Offered Load Control in IEEE802.11 Multi-hop Ad-hoc Networks*

Ping Chung Ng and Soung Chang Liew
Department of Information Engineering
The Chinese University of Hong Kong
{pcng3, soung}@ie.cuhk.edu.hk

Abstract

In multi-hop ad-hoc networks, stations may pump more traffic into the networks than can be supported, resulting in high packet-loss rate, re-routing instability and unfairness problems. This paper shows that controlling the offered load at the sources can eliminate these problems. In addition, we provide an analysis to estimate the optimal offered load that maximizes the throughput of a multi-hop traffic flow. We use this result to devise schemes that can achieve fairness when there are multiple flows from different sources to different destinations. We believe this is a first paper in the literature to provide a quantitative analysis (as opposed to simulation) for the impact of hidden nodes, exposed nodes, and signal capture on sustainable throughput. The analysis is based on the observation that a large-scale 802.11 network with hidden nodes is a network in which the carrier-sensing capability breaks down partially. Its performance is therefore somewhere between a carrier-sensing network and an Aloha network. Indeed, our analytical closed-form solution has the appearance of the throughput equation of the Aloha network. Our approach allows one to identify whether the performance of an 802.11 network is hidden-node limited or spatial-reuse limited.

1. Introduction

A wireless multi-hop ad-hoc network provides quick and easy networking in circumstances that require temporary network services or when cabling is difficult. The IEEE 802.11 Distributed Co-ordination Function (DCF), based on Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA), is the most popular MAC protocol used in wireless ad-hoc networks.

In wireless networks, interferences are location-dependent. For a traffic flow from a source node to a destination node in a multi-hop network, the nodes in the middle of the path have to contend with more nodes when forwarding the traffic of the flow. Experiencing lighter contention, the source node may inject more traffic into the path than can be forwarded by the later nodes. This may result in excessive packet losses and re-routing instability. When there are multiple flows, unfairness may also arise when some flows experience higher contention than other flows.

The capacity of wireless networks has been studied extensively. Much of the previous work focused on computing theoretical throughput bounds (e.g. [1][2]). Some of these throughput limits are obtained under the assumption of global scheduling [3][4]. The popular IEEE 802.11 wireless networks in use today are not amenable to such global scheduling.

This paper primarily focuses on 802.11 and 802.11-like networks. Although there were also prior investigations [5][6] on how to modify the the 802.11 protocol to solve performance problems, we try not to perturb the protocol too drastically so that the same standard-based equipment can be used without major redesigns.

To devise schemes to achieve high throughput and fairness in multi-hop networks, it is important to be able to analyze the *contention experienced by a node* as a function of the network topology and traffic flows in a quantitative manner. Such an analysis is currently lacking in the literature, possibly due to the fact that the analysis is complicated by the existence of hidden-node, exposed-node and signal-capturing effects. This paper is a first attempt toward such a quantitative analysis. The analysis yields insight into the impact of different network parameters and properties on performance. As an example, we use our analysis to establish the optimal offered load for a traffic flow in this paper. We also show that the analytical approach can be used to achieve fairness when there are multiple flows in the network.

* This work was sponsored by the Areas of Excellence scheme established under the University Grant Committee of the Hong Kong Special Administrative Region, China (Project Number AoE/E-01/99).

Most previous studies of the hidden-node problem of 802.11 were conducted by simulations [2][7]. References [8] [9] extended the hearing graph framework in [10] to model hidden terminals and terminal mobility using a Markov chain. They established a relationship between the average number of stations hidden from each other and the likelihood of a station remaining in its Basic Service Area. Their results on the effect of hidden nodes on throughput, however, were obtained from simulations, not analysis. In addition, the signal capture property that allows a packet to be received successfully despite transmissions by hidden nodes was ignored.

The rest of this paper is organized as follows. Section 2 gives details of the simulation set-up assumed in this paper. In Section 3, we review the major performance problems in multi-hop ad-hoc networks and suggest possible solutions to them. Section 4 analyzes factors which degrade the throughput, and formulate a method to estimate the optimal offered load in a single-flow case. In Section 5, we show that our proposed scheme can achieve fairness of channel bandwidth usage among multiple flows. Section 6 concludes this paper.

2. Simulation Set-up

The simulations in this paper were conducted using NS2.1b9 [11]. All nodes communicate using identical, half-duplex wireless radio based on the 802.11 DCF, with data and basic rates set at 11Mbps. The RTS/CTS mechanism is turned off. Nodes are stationary. The transmission range is 250m, the carrier-sensing range is 550m, and the capture threshold *CPThreshold* is set to 10dB. The Ad-hoc On-Demand Distance Vector (AODV) routing protocol and the two-ray propagation model are used. All data sources are UDP traffic streams with fixed packet size of 1460bytes.

3. Performance Problems in 802.11 Multi-hop Networks: Single-Flow Investigation

In a multi-hop ad-hoc network, sources may inject more traffic into the network than can be supported. This may result in two problems: 1) high packet loss rate, and 2) re-routing instability. In this section, we use an 8-node string multi-hop network as an example to illustrate these problems. In Fig. 1, node 1 sends a UDP traffic stream to node 8. The traffic is generated at node 1 in a saturated manner in which as soon as a packet is transmitted to node 2, another is waiting in line. The traffic at later nodes all originates from node 1 and is not saturated.

3.1 High Packet-Drop Rate

Figure 2 shows the per-hop throughput of an 8-node flow obtained from simulations. The throughputs plotted are obtained by averaging over one-second intervals.

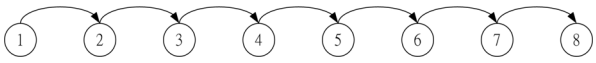

Figure 1. UDP traffic flow with node 1 as the source and node 8 as the destination in an 8-node multi-hop traffic flow

In Fig. 1, node 1 can sense the transmissions from nodes 2 and 3. This means node 1 must share the channel capacity with them. As a result, the throughput of the first hop is approximately 1/3 of the total channel capacity. Node 2, on the other hand, can be interfered by nodes 1, 3 and 4. This results in approximately 1/4 of the total channel capacity for the second hop. After that, each node must compete with four other nodes. The per-hop throughput stabilizes from the third hop to the last hop with approximately 1/5 of the total channel capacity. The first and the second nodes pump more packets to the following nodes than they can forward. This results in excessive packet drops at the second and the third node.

As shown in Fig. 2, the average throughput drops from 1.86Mbps at the first hop to 1.13Mbps at the last hop. In other words, about 40% of packets are lost in transit. This high packet-loss rate is undesirable, especially for real-time traffic without a retransmission mechanism at the upper protocol layer.

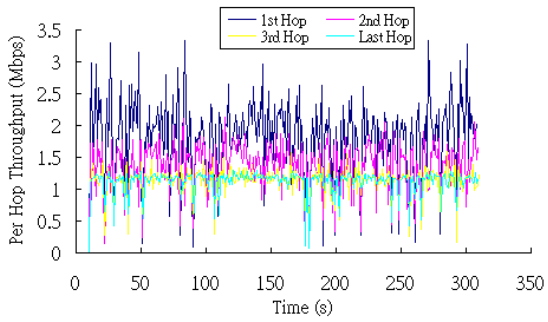

Figure 2. Per-hop throughputs of an 8-node flow

3.2 Re-routing Instability

Figure 2 also shows that the throughputs tend to oscillate widely over time. The throughput oscillations are caused by triggering of the re-routing function. In the multi-hop path, nodes 1 and 2 sense fewer interfering stations than later nodes. As a result, they pump more traffic into the network than it can support. This results in a high contention rate at the later nodes. When one of the

later nodes fails to transmit a packet after a number of retries, it declares the link as being broken. The routing agent is then invoked to look for a new route. Before a new route is discovered, no packet can be transmitted, causing the throughput to drop drastically. In the string network topology under study, there is only one route from node 1 to node 8, so the routing agent will eventually "re-discover" the same route again. The breaking and rediscovery of the path results in the drastic throughput oscillations observed. For a general network with multiple paths from source to destination, the same throughput oscillations will still be expected. This is because the declaration of the link failure is caused by self-interference of traffic of the same flow at adjacent nodes. More details on re-routing instability can be found in [12].

3.2.1 Hidden-Terminal Problem

Besides the collisions of packets among nodes inside a carrier sensing range, the hidden-terminal problem further increases the chance of link-failure declarations. Consider Fig. 3. When node 4 sends a packet to node 5, node 2 senses the channel to be busy while node 1 senses the channel to be idle, since node 4 is inside the carrier-sensing range of node 2 but outside that of node 1. Once node 1 senses the channel as idle, it may count down its back-off contention window until zero and transmit a packet to node 2.

If the transmission from node 4 is still in progress, node 2 will continue to sense the channel as busy, and it will not receive the packet from node 1. As a result, node 2 will not return an ACK to node 1. Node 1 may then time out and double the contention window size for retransmission later.

Meanwhile, node 4 transmits the packet successfully and is not aware of the collision at node 2. When transmitting the next packet, node 4 will use the minimum contention window size. The hidden-terminal scenario favors node 4, and the chance of collision at node 2 can not be reduced even though node 1 backs off before the next retry. The hidden-terminal problem increases the chance of multiple retries by node 1, making the wrong declaration of link failures and therefore re-routing instability more likely.

Note that the negative effect of a hidden terminal is much more than that of a contending terminal within the carrier-sensing range. This is because the carrier-sensing capability in the CSMA protocol breaks down with respect to the hidden terminal, making collisions much more likely.

3.2.2 Ineffectiveness of Solving Hidden-Terminal Problem with RTS/CTS

The RTS/CTS mechanism in 802.11 is designed to solve the hidden terminal problem. However, using RTS/CTS in multi-hop networks does not eliminate the hidden terminal problem. The effectiveness of RTS/CTS mechanism is based on the assumption that transmissions by mutually hidden terminals are to a common receiver. Before the transmission of a hidden terminal begins, the receiver will forewarn other hidden terminals to prevent them from transmitting. This assumption may not hold in a multi-hop network.

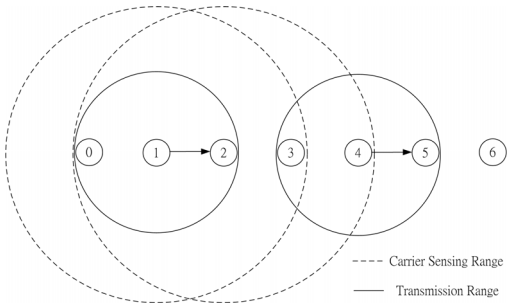

Figure 3. Node 4 as a hidden terminal to node 1

Consider the scenario in Fig. 3 again. The RTS transmitted by node 4 will cause a CTS to be returned by node 5. However, this CTS cannot be received by node 1. Therefore, node 1 may still transmit a packet to node 2 while the transmission of node 4 is in progress. The hidden-terminal effect as described in the previous subsection cannot be eliminated. For more details, the interested reader is referred to [5], in which it was argued that when the carrier-sensing range is larger than two times of the transmission range, RTS/CTS is no longer needed. In this paper, we assume the use of the basic access mode without RTS/CTS.

3.3 Solutions to High-Packet Loss Rate and Re-routing Instability

Reference [13] demonstrated the existence of an instability problem for a TCP traffic flow in a multi-hop network. It provided a solution to solve TCP instability by limiting the traffic at the transport layer. The solution assumes TCP Vegas and limits the TCP window size to at most 4. As a result, only a maximum of four packets can be in transit in the path at any one time. This prevents a node from hogging the channel for a long period of time.

Two observations are as follows. First, it is not clear that the solution is effective when there are multiple TCP flows along the same path, or when TCP flows on

adjacent paths may interfere with the flow on the path. Second, the instability problem is caused by false declaration of link failures which is rooted at the link layer. This problem is not a phenomenon for TCP traffic only, but also for other types of traffic. Therefore, we believe a more general approach should attempt to solve this problem at the link layer.

There are two possible link-layer solutions: 1) do not declare link failures before a new path can be discovered; or 2) control the offered load at the source to reduce contention rate.

3.3.1 Link-Failure Re-routing

Strictly speaking, in the above scenario the link has not failed, although it is congested and the attempt to look for a new path is definitely warranted. However, before a new route can be discovered, one should continue to use the old route. That is, a "don't-break-before-you-can-make" strategy should be adopted.

To show that the throughput oscillations are in fact due to triggering of re-routing, we disabled the link-failure triggered re-routing function in one of our simulations. Figure 4 shows the result. The throughput becomes more stable and the drastic drops in throughput are eliminated.

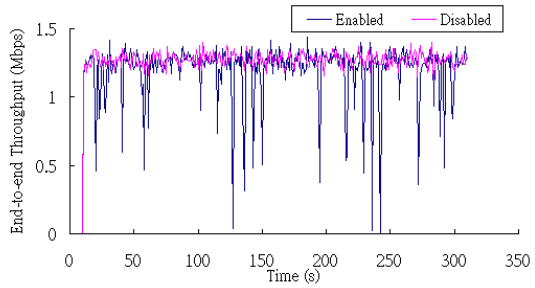

Figure 4. End-to-end throughputs with link-failure declarations enabled/disabled

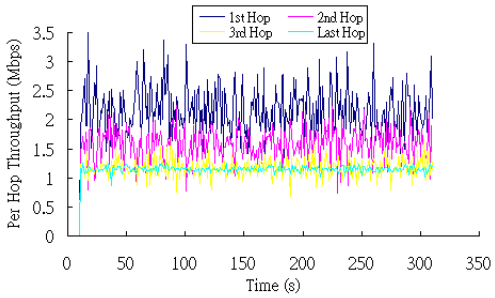

Figure 5. Per-hop throughputs of an 8-node flow after disabling link-failure re-routing

The study of multi-hop routing is beyond the scope of this paper. Here, we just want to point out that false triggers of re-routing should be studied as a separate problem. It could be more effectively dealt with directly rather than indirectly through higher-layer mechanisms. We refer interested readers to [12] in which the "don't-break-before-you-can-make" strategy was implemented. Simulation results in the paper showed that the strategy can prevent the re-routing instability problem and reduce the throughput variations in multi-hop ad-hoc networks drastically.

Figure 5, however, shows that the average throughput still drops from 2.14Mbps in the first hop to 1.15Mbps in the last hop even when re-routing is disabled. The high packet-loss rate remains.

3.3.2 Controlling Offered Load

To prevent high packet loss rate for a flow, the offered load must be controlled. Figure 6 plots the end-to-end throughput of a 12-node multi-hop path versus offered load. The peak throughput is obtained at offered load of 1.18Mbps. Offered load beyond this is unsustainable and high loss rate results because Throughput < Offered Load. This existence of an optimal offered load for a multi-hop path was also pointed out in [2]. In this paper, we provide an analysis to estimate the maximum sustainable throughput, and in doing so, reveal the factors that govern it.

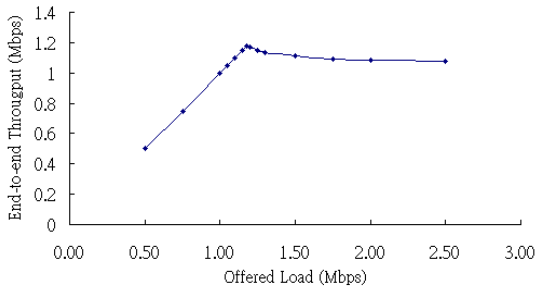

Figure 6. End-to-end throughput versus offered load in a 12-node flow

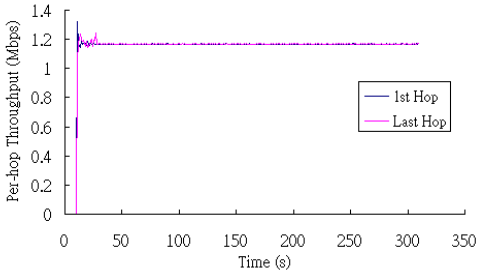

Figure 7. Per-hop throughputs with offered load control (at 1.18Mbps).

Controlling offered load also prevents the instability problem even when the link-failure-triggered re-routing in the routing agent is enabled. Figure 7 shows that the instability problem is eliminated by setting the offered load at the optimal sending rate (1.18Mbps). However, the instability problem is solved by avoiding congestion condition rather than the removal of the problematic strategy of suspending the link usage before a new route can be discovered. A temporary external interference source (e.g., a nearby microwave oven) can easily cause the condition to arise again. We believe that even when offered-load control is exercised, a mechanism to deal with re-routing instability, such as that in [12], is sill needed.

4. Offered-Load Control in 802.11 Networks: Single-Flow Analysis

We now consider the problem of determining the optimal offered load (i.e., the maximum sustainable throughput) for a single flow in a multi-hop network. The throughput is limited by two factors: 1) the hidden-terminal and exposed-terminal problems; and 2) the carrier sensing mechanism. We first analyze the impact of these two factors. After that, we present numerical results showing that the analytical results match the simulation results closely. Our analysis yields a closed-form solution, which we believe provides the insight and foundation for the study of more complex situations involving multiple flows in future work.

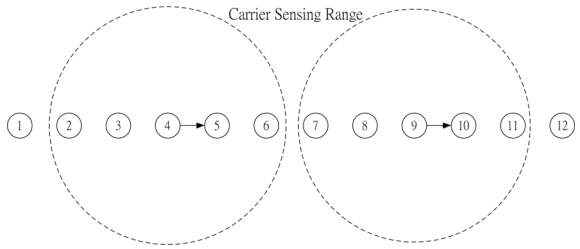

Figure 8. A 12-node string multi-hop network

4.1 Capacity Limited by the Hidden-terminal and Exposed-terminal Problems

We will express the throughput of a single flow in terms of the airtime used by a node. Figure 8 shows a chain of 12 nodes. The traffic flows from left to right. Imagine that this is a longer chain with more nodes extending to the left of node 1 and the right of node 12. By the time the traffic reaches node 1, a "steady state in space" has been reached in which all nodes experience the same situation without the boundary effects. The question we ask is "What is the maximum throughput that can flow through this chain?"

Consider a long stretch of time in the interval [0, *Time*]. Let Si be the airtime within this interval that a "steady-state" node i transmits. This airtime includes the transmission times of the data packets (PACKET), the transmission times of the acknowledgements (ACK) from node (i+1), the durations of the distributed interframe space (DIFS) and the durations of the short interframe space (SIFS). Also, included in Si are the times used up for retransmissions in case of collisions. However, Si does not include the count-down of the idle slots of the contention window, since adjacent nodes can count down together and these count-down times are not unshared resources used up exclusively by node i.

Let $x = |Si| / Time$, T = traffic throughput (in Mbps) flowing through the a "steady-state" node (and therefore also the end-to-end throughput), and ρ = the collision probability for a transmission. Then, we have.

$$T = x \cdot (1 - \rho) \cdot d \cdot data\_rate \quad (1)$$

where $d = DATA / (DIFS + PACKET + SIFS + ACK)$ which is the proportion of time within x that is used to transmit the data payload; and $data\_rate$ is the data transmission rate. Note that *DATA* is the pure payload transmission time of a packet, while *PACKET* includes transmission times of the physical preamble, MAC header, and other higher-layer headers.

For simplicity, we assume that the carrier-sensing mechanism eliminates collisions to the extent that they are negligible, and that collisions are predominantly caused by hidden and exposed nodes. Consider node 4 in Fig. 8. Our assumption means that the transmission of node 4 will not collide with the transmissions of nodes 2, 3, 5, and 6; but node 1 and node 7 may cause collisions at node 4 due to the exposed and hidden-node effects, respectively.

To derive ρ, we consider the "vulnerable period" induced by the hidden and exposed nodes. During a vulnerable period, a node may suffer a collision if it transmits a packet. ρ can be decomposed into two factors: 1) the collision probability due to a hidden node (ρ_{HT}) ,and 2) the collision probability due to an exposed node (ρ_{ET}). They are related as follows:

$$\rho = 1 - (1 - \rho_{HT})(1 - \rho_{ET}) \quad (2)$$

In the following subsections, we first explain the effect of the packet arrival order on signal capture. Then,

we derive ρ_{HT} and ρ_{ET}. We show that the later is relatively small and can be ignored.

Our analysis is based on the following assumptions:
(A.1) The transmission of a node is independent of the transmissions of nodes outside its carrier sensing range.
(A.2) The packet collision probability of a node with nodes inside its carrier sensing range is negligible, thanks to the carrier-sensing property of CSMA.

4.1.1 Signal Capture

In Fig. 9, both nodes 4 and 7 have a packet to transmit. This may cause the aforementioned hidden-terminal collision. However, the signal capturing property may still allow a packet from node 4 to be received successfully, provided it transmits before node 7.

More specifically, suppose that node 4 transmits first and the signal power of the transmission received at node 5 is P_4. Node 7 then transmits a packet with power of P_7 at node 5. If $P_4 > P_7 + CPThreshold$, where *CPThreshold* is the capture threshold, then no collision occurs, and node 5 can still receive the packet from node 4 successfully.

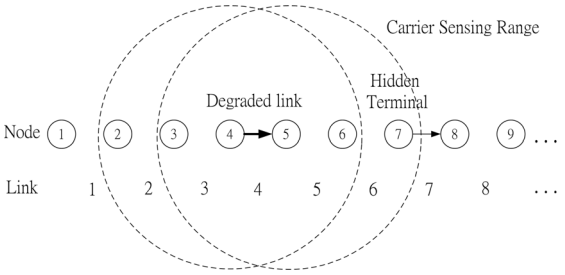

Figure 9. Node 7 as a hidden-terminal to node 4

However, if node 7 transmits first, node 5 senses the signal from node 7 and declares the channel to be busy. In that case, a newly arriving packet from node 4 can not be received even if $P_4 > P_7 + CPThreshold$. Effectively, the packet from node 4 to node 5 experiences a collision.

For the sake of argument, suppose that *CPThreshold* is set to be 10dB. Let d be the fixed distance between nodes. In this case, node 4 and node 7 are separated by a distance larger than the carrier sensing range. Thus, node 4 and node 7 can send packets at the same time. From [14], in a two ray propagation model, the signal-to-noise ratio at node 5 is

$$SNR = P_4 / P_7 = (2d/d)^4 = 2^4 = 16 > CPThreshold$$

This means that the power level of the packet transmitted by node 4 and received at node 5 is always more than *CPThreshold* higher than the power level of the received signal from node 7.

4.1.2. Analysis of Vulnerable Period induced by Hidden Nodes

In the analysis of the hidden-node problem, the key is to identify the vulnerable period during which the transmission of a node will collide with the transmission of a hidden node. This is illustrated in Fig. 10. Note that a hidden-node collision only occurs if the transmissions of nodes 4 and 7 overlap and that the transmission of node 7 precedes that of node 4.

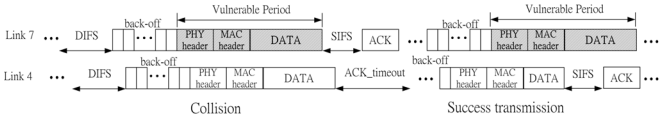

Figure 10. Collision occurs when the transmission of node 4 begins inside the vulnerable period.

If this were an Aloha network, nodes 4 and 7 could collide at anytime during the interval [0, *Time*]. However, in a carrier-sense network, some of the times during this interval must be removed from the "sample space" in the analysis of collision probability.

Consider Fig. 8. When node 5 or 6 transmits, node 4 and node 7 will not by assumption (A.2). This means that *S4*, *S5*, and *S6* are non-overlapping; and *S5*, *S6*, and *S7* are non-overlapping. In particular, node 7 cannot cause collision on node 4 during *S5* and *S*6. Now, nodes 5 and 6 use up $2 \cdot x$ fraction of the airtime during [0, *Time*]. The remaining fraction of airtime where node 4 and node 7 may collide is (1- $2 \cdot x$). Since node 7 uses x fraction of remaining airtime for transmissions, the vulnerable period induced by node 7 on node 4 is

$$\rho_{HT} = \frac{x}{1-2x} \cdot a \qquad (3)$$

by assumption (A.1), where
$a = (PACKET)/(DIFS + PACKET + SIFS + ACK)$
is fraction of time used for transmitting the data packet.

4.1.3. Analysis of Vulnerable Period induced by Exposed Nodes

In Fig. 11, nodes 1 and 4 are outside the carrier-sensing range of each other. At a given time, both nodes 1 and 4 attempt to send a packet to nodes 2 and 5, respectively.

Node 1 is outside the carrier-sensing range of node 4, so the transmission of node 1 does not affect the transmission of node 4. However, node 2 is inside the carrier-sensing range of node 4. Node 4 can sense the

ACK returned from node 2 to node 1. When the ACK from node 5 overlaps with the ACK from node 2 at node 4 and the ACK from node 5 reaches node 4 later than that of node 2 as shown in Fig. 12, a collision occurs.

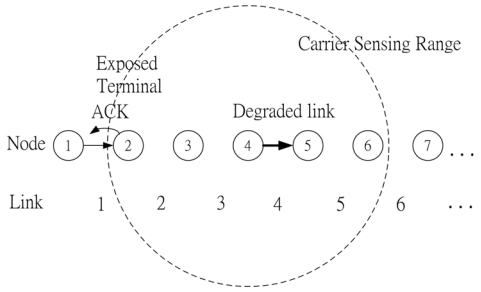

Figure 11. Node 2 as an exposed-terminal to node 4

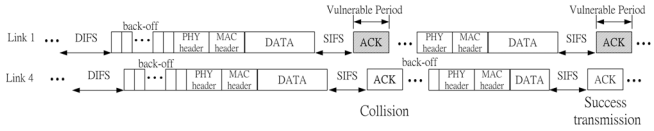

Figure 12. Collision occurs when the ACK from node 5 begins inside the vulnerable period.

However, this ACK-ACK collision can only occur if the transmission of node 4 begins at time t < SIFS later than the transmission of node 1. When t > SIFS, the transmission of node 4 is still in progress and node 4 is not aware of the transmission of ACK from node 2: that is, node 4 will not be able to read the physical preamble in ACK from node 2 and initiate the physical carrier-sensing mechanism that prevents node 4 from receiving the ACK from node 5 later. Therefore, no collisions can occur if t > SIFS. Under the randomization assumption of (A.1), the chance for t < SIFS equals: $SIFS/(DIFS + PACKET + SIFS + ACK) = 0.0064$ under the settings in Table I. Therefore, the ACK-ACK collision rarely happens. This has been borne out by our simulations, in which we could not detect collisions due to the exposed-node problem. We will therefore assume that the degradation caused by exposed nodes is negligible in our analysis henceforth. That is, equation (2) becomes

$$\rho \approx \rho_{HT} \tag{4}$$

4.1.4. Sustainable Throughput

Substituting equations (3) and (4) in (1), we have

$$T = x \cdot (1 - a \cdot \frac{x}{1-2x}) \cdot d \cdot data\_rate \tag{5}$$

Physically, there are two factors affecting T in the opposing directions. As x increases, more airtime is used by a node and there is less idling, and this should push T up. However, larger x also leads to a larger vulnerable period, pulling T down.

Differentiating (5) with respect to x and setting $dT/dx = 0$, the optimal value of x that maximizes the throughput is given by

$$x^* = \frac{(2+a) - \sqrt{a^2 + 2a}}{4 + 2a} \tag{6}$$

Substituting equation (6) in (1) yields the maximum sustainable throughput $T(x^*)$. The offered load should be set to a value smaller than $T(x^*)$ to prevent excessive packet loss.

4.2 Capacity Limited by Carrier Sensing Property

Carrier sensing prevents simultaneous transmissions of nodes within the carrier-sensing range of a node. This imposes a limit on channel spatial-reuse. Potentially, the throughput could be limited by carrier sensing rather than hidden nodes. The maximum throughput derived above is due to hidden nodes. We now consider whether carrier sensing further reduces the sustainable throughput.

Consider node 4 and nodes within its carrier-sensing range in Fig. 8. The total airtimes used up by these nodes cannot exceed *Time*. That is, $|S2 \cup S3 \cup S4 \cup S5 \cup S6| \le Time$.

Define $y = |S2 \cup S3 \cup S4 \cup S5 \cup S6| / Time$, to be the fraction of airtime used up by these nodes within the interval [0, *Time*]. Now, $|S2 \cup S3 \cup S4 \cup S5 \cup S6|$ can be decomposed using the inclusion-exclusion principle:

$$|S2 \cup S3 \cup S4 \cup S5 \cup S6| = |S2| + |S3| + ... + |S6| - |S2 \cap S3| - |S2 \cap S4| - ... + |S2 \cap S3 \cap S4| + ...$$

However, we note that the intersection of the airtimes used by any three nodes or above is null, thanks to carrier sensing. In addition, the intersections of airtimes used by two nodes are non-null only for $S2 \cap S5$, $S3 \cap S6$, and $S2 \cap S6$. We therefore have

$$y \cdot Time = \sum_{i=2}^{6} |Si| - |S2 \cap S5| - |S3 \cap S6| - |S2 \cap S6| \tag{7}$$

Consider the overlapped airtimes of node 2 and node 5. When node 3 or 4 transmits, node 2 and 5 do not, by virtue of carrier sensing. Following similar derivations as in Section 4.1.2, the remaining fraction of airtime where $S2$ and $S5$ may overlap is (1-2x). In particular, we have

$$|S2 \cap S5| = |S3 \cap S6| = \frac{x^2}{1-2x} \cdot Time \quad (8)$$

Nodes 3 and 6 face the same situation. Hence, $|S2 \cap S5| = |S3 \cap S6|$ in (8).

For $|S2 \cap S6|$, the amount of airtime of node 2 that may overlap with that of node 6 is ($|S2|$-$|S2 \cap S5|$), and the amount of airtime of node 6 that may overlap with that of node 2 is ($|S6|$-$|S3 \cap S6|$). The "sample space" within which $S2$ and $S6$ may overlap is [0, *Time*] – $S3$ – $S4$ – $S5$. As a result, we have

$$|S2 \cap S6| = \frac{(|S2| - |S2 \cap S5|) \cdot (|S6| - |S3 \cap S6|)}{Time - |S3| - |S4| - |S5|}$$

The above gives

$$|S2 \cap S6| = \frac{(x - x^2/(1-2x))^2}{1-3x} \cdot Time \quad (9)$$

Substituting equations (8) and (9) into (7), we have

$$y = 5x - \frac{2x^2}{1-2x} - \frac{x^2(1-3x)}{(1-2x)^2} \quad (10)$$

The value of x for $y > 1$ is an "infeasible region". Let the x at which $y(x) = 1$ be x'. If $x' > x^*$ in equation (6), then the system throughput is limited by hidden nodes. However, if $x' < x^*$, the system is limited by the spatial-reuse limit caused by the carrier-sensing mechanism. In the next subsection, we show that for the case under study, the system throughput is hidden-node limited.

4.3 Numerical Results

In Sections 4.1 and 4.2, we have provided analysis on the capacity limited by 1) hidden terminals and exposed terminals and 2) the carrier sensing property. We now examine the numerical results. Table I shows the system parameters assumed, and the associated analytical T and y.

Table I. System parameters and Max Throughput.

Packet payload (*DATA*)	1460 bytes
UDP/IP header	20 bytes
MAC header	28 bytes
PHY header	24 bytes
ACK size	14 bytes
Channel bit rate	11 Mbps
PHY header bit rate	1 Mbps
Slot time	20 *us*
SIFS	10 *us*
DIFS	50 *us*
$CW_{\min}$	32
$CW_{\max}$	1024
Retransmission limit	7

x^*	0.2291
$T(x^*)$	1.1193Mbps
$y(x^*)$	0.8959

For 1), Figure 13 shows the simulation results, which indicate that the optimal offered load (or sustainable throughout) decreases as the number of nodes increases in a string multi-hop topology. For chains with more than 20 nodes, the optimal offered load stabilizes at 1.16Mbps. Our analytical result yields 1.12Mbps, a close match.

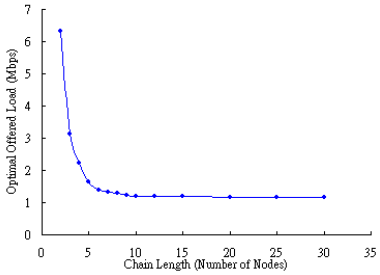

Figure 13. Optimal offered load versus number of nodes in a string multi-hop network.

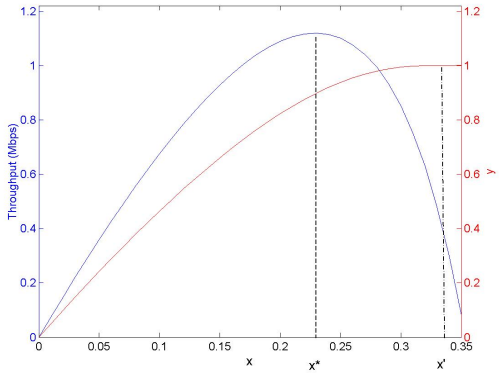

Figure 14. The flow throughput *T* in Mbps (left y-axis) and the fraction of airtime *y* used by all nodes within a carrier sensing range (right y-axis) versus the airtime *x* used by a node.

For the analytical results, Fig.14 plots network throughput T (left y-axis) versus x as limited by the hidden-node effect, and y (right y-axis) versus x as limited by carrier sensing. The maximum $T(x^*)$=1.12Mbps is achieved with x^*=0.229. For x^*, y = 0.896 < 1. This means that the capacity of the network is limited by hidden nodes rather than carrier sensing. Note that when the number of nodes within a carrier-sensing region is large and the number of hidden nodes is small, the capacity could in principle be limited by carrier sensing instead. This could be the case, for example,

when the carrier sensing range is much larger than that of the transmission range.

For the interested reader, reference [15] showed that the carrier-sensing mechanism of 802.11 may impose a constraint on channel spatial-reuse that is overly restrictive, making the network performance non-scalable. The same paper also provides a scheme that modifies 802.11 slightly to achieve scalable performance. We believe the scheme may relieve both the carrier-sensing and hidden-node effects being investigated here, although further study will be needed to validate this conjecture.

5. Achieving Fairness in Other Network Topologies: Multi-flow Investigation

We have shown that controlling the offered load at the source node of a single-flow path eliminates high packet-loss rate. In this section, we will show that controlling the offered load can achieve fairness of channel bandwidth usage among multiple flows. Due to space limitation, the detailed analysis will be deferred to another paper.

5.1 Lattice Topology

To study the interactions among multiple flows, we consider an *N* x *M* lattice network as shown in Fig. 15. All nodes are separated by 200m. The nodes in the first column are the source nodes, and each of them injects traffic into the networks destined for nodes in the last column. In our simulation, we set *M*=*N* for convenience sake.

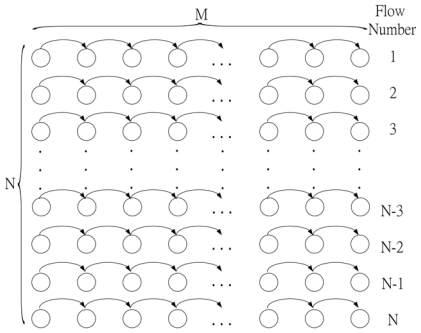

Figure 15. An *N* x *M* lattice topology with *N* traffic flows from left to right

Figure 16 shows that the average end-to-end throughput of all flows decreases as the size of the lattice increases. Reference [2] reported a similar trend in the lattice topology. In addition, we observe an unfairness problem between flows. Figure 17 shows the per-flow end-to-end throughput of a 4x4 lattice network. The flows on two sides (flow 1 and 4) have fewer interfering stations than the middle flows (flow 2 and 3). This causes the flows on two sides to pump more traffic into the network than the middle flows. In the 4x4 lattice network, flow 2 and flow 3 have to compete with the aggressive transmissions of flow 1 and flow 4, resulting in severe throughput degradations.

The uneven numbers of competing stations in the lattice structure severely degrades the performances of flows in the middle. Controlling the offered load in lattice networks prevents aggressive transmissions from two sides to give more chances for nodes in the middle to transmit.

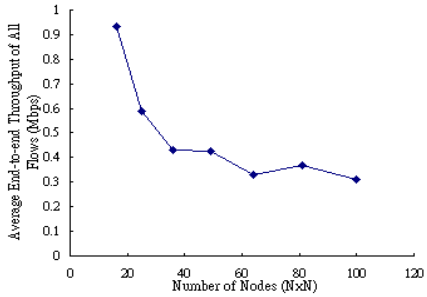

Figure 16. Average end-to-end throughput of all flows versus number of nodes in an *N* x *N* lattice network when the source nodes inject traffic into the network in a saturated manner

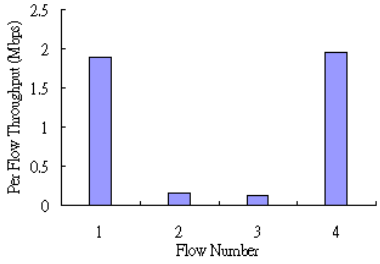

Figure 17. Per-flow end-to-end throughput of a 4x4 lattice network with saturated traffic sources

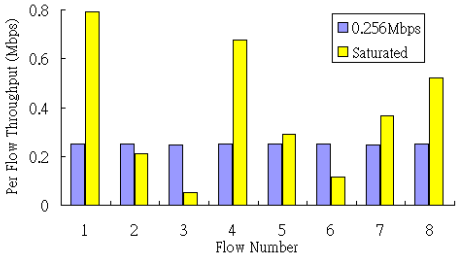

Figure 18. Per-flow throughput of an 8x8 lattice network with the offered load of 0.256Mbps and saturated traffic sources.

Figure 18 shows that a fair share of the channel throughput among the flows in an 8x8 lattice can be achieved when the offered loads at the sources are limited to 0.256Mbps. This sustainable offered load is obtained by extending the single-flow analysis given in the preceding sections. Although the average end-to-end throughput is slightly lower than that of using saturated traffic sources, controlling the offered load can prevent unacceptable per-flow throughput performance and achieve fair bandwidth allocation.

6. Conclusion

This paper is an attempt to identify the maximum throughput that can be sustained in an 802.11 multi-hop network. Our contributions are two-folds:

1. We have shown that uncontrolled, greedy sources can cause unacceptably high packet-loss rate, large throughput oscillations, and unfair bandwidth allocations among traffic flows. Judicious offered load control at the sources, however, can eliminate these problems effectively without modification of the 802.11 multi-access protocol.
2. We have established an analytical framework for the study of the effects of hidden nodes and carrier-sensing operation. This analysis allows one to determine whether the system throughput is hidden-node limited or spatial-reuse limited. In particular, we have shown that the maximum sustainable throughput is limited by two factors: (i) the vulnerable periods which depend on the numbers of hidden nodes and the fraction of airtime in the time horizon when hidden-node collisions may occur; (ii) the number of nodes within a carrier-sensing region and the total airtime used up by them.

We believe that this is a first paper in the literature to provide a *quantitative analysis* on the fundamental impact of hidden nodes and carrier sensing on system throughput.

The single-flow analysis in this paper serves as a "building block" for the study of the multiple-flow case, in which besides self-interference induced by traffic of the same flow, there are also mutual interferences among traffic of different flows. By way of an example, we have shown how to apply the single-flow result to control the offered loads of multiple non-overlapping flows in a lattice network. More complicated situations with overlapping multiple flows remain to be further investigated. We believe the approach in this paper provides a good foundation for such an extension.

References

[1] P. Gupta, P. R. Kumar, "The Capacity of Wireless Networks", *IEEE Trans. Inform. Theory*, Vol.46, No.2, pp.388-404, Mar. 2000.

[2] J. Li, C. Blake et al., "Capacity of Ad Hoc Wireless Networks", *ACM MobiCom'01*, Rome, Italy, July 2001.

[3] K. Jain et al. "Impact of Interference on Multi-hop Wireless Network Performance", *ACM MobiCom'03*, San Diego, USA, Sept. 2003.

[4] M. Kodialam, T. Nandagopal, "Characterizing Achievable Rates in Multi-hop Wireless Networks: The Joint Routing and Scheduling Problem", *ACM MobiCom'03*, San Diego, USA, Sept. 2003

[5] K. Xu, M. Gerla, S. Bae, "How Effective is the IEEE 802.11 RTS/CTS Handshake in Ad Hoc Networks?", *IEEE GLOBECOM '02*, Vol. 1 , pp. 17-21, Nov. 2002.

[6] S. Ansari et al. "Performance Enhancement of TCP on Multihop Ad hoc Wireless Networks", *IEEE ICPWC'02*, pp. 90-94, Dec. 2002.

[7] Z. Hadzi-Velkov, L. Gavrilovska, "Performance of the IEEE 802.11 Wireless LANs under Influence of Hidden", *IEEE PWCS'99*, pp. 221-225, Feb. 1999.

[8] S. Khurana et al., "Effect of Hidden Terminals on the Performance of IEEE 802.11 MAC Protocol", *IEEE LCN'98*, pp. 12-20, Oct. 1998.

[9] S. Khurana et al., "Performance Evaluation of Distributed Co-Ordination Function for IEEE 802.11Wireless LAN Protocol in Presence of Mobile and Hidden Terminals", *IEEE MASCOTS'99*, pp.40-47, Oct. 1999.

[10] F. A. Tobagi, L. Kleinrock, "Packet switching in radio channels: Part ii - the hidden terminal problem in carrier sense multiple-access and the busy-tone solution", *IEEE Trans. on Commun.*, pp.1417–1433, December 1975.

[11]"The Network Simulator–ns2",http://www.isi.edu/nsnam/ns

[12] P. C. Ng, S. C. Liew, "Re-routing Instability in IEEE 802.11 Multi-hop Ad-hoc Networks", *IEEE WLN'04*, Nov. 2004, Tampa, USA.

[13] S. Xu, T. Saadawi, "On TCP over Wireless Multi-hop Networks", *IEEE MILCOM 2001*, Vol.1, pp.282-288, Oct. 2001.

[14] T. Rappaport, "Wireless Communications: Principles and Practice", Prentice Hall, New Jersey, 2002.

[15] P.C. Ng, S. C. Liew, L. B. Jiang, "A Performance Evaluation Framework for IEEE 802.11 Ad-hoc Networks", *ACM PE-WASUN'04*, Venice, Italy, Oct. 2004.