaturated network, it is unlikely that a particular station cannot “win” the medium access in many time slots. According to [27], if we assume 802.11b, the probability of packet collision in a saturated network is less than 0.15 if there are four clients in a WLAN (in our case, we have four clients in each time slot if $n = 3$). The collision probability is even lower for our unsaturated network. Hence, the portion of time slot wasted resulted from the fact that packet collision is also small. So, the maximum delay is then $\Delta T + B$. Thus,

$$\Delta T + B \leq DB, \quad (9)$$

where DB is the delay budget. Since $C\Delta T = BI$ (see Fig. 12), we have

$$BI/C + B \leq DB \Rightarrow C \geq \frac{BI}{DB - B}. \quad (10)$$

Suppose we set a local delay budget of 30 ms for VoIP applications [1]. A typical value of B is 0.5 ms. With the default separation time between two beacons of 100 ms, a BI is 99.5 ms. Hence, the number of frames C in a beacon interval is at least $99.5/29.5 = 3.37$. C should be a positive integer, so the smallest C is 4.

Assuming that $n = 3$, $C = 4$, and the use of GSM 6.10 codec and 802.11b ($R_{VoIP} = 50$, $C_{AP-1} = 12$), from (8), we find that 90 percent of capacity is utilized. That is, at most $\lfloor 12 \times 90 \text{ percent} \rfloor = 10$ VoIP sessions can be admitted per AP in 802.11b networks. If 802.11g is assumed ($C_{AP-1} = 60$), 98 percent of capacity is utilized. That is, 58 VoIP sessions can be admitted per AP in 802.11g networks.

Take 802.11b WLAN. With three-frequency-channel CoTDMA, the per-AP capacity over multiple WLANs is 10 VoIP sessions. This is a 35.3 percent improvement over the per-AP capacity of 7.39 sessions for the three-frequency-channel clique-analytical call admission control strategy in Section 3.4.

Another possibility for implementing the concept of time division is to use the “polling mechanism” of Point Coordination Function (PCF) [12] to imitate the time slot assignment in CoTDMA. In PCF, traffic is scheduled by the AP so that no extra guard-time overhead is needed for time-slot switching. If we assume PCF, we could set $n = C_{AP-1}$ (i.e., $k = 1$) in our CoTDMA call admission scheme so that we essentially have the extreme case of Fine-grained Time-Division Multiple Access (FiTDMA). In PCF, an AP maintains a polling list containing all the wireless stations in its WLAN. In the contention-free period, the AP polls the stations on the polling list one by one since only one VoIP session is active in each time slot in a cell in FiTDMA. Only the polled station can transmit packets to AP (the downstream packet is attached in the polling packet). The position of a station in the polling list corresponds to the time slot assigned to the station. That is, there is still the issue of selecting the time slot to poll the clients. Through our two-layer-color assignment approach, time slots can be selected in a way that isolates co-channel interference among different WLAN cells. In this way, stations which may interfere with each other in adjacent cells will not be polled at the same time. To do

this, we can fix $n = C_{AP-1}$ in our two-layer-color assignment, and then perform (first-layer color, second-layer color) assignment. With $n = C_{AP-1}$, we know from the experiments in Section 4.1.3 that 100 percent of clients can be successfully colored. The system performance can be increased to 12 sessions per AP in 802.11b and 60 sessions per AP in 802.11g.

With FiTDMA, the number of vertices successfully colored can be increased. Furthermore, since only one VoIP session is active in each time slot in a cell, carrier-sensing mechanism can be deactivated. Without backoff countdown in contention period, C_{AP-1} , the per-AP capacity in single isolated WLAN, can be boosted in PCF (in 802.11b, C_{AP-1} increases from 12 to 17, while in 802.11g, C_{AP-1} increases from 60 to over 90). From this point of view, FiTDMA in PCF may increase VoIP capacity more than CoTDMA in DCF does. A major concern, however, is that PCF is seldom used in practice and many 802.11 devices do not support it—unlike DCF, the robustness of PFC in field deployment has not been well tested. Another concern is that FiTDMA is essentially a pure TDMA scheme, so it has the disadvantages common to TDMA, such as tight synchronization requirement, wastage of the bandwidth resource if the traffic of nodes is not saturated, and so on.

5 CONCLUSION

In this paper, we have shown that when there are multiple 802.11 WLANs within the vicinity of each other, the already low VoIP capacity in the single isolated WLAN case (around 12 and 60 VoIP sessions per AP in 802.11b and 802.11g, respectively) is further eroded in a very significant manner. It is possible that only less than one percent goodput can be supported. For example, in 802.11b, less than two VoIP sessions per AP can be supported, while computation based on the raw bandwidth of the WLAN yields 200 sessions per AP. In 802.11g, around 10 VoIP sessions per AP can be supported, while 900 sessions are suggested by raw-bandwidth computation. The dismal performance and inefficiency imply that there is much room for improvement within the 802.11 standard as far as the support for VoIP is concerned.

The low VoIP capacity in the single WLAN [1], [2] is due largely to header overheads, and packet aggregation [1], [2] is an effective solution to reduce the header penalty. The further degradation of VoIP capacity in the multiple-WLAN case, however, is due to mutual interferences among the WLANs and requires additional solutions. In essence, CSMA is rather inefficient when there are multiple WLANs in the vicinity of each other. This paper suggests a two-pronged approach: **1) call admission control** and **2) virtual channelization**.

Regarding 1), we have formulated a clique-analytical call admission control algorithm and shown that (compared with a simplistic call admission scheme) it can improve the VoIP capacity in a 5×5 , 25-cell 802.11b WLAN by 52.1 percent from 1.63 sessions to 2.48 sessions per AP. The improvement is 36.8 percent in 11g. If three orthogonal frequency channels are used, such as those available in 802.11b/g, the capacity can be increased to

7.39 (802.11b) and 44.91 (802.11g) VoIP sessions per AP by careful frequency channel assignment to the cells.

Regarding 2), the three orthogonal frequency channels in 802.11b/g are not enough to completely isolate interferences among WLANs. This, in turn, requires the carrier-sensing range of 802.11 to be set rather large to prevent packet collisions; but doing so also increase the exposed-node problem that degrades the VoIP capacity. In this paper, we have shown that “virtual channels” (or time slot channels) could be created to combat this problem. By assigning the virtual channels judiciously to the VoIP stations, we could effectively isolate the interferences between cells. Specifically, we have proposed a scheme called CoTDMA for virtual channelization that is compatible with the basic 802.11 CSMA protocol. In CoTDMA, the time dimension is divided into multiple coarse time slots; multiple VoIP sessions are assigned to each time slot, and the VoIP sessions assigned to the same time slot in the same WLAN make use of the basic 802.11 CSMA protocol to coordinate channel access. The basic idea is that VoIP sessions of adjacent WLANs that may interfere with each other in a detrimental way should be assigned different time slots or frequency channels. In essence, the VoIP sessions of different WLANs do not need to use CSMA to coordinate transmissions among them. This means that CSMA is in use only within a cell, thus bypassing the inefficiency of CSMA in the multicell scenario.

From our simulation experiments, we find that CoTDMA could improve the VoIP capacity meaningfully. Our results indicate that with a small number (3-4) of coarse time slots in CoTDMA, the per-AP VoIP capacity can be increased to 10 sessions in 802.11b and 58 sessions in 802.11g (another 35.3 and 29.15 percent improvement over clique-analytical admission control with three orthogonal frequency channels, respectively).

We note that CoTDMA is fundamentally a technique in which stations contending for a common resource (i.e., airtime) are compartmentalized into subsets so that only stations within each subset contend with each other. The partitioning is done in such a way that the subnetwork consisting of stations within each subset is less susceptible to detrimental interference/carrier-sensing pattern so that the resource could be used more efficiently. This principle is applicable not just to voice traffic, but to wireless networking, in general, with or without voice traffic. The graph-theoretic formulation of the two-layer coloring problem in Section 4 may serve as the starting point for the exploration of the general case.

APPENDIX

The clique-analytical call admission control presented in Section 3 can be generalized if we calculate the sum of airtime of sessions instead of counting the number of sessions in the maximal cliques. Here, we also assume that only VoIP traffic is present in the network. The model can be directly extended to the scenario of coexistence with other traffic, like best-effort data traffic, by preallocating some airtime to the other traffic.

In *each* second, the transmission time needed (airtime AT_{v_i}) for VoIP session v_i can be computed as (A1):

$$AT_{v_i} = 2T_{avg-v_i} R_{VoIP-v_i}. \quad (A1)$$

In the above, R_{VoIP-v_i} is the number of VoIP packets that v_i generates per second, which depends on the VoIP codec used by v_i . T_{avg-v_i} is the average transmission time of a VoIP packet in v_i . It is calculated from the VoIP packet payload size, 802.11 data rate, and various header overheads [1]. There are both uplink and downlink in a VoIP session, so there is a factor of 2 in (A1).

If the VoIP sessions are in the same clique in the corresponding conflict graph, their airtimes cannot overlap. Therefore, (A2) sums up all the airtime AT_{v_i} , $\forall v_i \in C_{v_i,x}$, where $C_{v_i,x}$ is one of the maximal clique which v_i belongs to (refer to Section 3.3):

$$AT_x = \sum_{v_i \in C_{v_i,x}} AT_{v_i}, \quad (A2)$$

$$1 \leq m_{AT_{v_i}} = \max_x(AT_x). \quad (A3)$$

Obviously, AT_x cannot be larger than 1 second. So, (A3) sets the condition for call admission control. Instead of counting the maximum number of VoIP session, as in Section 3.3, where $m_{v_i} = \max_x |C_{v_i,x}|$, we now generalize the call admission control to find the maximum total airtime used, $m_{AT_{v_i}} = \max_x(AT_x)$, in a clique.

In the above derivation, we have taken different VoIP codecs (different payload sizes, packetization intervals, etc.) and different WLAN data rates (could be caused by ARF) into account.

The total airtimes calculated above did not include the overhead time due to transmission failures. To take into account transmission failures due to collisions, the total available airtime can be set to $1 - p_c$, where p_c is the collision probability in saturated or unsaturated networks, the derivations of which can be found in previous work [27], [28]. Transmission failures can also be due to background random noise. To take them into account, p_c should be replaced by p , where p is the loss probability due to random noise and collisions. The loss probability due to random noise depends heavily on the environment the network operates in [29]. The experimental results in [29] showed that in a typical office environment, there is no visible packet loss when both the client station and AP are in the same room. It is worth noting that after applying CoTDMA, all hidden nodes are removed and collision probability is small (see discussion in Section 4.2). Hence, the overhead due to transmission failures is largely reduced. To the extent that random noise is small and negligible, we could still use *one* second as the total available airtime and the actual airtime used will be very close to what we calculate in (A2).

ACKNOWLEDGMENTS

This work was supported by the Competitive Earmarked Research Grant (Project Number 414106) established under the University Grant Committee of the Hong Kong Special Administrative Region, China.

REFERENCES

- [1] W. Wang, S.C. Liew, and V.O.K. Li, "Solutions to Performance Problems in VoIP over 802.11 Wireless LAN," *IEEE Trans. Vehicular Technology*, vol. 54, no. 1, pp. 366-384, Jan. 2005.
- [2] W. Wang, S.C. Liew, Q.X. Pang, and V.O.K. Li, "A Multiplex-Multicast Scheme That Improves System Capacity of Voice-over-IP on Wireless LAN by 100%," *Proc. Ninth IEEE Symp. Computers Comm.*, June 2004.
- [3] D.P. Hole and F.A. Tobagi, "Capacity of an IEEE 802.11b Wireless LAN Supporting VoIP," *Proc. IEEE Int'l Conf. Comm. (ICC '04)*, vol. 1, pp. 196-201, June 2004.
- [4] F. Anjum et al., "Voice Performance in WLAN Networks—An Experimental Study," *Proc. IEEE Global Comm. Conf. (GLOBECOM '03)*, vol. 6, pp. 3504-3508, Dec. 2003.
- [5] S. Garg and M. Kappes, "An Experimental Study of Throughput for UDP and VoIP Traffic in IEEE 802.11b Networks," *Proc. IEEE Wireless Comm. and Networking Conf. (WCNC '03)*, vol. 3, pp. 1748-1753, Mar. 2003.
- [6] H. Wu et al., "SOFTMAC: Layer 2.5 Collaborative MAC for Multimedia Support in Multihop Wireless Networks," *IEEE Trans. Mobile Computing*, vol. 6, no. 1, pp. 12-25, Jan. 2007.
- [7] D. Niculescu et al., "Performance of VoIP in a 802.11 Wireless Mesh Network," *Proc. IEEE INFOCOM*, Apr. 2006.
- [8] S. Ganguly et al., "Performance Optimizations for Deploying VoIP Services in Mesh Networks," *IEEE J. Selected Areas in Comm.*, vol. 24, no. 11, pp. 2147-2158, Nov. 2006.
- [9] *The Network Simulator—ns2*, <http://www.isi.edu/nsnam/ns/>, 2007.
- [10] Cisco System, Data Considerations and Evolution of Transmission Network Design, http://www.cisco.com/en/US/prod/collateral/optical/ps5724/ps2006/prod_white_paper0900aecd803faf8f_ps2001_Products_White_Paper.html, 2009.
- [11] A. Sfairopoulou, C. Macián, and B. Bellalta, "VoIP Codec Adaptation Algorithm in Multirate 802.11 WLANs: Distributed vs. Centralized Performance Comparison," *Dependable Adaptable Networks Services*, pp. 52-61, Springer, 2007.
- [12] M.S. Gast, *802.11 Wireless Networks: Definitive Guide*. O'Reilly, 2002.
- [13] L.B. Jiang and S.C. Liew, "Hidden-Node Removal and Its Application in Cellular WiFi Networks," *IEEE Trans. Vehicular Technology*, vol. 56, no. 5, pp. 2641-2654, Sept. 2007.
- [14] Cisco System, Extending the Reach of a LAN, <http://www.ciscopress.com/articles/article.asp?p=1156068&seqNum=2>, 2009.
- [15] L.B. Jiang and S.C. Liew, "Improving Throughput and Fairness by Reducing Exposed and Hidden Nodes in 802.11 Networks," *IEEE Trans. Mobile Computing*, vol. 7, no. 1, pp. 34-49, Jan. 2008.
- [16] *Clique*, http://en.wikipedia.org/wiki/Clique_%28graph_theory%29, 2009.
- [17] *Maximal Clique*, http://en.wikipedia.org/wiki/Maximal_clique, 2009.
- [18] *IEEE Std. 802.11e, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. Amendment 8: MAC Quality of Service Enhancements*, IEEE, Nov. 2005.
- [19] J. Snow, W. Feng, and W. Feng, "Implementing a Low Power TDMA Protocol over 802.11," *Proc. IEEE Wireless Comm. and Networking Conf. (WCNC '05)*, vol. 1, pp. 75-80, Mar. 2005.
- [20] S. Sharma et al., "Implementation Experiences of Bandwidth Guarantee on a Wireless LAN," *Proc. ACM/SPIE Multimedia Computing and Networking*, Jan. 2002.
- [21] T. Chiueh and C. Venkatramani, "Design, Implementation, and Evaluation of a Software-Based Real-Time Ethernet Protocol," *ACM SIGCOMM Computer Comm. Rev.*, vol. 25, no. 4, pp. 27-37, 1995.
- [22] K.M. Sivalingam et al., "Design and Analysis of Low-Power Access Protocols for Wireless and Mobile ATM Networks," *Wireless Networks*, vol. 6, no. 1, pp. 73-87, Feb. 2000.
- [23] F.N. Ali et al., "Distributed and Adaptive TDMA Algorithms for Multiple-Hop Mobile Networks," *Proc. IEEE Military Comm. Conf. (MILCOM '02)*, pp. 546-551, Oct. 2002.
- [24] A. Kanzaki, T. Hara, and S. Nishio, "An Adaptive TDMA Slot Assignment Protocol in Ad Hoc Sensor Networks," *Proc. ACM Symp. Applied Computing (SAC '05)*, Mar. 2005.
- [25] Z. Cai and M. Lu, "SNGR: A New Medium Access Control for Multi-Channel Ad Hoc Networks," *Proc. IEEE Vehicular Technology Conf. (VTC '00)*, May 2000.
- [26] D.J.A. Welsh and M.B. Powell, "An Upper Bound for the Chromatic Number of a Graph and Its Application to Timetabling Problems," *Computer J.*, vol. 10, pp. 85-86, 1967.
- [27] G. Bianchi, "Performance Analysis of the IEEE 802.11 Distributed Coordination Function," *IEEE J. Selected Areas in Comm.*, vol. 18, no. 3, pp. 535-547, Mar. 2000.
- [28] O. Tickoo and B. Sikdar, "Modeling Queuing and Channel Access Delay in Unsaturated IEEE 802.11 Random Access MAC Based Wireless Networks," *IEEE/ACM Trans. Networking*, vol. 16, no. 4, pp. 878-891, Aug. 2008.
- [29] C. Hoene et al., "Measuring the Impact of Slow User Motion on Packet Loss and Delay over IEEE 802.11b Wireless Links," *Proc. IEEE Conf. Local Computer Networks (LCN '03)*, 2003.



protocols. He is a student member of the IEEE.



Recently, he and his students won the Best Paper Awards in the First IEEE International Conference on Mobile Ad-Hoc and Sensor Systems (IEEE MASS 2004) and the Fourth IEEE International Workshop on Wireless Local Network (IEEE WLN 2004). Separately, TCP Veno, a version of TCP to improve its performance over wireless networks proposed by him and his students, has been incorporated into a recent release of Linux OS. His publications can be found at www.ie.cuhk.edu.hk/soung. He is a senior member of the IEEE.

► For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.