Modeling Per-flow Throughput and Capturing Starvation in CSMA Multi-hop Wireless Networks

Michele Garetto, Theodoros Salonidis, and Edward W. Knightly
Department of Electrical and Computer Engineering
Rice University, Houston, TX 77005

***Abstract*— Multi-hop wireless networks employing random access protocols have been shown to incur large discrepancies in the throughputs achieved by the flows sharing the network. Indeed, flow throughputs can span orders of magnitude from near starvation to many times greater than the mean. In this paper, we address the foundations of this disparity. We show that the fundamental cause is not merely differences in the number of contending neighbors, but a generic coordination problem of CSMA-based random access in a multi-hop environment. We develop a new analytical model that incorporates this lack of coordination, identifies dominating and starving flows and accurately predicts per-flow throughput in a large-scale network. We then propose metrics that quantify throughput imbalances due to the MAC protocol operation. Our model and metrics provide a deeper understanding of the behavior of CSMA protocols in arbitrary topologies and can aid the design of effective protocol solutions to the starvation problem.**

I. Introduction

In a multihop wireless network in which not all nodes are within radio range of each other, Carrier Sense Multiple Access protocols (CSMA) employing any variant of hand shaking (pure CSMA, MACA, MACAW, IEEE 802.11 with two- or four-way handshake, etc.) are known to yield severe unfairness and even complete flow *starvation*. Predicting such behavior using an analytical model is of crucial importance in the development of appropriate counter-measures as well as in wireless system design.

In this paper we develop a novel analytical model that is able to compute the individual throughputs of all flows in an arbitrary network topology. Existing techniques for arbitrary topologies either do not capture the behavior of sophisticated access protocols such as IEEE 802.11 [4], [9], [21], [23] or do not take into account the lack of coordination inherent in the asymmetry of the multi-hop setting [6], [8].

The key to predict the throughput of each transmitting node is to correctly derive the fraction of time sensed busy by the node and the probability of a successful transmission attempt. In contrast to previous work, we accurately analyze and compute these quantities for any node in an arbitrary topology. To do so, we use a decoupling technique to describe the behavior of each node based on its private view of the channel state. In addition to the backlogged source analysis common to the 802.11 modeling literature [3], [5], [6], our model allows computation of flow throughputs for any set of input rates in the network. This is an important characteristic of our analysis as it also enables using the model in applications related to admission control and analysis of rate-control mechanisms.

Starvation manifests as a throughput distribution in which a few dominating flows receive very high throughput and many starving flows receive very low (sometimes zero) throughput. Flow starvation is naturally captured by our model as extended periods of carrier sensing or high packet loss probability or both. We show that such effects are not due to a particular CSMA protocol such as IEEE 802.11 but rather have their foundations in a generic coordination problem of CSMA random access when applied to multi-hop wireless networks, and in the use of carrier sense itself.

To evaluate any solution to the starvation problem we must be able to quantify the effect of the MAC-related starvation problem on the throughput distribution. As different solutions can produce radically different per-flow throughput distributions, we show that traditional metrics such as aggregate utility are not sufficient to quantify starvation. Instead, we utilize Lorenz curves and the Gini index, used in the economics literature to quantify a society's distribution of wealth to individuals, to evaluate the network's distribution of throughput to flows. These metrics, as well as new metrics including flow-preference graphs, a "poverty index," and a "disproportionality index" accurately capture throughput inequalities caused by a specific protocol solution; they also distinguish starvation effects due to lack of coordination from inherent throughput imbalances due to differences in the number of contending neighbors.

The remainder of this paper is structured as follows: In Section II we review the state of the art of modeling CSMA random access protocols. Section III presents the analytical model and pinpoints the fundamental sources of starvation. In Section IV we compare the predictions of our model to *ns* simulations and evaluate the contribution of various factors to throughput prediction. Section V introduces the starvation metrics. In Section VI we present an application example of our model and metrics to the study of a simple rate-limiting policy to combat starvation. We conclude in Section VII.

II. Related work

Analytical modeling of random access protocols in *single-hop* wireless networks originates in models for Aloha [1] and CSMA [14]. More recently, models for 802.11 have been proposed that characterize the throughput achieved by saturated sources [3], [5]. Analysis of single-hop networks is greatly simplified by the fact that all contenders within the

This research was supported by NSF ITR Grant ANI-0325971, by the Cisco ARTI Program, and by the Lagrange Project of the CRT foundation.

single contention region have a common view of the channel and can coordinate their transmissions. Unfortunately, this is not the case for the multi-hop topologies considered here.

Models for *multi-hop* wireless networks come in varying degrees of analytical detail and topology assumptions. Detailed analyses exist for restricted topology configurations. Capacity and delay for both Slotted Aloha and CSMA for a two-hop wireless access network have been studied in [10] and [20], respectively. For 802.11, [7] contains a detailed Markov chain throughput analysis for the scenario in Fig. 3. In [17], a queuing-theoretic analysis of the scenario in Fig. 2 is developed and extended to a chain topology of single-hop flows. However, these techniques exploit the specialized topology structure and do not predict per-flow throughput in a general topology.

Existing models of CSMA protocols for arbitrary topologies can be classified in *transmission set* approaches and *node-based* approaches. The transmission set approach was first proposed in [4] to check if a given set of input rates is feasible, assuming Poisson packet arrivals on the links. The system behavior is described by a continuous time Markov chain where each state is a set of nodes that can transmit simultaneously (an independent set in the communication graph). The analysis leads to a product form solution and an iterative method determines if the achieved throughputs match the input rates. Extensions to the model of [4] include using links instead of nodes as transmitting entities [21], [23] and CDMA signaling [9]. While the model of [4] accounts for carrier sense mechanisms, it does not capture the behavior of sophisticated random access protocols such as 802.11. In particular, it neither models packet losses due to the MAC protocol operation nor the Binary Exponential Backoff (BEB) mechanism. Another drawback is that it requires computation of all independent sets in the network, an NP-complete problem.

In the node-based approach, each flow is viewed in isolation; the conditional packet loss probability of a flow is a function of the product of the transmission probabilities of interfering flows in the immediate neighborhood. Based on this approach, [8] presents a model for 802.11 without BEB that takes into account capture effects and hidden terminals. Reference [6] proposes an enhanced model for computing throughputs of saturated link flows taking into account the 802.11 intricacies, RTS/CTS and virtual carrier sense; it also incorporates channel errors for each link. Since no computation of independent sets is involved, the node-based approach is computationally efficient. Unfortunately, this approach does not take into account the coordination problems due to carrier sensing and topology asymmetries. In section IV we will show that such problems play a dominant role in the resulting throughput distribution.

III. Analytical Model

In this section we develop a new analytical framework to study the behavior of CSMA/CA random access protocols in ad hoc networks with arbitrary topology. While we focus on the 802.11 protocol family, our analysis is general enough to study different variations of CSMA/CA. Our objective is to predict the throughput of individual stations in an arbitrary topology. We consider the case in which all traffic produced or relayed by a station is directed to a single neighboring node. This is the necessary first step before analyzing more complex traffic patterns.

To understand the role of the MAC protocol, we purposely neglect physical layer issues that would complicate the analysis without eliminating the critical MAC issues. Therefore, we consider a simplified physical layer model as follows. (i) The *transmission range* R_T of each station is fixed, and all stations transmit at the same power. Only stations within radius R_T of the sending station are capable of successfully decoding its packets. (ii) The *sensing range* R_S of each station is fixed. Only stations within radius R_S of a transmitting station receive enough power to trigger carrier sense and detect a busy channel. (iii) There are no power capture effects: a packet is not received correctly at a station if it overlaps with at least one other packet transmitted by a station within distance R_S. (iv) The communication channel is error-free : a received packet is always decoded correctly in absence of collision. While it is possible to incorporate a more realistic physical layer in our model, in Section IV we will show simulation results indicating that this would not affect the core starvation issue due to MAC interactions that we study here.

The remainder of the section is organized as follows. In Section III-A we describe the behavior of an individual station employing the CSMA/CA protocol and derive the basic variables needed to predict its throughput, under the assumption that the station is backlogged. The computation of these variables is described in two sections: in Section III-B we study how a station senses the activity of other stations in its neighborhood; in Section III-C we describe how to compute the collision probability of a station. The extension of the model to analyze the case of non-backlogged sources is presented in Section III-D. In Section III-E we explain how to solve for the entire system, i.e. how to compute the throughputs of all sources in the network, considering their mutual interaction. Finally, we present numerical results and simulations in Section IV.

A. General model of an individual station

Our modeling framework contrasts with existing techniques (e.g., [3]) because we account for the fact that in an arbitrary topology the perceived channel state is different from station to station. We now summarize a general decoupling technique introduced in [18] that allows us to study the behavior of an individual station based on its private view of the channel.

The channel state as seen by a single source. The behavior of a station employing a CSMA protocol is dictated by the occupation of the 'air' around it (the channel). The evolution of its private view of the channel state can be abstractly represented by a temporal diagram such as the one exemplified in Fig. 1.

We identify 4 different states : *(i)* idle channel; *(ii)* channel occupied by a successful transmission of the station; *(iii)* channel occupied by a collision of the station; *(iv)* busy channel due to activity of other stations, detected by means of either physical or virtual carrier sensing (the NAV). The time

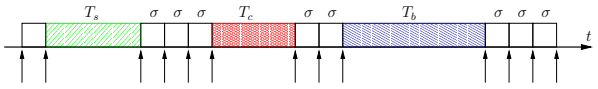


Fig. 1. The channel view of a station and the discrete time renewal process embedded in it.

intervals during which the station remains in each of the four states above are denoted by σ, T_s, T_c, and T_b, respectively. In Fig. 1, the arrows placed below the temporal axis point the time instants of a possible state change. While σ is constant, equal to the backoff slot, the duration of the other intervals can be variable (with general distribution), depending on the access mechanism (basic access or RTS/CTS), the frame size, and the sending rate of the transmitting station(s).

Analysis of the behavior of a single source. To analyze the behavior of the station, we make the fundamental assumption that the channel evolution can be described by a renewal process: at each switching time the next state does not depend on the current state. Let Π_σ, Π_s, Π_c, Π_b be, respectively, the occurrence probabilities of the four states described above. Let τ be the probability that the station sends out a packet after an idle slot, under the assumption that it is always backlogged (we remove this assumption later). Let p be the probability that a transmission of the station is not successful. The probability p is usually referred to as the *conditional* packet loss probability [3]. We also introduce b, the probability that the channel becomes busy after an idle slot due to activity of other stations (assuming that the station does not start transmitting). Using these probabilities, we can specify the occurrence probability of each of the four channel states at the switching instants: $\Pi_\sigma = (1-\tau)(1-b)$, $\Pi_s = \tau(1-p)$, $\Pi_c = \tau p$, and $\Pi_b = (1-\tau)b$.

Computation of the throughput. Using standard renewal theory, the throughput of the station (expressed in packet/s), is given by $T_P = \frac{\Pi_s}{\Delta}$, where Δ is the average duration of a channel state (in seconds). Thus the final formula for the throughput of a station is,

$$T_P = \frac{\tau(1-p)}{\tau(1-p)\bar{T}_s + \tau p\,\bar{T}_c + (1-\tau)(1-b)\sigma + (1-\tau)b\,\bar{T}_b} \quad (1)$$

The probability τ is a deterministic function of p, which depends only on backoff parameters such as the window size, the number of backoff stages, etc. For 802.11, the expression of τ as a function of p is given in [3]. Recently, it has been shown in [15] that similar expressions of τ as a function of p can be derived for a large class of backoff mechanisms, employing arbitrary window distributions and backoff multipliers. The complete expression of τ for 802.11 that takes into account the maximum retransmission limit jointly with the maximum window size, is given by

$$\tau = \frac{2q(1-p^{m+1})}{q(1-p^{m+1}) + W_0\left[1-p-p(2p)^{m'}\left(1+p^{m-m'}q\right)\right]} \quad (2)$$

where $q = 1-2p$, W_0 is the minimum window size, m is the *maximum retry limit*, and m' is the backoff stage at which the window size reaches its maximum value, $m' \leq m$.

The average durations $\bar{T}_s$ and $\bar{T}_c$ of a successful transmission or of a collision in which the station is involved are also known *a priori* (see [3]). It turns out that the only unknown variables in Equation (1) are: i) the occurrence probability b of a busy period, and its average duration $\bar{T}_b$; ii) p, the conditional packet loss probability. These variables are specific to each station and result from the interaction of the station with the rest of network.

Some remarks on the throughput formula (1): as expected, throughput is a decreasing function of the collision probability p (notice that τ is also a decreasing function of p). Moreover, it is a decreasing function of both b and $\bar{T}_b$ (notice that $\bar{T}_b$ is much larger than σ). Intuitively, the larger the fraction of time the station senses the channel busy (which is proportional to both b and $\bar{T}_b$), the fewer its transmission opportunities. A transmitter is expected to *starve* when either p or the product $b\bar{T}_b$ are disproportionally large as compared to neighbors. Notice that these two reasons are not mutually exclusive, and they can combine to yield even lower throughput.

Now, one would expect that, as a result of fair contention among stations, nodes relatively close to each other (e.g., in the same transmission region or clique) would have similar values of packet loss probability and a similar number of transmission opportunities. Instead, CSMA-based protocols like IEEE 802.11 DCF can introduce large discrepancies between the observed values of these quantities in neighboring nodes. This fact can be demonstrated even in very simple topologies. For example, Fig. 2 depicts a topology with two flows, where the receiver of Flow 1 (node a) is in range of the transmitter of Flow 2 (node B), whereas the senders are not in range of each other. This is a well-known problematic scenario for a random access protocol [2], [12]. The two flows form a clique, but Flow 1 achieves significantly less throughput than flow 2 (by a factor of 10 or higher) under saturated conditions, regardless of whether the RTS/CTS handshake is used or not (the RRTS message does not solve the problem either). Indeed flow 1 experiences a very large packet loss probability (in same cases $p(A)$ is close to 100%) whereas flow 2 suffers no loss. As a consequence, the throughput of flow 1 approaches zero, even if the fraction of time sensed busy by sender A is null (i.e., $b(A) = 0$). We refer to this problem as **Information Asymmetry (IA)**, because the root cause of the observed behavior resides in an asymmetric view of the channel state as perceived by senders. In [18] we have analyzed this case in detail using the modeling framework described in this Section.

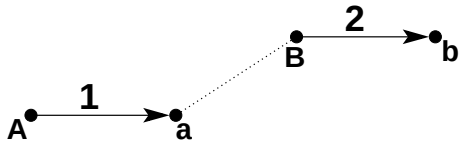


Fig. 2. Information Asymmetry (IA) example

Another emblematic case is illustrated in Fig. 3. Here there are three links, two of which (flow 1 and 3) are out of range from each other. If all flows are backlogged, the middle flow (flow 2) receives very low throughput, while the outer flows (1

and 3) receive a throughput close to the maximum. This time the cause of the observed behavior is not the loss probability (indeed, the value of p is negligible for all three flows) but the lack of transmission opportunities for the middle flow. The temporal diagram reported in Fig. 4 illustrates how the activity of the outer links is sensed by node B. Senders A and C are not coordinated so that their transmissions overlap randomly at node B, which senses the channel busy for prolonged periods. Flow 2 has a chance to capture the medium only when *both* outer flows are in the back-off phase (in Fig. 4, the channel is sensed idle by sender B only during the small interval comprised by the vertical lines). This problem has been already identified by researchers (see for example [7], [23]), and we refer to it as **Flow-in-the-Middle (FIM)**, as it occurs whenever a sender senses the activity of two other flows that do not sense each other.

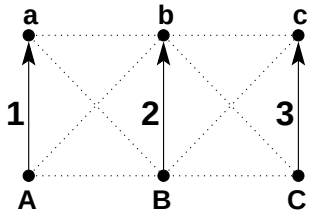


Fig. 3. Flow-In-the-Middle (FIM) example

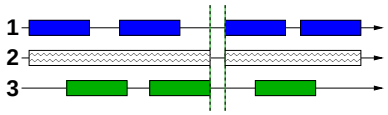


Fig. 4. The channel activity of links 1 and 3 as sensed by sender node B

Both IA and FIM scenarios are known to be especially problematic when flows are backlogged. In a general topology with many active links, nodes can be part of multiple cliques and they interact with many other nodes in their vicinity. Thus, a flow can potentially suffer from many IA and FIM problems, but with respect to flows that do not send at their maximum possible rate. Therefore, one could question what is the real impact of these problems in the case of a large, dense network. Our main contribution is indeed a model that incorporates both problems in a general setting, allowing us to assess their real influence for an arbitrary network.

Sections III-B and III-C are devoted to to the computation of b, T_b, p of a single station as a function of the channel activity of neighboring stations. In the following derivations, we restrict ourselves to the computation of $b(i)$, $T_b(i)$ and $p(i)$ for a tagged station i, assuming that the behavior of all other stations is completely specified (i.e., variables $b(j)$, $T_b(j)$ and $p(j)$ of every station j other than i are known). This computation is a fundamental part of the system-wide procedure that determines all station throughputs as described in Section III-E.

B. Computation of the fraction of busy time of a node

In this section we compute the amount of time during which a tagged node i senses the channel busy due to the activity of other stations in its vicinity. In 802.11 the channel is sensed busy as a result of two different functions: physical carrier-sensing or virtual carrier-sensing. The latter is provided by the Network Allocation Vector (NAV), which is updated by inspecting the duration fields of correctly received frames.

In single-hop 802.11 systems all stations are synchronized and the duration of a busy period equals (for all stations) the duration T_s of a successful packet exchange or the duration T_c of a collision. However, in the multi-hop case, determination of this quantity is more involved: due to lack of coordination, transmissions of different stations can overlap randomly at the tagged station. This can lead to busy channel periods much longer than T_s or T_c for the tagged station, as we have seen, for example, in the FIM scenario of Fig. 3.

Since at this point we assume that the activity of the other stations is known, all we need to do is characterize the overlapping of these activities as perceived by the tagged station, in terms of the probability b and the average duration $\bar{T}_b$ as defined in Section III-A. Our analysis is divided into 4 steps, described in the following subsections.

1) **Step 1: Computation of the maximal cliques:** The first step is the computation of all maximal cliques that the tagged station i belongs to. A maximal clique of station i is defined as a set of transmitting stations such that: 1) station i senses the activity of all other stations in the set; 2) any two pairs of stations in the set are in conflict; 3) the set is not contained in any other set that satisfies properties 1 and 2. Exact computation of all maximal cliques of an arbitrary graph is known to be a hard problem. However, efficient techniques exist to find all maximal cliques in the contention graph induced by a wireless network. In particular, we modified the approximate technique proposed in [11], and developed a *scanning disk* algorithm to find all maximal cliques of a station according to our definition.

2) **Step 2: Cliques reduction:** The second step of our analysis consists of a clique reduction algorithm aimed at finding the minimum set of maximal cliques covering all stations sensed by station i. This is a set covering problem which is also difficult to solve optimally; thus we resort to the following greedy algorithm: we compute the aggregate channel activity of all maximal cliques found in the first step of the analysis. We take the clique with the largest aggregate activity, and add it to the set of the covering cliques (initially empty). Then we remove from the remaining cliques all stations belonging to the extracted clique and recompute the aggregate activity of each clique. We repeat the previous steps until all stations are covered.

Using a geometric argument, it can be shown that in the worst case, the maximum number of cliques required to cover all stations sensed by station i is 6. In practice, the average number of maximal cliques to be considered is much smaller than 6. For example, in the dense topology of Fig. 13 the average number of cliques found by our algorithm is 2.74, and no station requires more than 4.

3) **Step 3: Computation of the active regions:** In the third step of the analysis, we compute the sets of stations resulting from the intersection of any combination of the

cliques determined during the second step. We discard all sets containing no stations, and call the remaining sets the *active regions* around station i. It is seen that the maximum number of regions is $2^M - 1$, where M is the number of cliques found by the second step. Next, we compute the aggregate activity of each region by combining the activity of each individual station in the region. This can be done easily, since the activity of stations in a region do not overlap in time. The activity of station j in a region can be described in terms of an *on-off* process: the duration of an *on* interval is the amount of time station i senses the channel busy because of the activity of j. During the *off* interval, the channel would be sensed idle by station i, if j were the only station sensed by i. We describe the resulting on-off process of station j by two variables: the average duration $\bar{T}_{ON}(j)$ of the on period, and the rate $\lambda(j)$ with which *on* events are generated by station j. These two variables can be readily computed, because we assume that the behavior of station j is completely specified. The exact computation of $\bar{T}_{ON}(j)$ and $\lambda(j)$ depends on the geometric relationships between the links used by i and j, and on the protocol behavior. For example, if i only hears the receiver of the link of station j, it does not sense the first packet of a data exchange (RTS or DATA) over that link. If station i does not decode the last packet of a data exchange over link j, the standard mandates that an additional time equal to EIFS has to be waited before i can resume decrementing its counter, thus this additional EIFS can be considered part of $\bar{T}_{ON}(j)$. Here we omit such details and consider that the sum of the activity of all stations in a region U generates an aggregate on-off process, whose parameters are $\lambda(U) = \sum_{j \in U} \lambda(j)$ and $\bar{T}_{ON}(U) = \frac{\sum_{j \in U} \lambda(j)\, \bar{T}_{ON}(j)}{\lambda(U)}$. After computing these values, we can substitute the entire active region U with a single *virtual* node u having the same on-off process. This concludes the analysis of the active regions of station i (step 3).

4) **Step 4: Computation of the busy period:** In the fourth step of the analysis, we apply a Markovian modeling technique to analyze the network of *virtual* nodes computed in the third step. For simplicity, we assume that the durations of the *on* and *off* period of each virtual node are exponentially distributed.

A virtual node is characterized by the *activation rate* g_u and the de-activation rate $\mu_u = 1/\bar{T}_{ON}(u)$. While μ_u is known, g_u has to be computed iteratively, such that the "effective" activation rate of the virtual node is equal to the desired rate λ_u.[1]

We consider independent sets D of virtual nodes (an independent set comprises virtual nodes that can be active simultaneously) and compute the steady-state probability that nodes in the set are jointly in the on period, denoted by $Q(D)$. The system allows a product form solution of the type

$$Q(D) = \left(\prod_{u \in D} \frac{g_u}{\mu_u} \right) Q(\phi) \tag{3}$$

where $Q(\phi)$, the probability that none of the virtual nodes is active, can be obtained by normalizing the sum of all probabilities $Q(D)$ to one. The quantity $1 - Q(\phi)$ is nothing else than the fraction of time sensed busy by station i, the basic quantity that we want to compute. From the solution of the system we can also derive the average duration of a busy period of station i, $\bar{T}_b(i)$, as required by the station model described in Section III-A. Indeed, the duration of an idle period of station i is exponentially distributed with parameter $\lambda_{idle}(i) = \sum_u g_u$, so that the average duration of an idle period is $\bar{T}_{idle}(i) = 1/\lambda_{idle}(i)$. It follows that the average duration of a busy period of i is $\bar{T}_b(i) = \frac{\bar{T}_{idle}(i)\,[1-Q(\phi)]}{Q(\phi)}$.

To compute probability $b(i)$, which is the other variable that we need to specify for station i, we need another quantity $n_e(i)$, which is the average number of events generated by the stations sensed by i during a busy period of i. Notice that these events can overlap in time while they are sensed by i, thus $n_e(i) \geq 1$. Since the aggregate event rate is known a priori, we have $n_e(i) = \sum_u \lambda(u)/[\bar{T}_{idle}(i) + \bar{T}_b(i)]$. Having computed this quantity, we can determine the probability $b(i)$ that a busy period starts at station i after an idle slot has elapsed, under the condition that station i does not start its own transmission. The following equality has to be satisfied:

$$\sum_u \lambda(u) = \frac{[1 - \tau(i)]\, b(i)\, n_e(i)}{\Delta(i)} \tag{4}$$

which is a throughput expression stating that station i must "see" the aggregate event rate of its neighbors while in the backoff phase. From Equation (4) we can derive $b(i)$ as a function of all known variables.

Example: We illustrate the analysis presented in this section through an example. The left part of Fig. 5 depicts a portion of the network around a tagged station i, of which we want to compute the fraction of busy time. The network area around i comprises all stations whose activity is sensed by i, i.e. either transmitters or receivers of active flows. Suppose that, after computing all maximal cliques of i (step 1), our clique reduction algorithm identifies 3 maximal cliques A, B, C covering all of the stations (step 2). There are 7 possible intersections among these cliques, and in our example all of them comprise at least one station, forming 7 active regions labeled $1 \ldots 7$. We describe the activity of each region as sensed by station i by an on-off process, and we substitute each region with a *virtual* node (step 3). The conflict graph between the virtual nodes is shown on the right part of Fig. 5. An edge exist between two virtual nodes if their respective regions cannot be active at the same time. From the conflict graph we derive all possible independent sets of virtual nodes, which are the following: $\{\phi\}$, $\{1\}$, $\{2\}$, $\{3\}$, $\{4\}$, $\{5\}$, $\{6\}$, $\{7\}$, $\{1,2\}$, $\{1,5\}$, $\{1,3\}$, $\{2,3\}$, $\{2,6\}$, $\{3,4\}$, $\{1,2,3\}$.

The activation rate g_u of each virtual node is computed iteratively. For example, consider virtual node 4. The "effective" activation rate of this node is equal to $g_4[Q(\phi) + Q(3)]$, because region 4 (see Fig. 5) can become active only when none of the other regions is active, or when only region 3 is active. Since we want the effective rate of node 4 to be equal to λ_4, we must have $g_4 = \lambda_4/[Q(\phi) + Q(3)]$. The values g_u

[1]The activation rate g_u is usually larger than the (desired) rate $\lambda(u)$, because a virtual node cannot become active in any state of the system model.

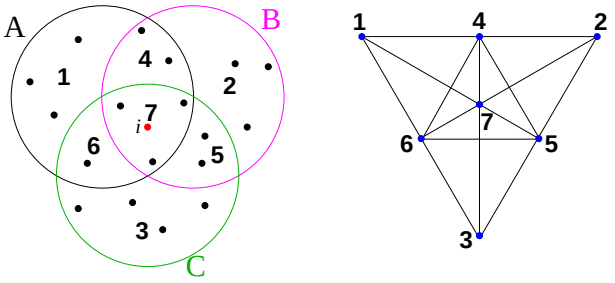


Fig. 5. Example of computation of the fraction of busy time of a tagged station i

typically converge in a few iterations. After convergence has been reached, we know exactly the fraction of busy time of station i, and we can compute variables $\bar{T}_b(i)$ and $b(i)$ to be used in the model of station i.

We remark that our technique to analyze the fraction of busy time of a node has some aspects in common with the model originally proposed in [4]. There, the authors study the throughput of CSMA networks by considering all independent sets of transmitting nodes, and develop a product-form solution similar to ours. They assume that the throughputs λ_i of nodes are known a priori: if the iterative computation of rates g_i converges, they conclude that the set of rates λ_i can be supported by the network (i.e., it's feasible). They do not consider any specific random access protocol (just deferral due to carrier sense), and assume that all transmissions are successful. Our analysis differs substantially from the one in [4]. Our goal is not to check if a given set of rates is feasible, but to compute the actual throughput of each node under a given random access mechanism (in our case, 802.11 DCF). Instead of considering all independent set of nodes in the network (which can be computationally very expensive), we consider independent cliques around each node (which is much more scalable), and derive information about the channel view of the node which is used in a global procedure to compute its throughput. Moreover, our analysis is computationally more efficient (it allows to analyze large, dense networks) and more accurate (because it considers the behavior of a real protocol). For example, under a realistic random access mechanism the channel is occupied not only by successful transmissions, but also by collisions, while the access rate of nodes depends on the backoff window mechanism, which can adapt to congestion (e.g. through the binary exponential backoff). These are fundamental aspects that have to be taken into account to analyze the behavior of multi-hop wireless network.

C. Computation of the conditional packet loss probability

In this section we compute the conditional packet loss probability p of a tagged station i. This turns out to be one of the most complicated variables to be computed in an arbitrary topology. Previous work yields only coarse approximations of this probability and, more importantly, neglects fundamental MAC layer effects that can cause one station to have surprisingly large values of p, e.g., the IA scenario of Fig. 2.

We identify four different categories of packet losses due to the MAC protocol: (i) Losses due to collisions between **coordinated stations**, which occur with probability p_{co}. (ii) Losses due to **Information Asymmetry**, which occur with probability p_{ia}. (iii) Losses due to **near hidden terminals**, which occur with probability p_{nh}. (iv) Losses due to **far hidden terminals**, which occur with probability p_{fh}. Our classification is exhaustive, i.e., it accounts for all possible losses that can occur between any two links in *conflict*. We independently analyze each category in the next subsections, precisely defining the geometry conditions under which they occur.

The way in which the different loss probabilities are combined to obtain the overall packet loss probability of a station is through their complementary successful probability: a transmission is successful if it does not not suffer from any of the four type of losses mentioned above (which can happen concurrently):

$$p(i) = 1 - [1 - p_{co}(i)]\,[1 - p_{ia}(i)]\,[1 - p_{nh}(i)]\,[1 - p_{fh}(i)] \quad (5)$$

1) **Losses between coordinated stations:** Collisions between coordinated stations are the traditional type of losses due to the MAC protocol considered in the design and analysis of 802.11 networks. Indeed, when all sources are in range of each other, the DCF function is able to synchronize all stations in such a way that all transmission attempts happen at well defined slot boundaries recognized by all stations. As a result, in this network scenario the conditional packet loss probability is simply given by $1 - (1 - \tau)^{n-1}$, where n is the total number of stations [3]. Similar formulas have been proposed to predict the packet loss probability in general networks, considering all sources around the transmitter or the receiver of a given link [6]. Unfortunately, in a multi-hop topology the DCF function fails to synchronize all stations, and the analysis is complicated. Indeed, in an arbitrary network, collisions still occur between transmitting stations locally coordinated by the protocol function, but such losses usually account only for a small part of the overall packet loss probability, as we will show.

In our model, we define "losses between coordinated stations" as collisions that occur at a station simultaneously receiving two packets from stations that are in range of each other. Indeed, transmitters must be at least within sensing range of each other in order for this kind of loss to occur. We now provide an expression for the collision probability $p_{co}(i, i')$ suffered by a given link $l(i, j)$ because of simultaneous transmissions of station i', assuming $d(i, i') < R_S$, $d(j, i') < R_S$ (here $d(m, n)$ denotes the euclidean distance between nodes m and n). One computed, the probabilities $p_{co}(i, i')$ can be combined to obtain the overall collision probability $p_{co}(i)$ of link $l(i, j)$.

To compute $p_{co}(i, i')$ we need the probability $c(i'|i)$ that station i' can transmit a packet, given that station i is transmitting (recall that $p(i)$ is a conditional collision probability, subject to the fact that station i is actually transmitting). The probability $p_{co}(i, i')$ is then given by $p_{co}(i, i') = c(i'|i)\,\tau(i')$. The probability $c(i'|i)$ can be obtained as a by-product of the

analysis of the fraction of busy time of station i' (Section III-B). Indeed it is given by

$$c(i'|i) = \frac{Q(\phi)}{\sum_{D \in \mathcal{A}(i)} Q(D)} \tag{6}$$

where $\mathcal{A}(i)$ is the set of states in which station i (more precisely, the active region of i' where i is located) can become active.

2) **Losses due to Information Asymmetry:** The Information Asymmetry problem can cause one flow to experience a very large packet loss probability, in some cases equal to 1. Thus, it is of crucial importance to accurately model these losses, as they are a major cause of starvation.

Link $l(i,j)$ suffers from the Information Asymmetry problem from link $l'(i',j')$ if the following geometric relationships among the stations are satisfied:

1) $d(i,i') > R_S$: transmitters are not in range.
2) $d(j,i') < R_S$: the receiver of link l is in range of the transmitter of link l'.
3) $d(i,j') > R_S$: the receiver of link l' is not in range of the transmitter of link l.

As already explained in Section III-A, the only chance sender i has to successfully transmit is when the initial packet of the two-way or four-way handshake (a DATA frame or an RTS frame) happens to arrive at j during a "gap" in the activity of link l'. This is indeed the fundamental observation that allows us to compute the loss probability of i.

Now we provide an expression for the loss probability $p_{ia}(i,i')$ suffered by a link $l(i,j)$ because of the activity of link $l'(i',j')$, given that the geometric conditions of the IA scenario are satisfied. Such probabilities can then be combined to obtain the overall loss probability $p_{ia}(i)$ suffered by i because of IA problems.

We model the activity of link l' as perceived by station j while station i attempts to transmit as an on-off process: the *on* period is the interval during which station j would not be able to receive the first packet sent by i and reply with a control message, provided that the first bit arrives during the *on* interval. The *on* interval has a constant duration T_{ON} that depends only on the geometric relationship between the two links. The *off* period is the gap in the activity of link l' that sender i has to discover randomly.

The average duration $\bar{T}_{OFF}$ can be computed as

$$\frac{1}{T_{ON} + \bar{T}_{OFF}} = \frac{\lambda(i')}{\sum_{D \in \mathcal{A}(i)} Q(D)} \tag{7}$$

which states that the rate at which *on* periods occur (the inverse of the average duration of a cycle) must be equal to the sending rate of station i', conditioned by the probability that station i can start transmitting, which is again obtained from the analysis of the busy period of i'.

Assuming that the duration of the *off* period is exponentially distributed, and that transmission attempts of sender i arrive at random points in time, we obtain the following expression of $p_{ia}(i,i')$

$$p_{ia}(i,i') = 1 - \frac{\bar{T}_{OFF}}{T_{ON} + \bar{T}_{OFF}} e^{-\frac{d}{\bar{T}_{OFF}}} \tag{8}$$

where d is the duration of the first packet sent by i (RTS or DATA). Equation (8) states that a transmission attempt of i is successful if and only if the first packet arrives during an *off* period *and* it is fully received by j before the beginning of the next *on* period.

3) **Losses due to near hidden terminals:** Near hidden terminal losses occur between two links $l(i,j)$ and $l'(i',j')$ arranged in such a way that the following geometric relationships among the stations are satisfied:

1) $d(i,i') > R_S$: transmitters are not in range.
2) $d(j,i') < R_S$: the receiver of link l is in range of the transmitter of link l'.
3) $d(i,j') < R_S$: the receiver of link l' is in range of the transmitter of link l.

The difference with respect to the IA scenario lies in the fact that now both receivers are in range of the transmitter of the other flow, producing a symmetric scenario illustrated in the left part of Fig. 6. Collisions occur because, during the time a sender is transmitting its first packet, the other sender starts its own transmission, and both packets are destroyed at the receivers.

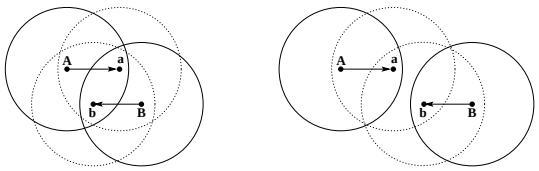


Fig. 6. Examples of near hidden terminals (left) and far hidden terminals (right)

Now we provide an expression for the loss probability $p_{nh}(i,i')$ suffered by a link $l(i,j)$ because of the activity of link $l'(i',j')$, assuming that the geometric conditions listed above are satisfied. Such probabilities can then be combined to obtain the overall loss probability $p_{nh}(i)$ suffered by i because of near hidden terminals.

The collision probability $p_{nh}(i,i')$ is computed as follows,

$$p_{nh}(i,i') = c(i'|i)\,[1 - (1 - \tau(i'))^m] \tag{9}$$

where $c(i'|i)$ is given by (6), whereas m is the number of transmission opportunities of i' during the duration of the first packet sent by i. We have $m = \lfloor d/\sigma \rfloor$, with d denoting the duration of the first packet sent by i. Since m can be large (equal to 13 for an RTS packet, and even larger for a DATA packet, depending on the payload size), it is clear that such losses can be significant, although they affect both flows, which consequently decreases their attempt probability τ, thereby reducing the chances of repeated collisions.

4) **Losses due to far hidden terminals:** Far hidden terminal losses occur between two links $l(i,j)$ and $l'(i',j')$ arranged in such a way that the following geometric relationships among the stations are satisfied:

1) $d(i,i') > R_S$: transmitters are not in range.
2) $d(j,i') > R_S$, $d(i,j') > R_S$: receivers are not in range of the transmitter of the other flow.
3) $d(i,j') < R_S$: the two receivers are in range of each other.

An example of this scenario is depicted in the right part of Fig. 6. In this case, the two links are connected only through their respective receivers, yet it is not possible to have two overlapping transmissions that complete successfully, because the control packets sent by one receiver interfere with the reception of packets at the other receiver. Although the configuration is symmetric, a loss event for one link does not correspond to a loss event for the other link. The reason is that whichever sender starts first can complete the transmission successfully even if the other sender attempts a transmission in the meantime.

We can express the loss probability $p_{fh}(i, i')$ of station i due to a far hidden terminal i' as the probability that i starts transmitting during an active period of i'. Viewing the activity of station i' as an on-off process, we compute the durations $T_{ON}(i')$ and $\bar{T}_{OFF}(i')$ in the same way as in Section III-C.2, and we have that

$$p_{fh}(i, i') = \frac{T_{ON}(i')}{T_{ON}(i') + \bar{T}_{OFF}(i')} \tag{10}$$

D. Extension to non-saturated conditions

The model of a single station (Section III-A) can be extended to the case of non-saturated conditions in which the transmission queue of a source can become empty. To this purpose, we assume that each sender station receives a given input rate Λ of packets to send on the channel, and that the actual throughput T_P achieved by the station is equal to the input rate Λ up to a saturation value T_{satur} (to be computed), after which it remains constant and equal to T_{satur}.

We add a new probability e to the description of the behavior of a single station as introduced in Section III-A, which is the conditional probability that the transmission queue of the station is empty, given that the station can potentially start a new transmission (i.e., when its backoff counter is zero and the channel has been sensed idle for a time slot). The occurrence probability of each of the four channel states described in Section III-A are modified as follows, $\Pi_s = \tau\,(1-p)(1-e)$, $\Pi_c = \tau\,p(1-e)$, $\Pi_\sigma = [(1-\tau)+\tau e]\,(1-b)$, $\Pi_b = [(1-\tau)+\tau e]\,b$, and the throughput formula (1) has to be changed accordingly. As a result, the only additional variable that we need to compute is the probability e. This can be achieved considering that $T_P = \min(\Lambda, T_{satur})$ as described above, and assuming that all other variables are known: if $T_P|_{e=0} < \Lambda$, then $e = 0$ (the source is saturated); otherwise e is equal to the value e^* such that $T_P|_{e^*} = \Lambda$, which is easily obtained by inverting the throughput formula.

E. Network solution

Determining the throughputs of all flows network-wide requires solving a coupled nonlinear multivariate system of equations. We adopt an iterative procedure: during each iteration we utilize the analysis of Sections III-B and III-C to update the variables $\{b(i), T_b(i), p(i), e(i)\}$ of every station as a function of the corresponding variables of its neighbors, as computed in the previous iteration. Extensive experiments have verified convergence to a single global solution for different topologies and initial conditions. Similar iterative techniques have been used in [3], [19] to analyze single-hop 802.11 networks.

It is interesting to note that, due to the link interdependencies imposed by the broadcast wireless medium, changes in the input rates of nodes in one part of a multi-hop wireless network can drastically affect the throughputs of nodes several hops away; indeed, we have observed that such changes may even propagate in the entire network (we do not show examples here due to space constraints). Our model can track the effect of such perturbations as it can predict the throughput distribution for any set of input rates.

IV. Simulations and model validation

To validate our model, we consider the topology of Fig. 7, where 50 nodes are randomly placed in a 1000 m x 1000 m area. Both transmission and sensing ranges are set equal to 200 m. In the following we refer to this scenario as "200/200" topology.

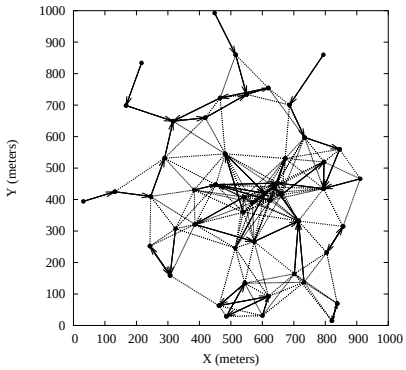


Fig. 7. 50-node topology 200/200

Each node attempts to transmit to a single, arbitrarily selected neighbor within its transmission range. In Fig. 7 the link flows are denoted by solid arrows while dotted lines connect nodes that are in range of each other. We consider saturated sources: each node sends at the maximum transmission rate using IEEE 802.11 DCF with RTS/CTS enabled.

We perform ns simulations using the 802.11b parameters listed in Table 1. All simulations use a 20s warm-up period and all quantities of interest are measured over the next 600s.

SIFS	10 μs
DIFS	50 μs
EIFS	364 μs
σ	20 μs
BasicRate	2 Mbps
DataRate	11 Mbps
PLCP length	192 bits @ 1 Mbps
MAC header (RTS,CTS,ACK,DATA)	(20,14,14,28) bytes @ BasicRate
Packet size	1000 bytes
(CW_{min}, CW_{max})	(31,1023)
Retry Limit (Short,Long)	(7,4)

TABLE I

Parameters setting for the MAC and physical layers.

Fig. 8 compare analytical and simulation results for the individual throughput (number of received packets divided by

total time) of the 50 flows in the 200/200 topology. We observe an excellent agreement between model and simulation results: the average relative error in the prediction of individual flow throughputs with respect to the maximum throughput on a link is 4.8%.

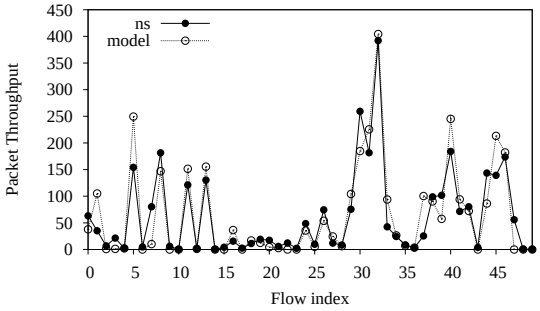


Fig. 8. Throughput of each individual flow in the 50-node topology 200/200

For a deeper validation of our analysis, we have measured in simulation detailed information about each flow. In particular, we have measured the fraction of busy time of each transmitter and the conditional packet loss probability, the two fundamental variables that affect the throughput of a flow (see Section III-A). The fraction of busy time of a node corresponds in the model to the quantity $\frac{(1-\tau)b\bar{T}_b}{\Delta}$. It is measured in simulation considering the amount of time the backoff counter decrement of the node has been suspended because of either physical or virtual carrier sense, divided by the duration of simulation. Fig. 9 reports the fraction of busy time of each node showing a good match between analytical and simulation results.

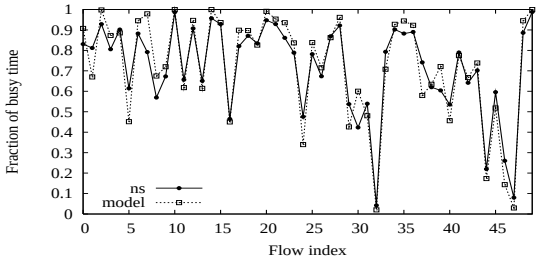


Fig. 9. The fraction of busy time of each node in the 200/200 topology, according to model and simulation.

Fig. 10 compares instead the conditional packet loss probability p of each node of the 200/200 topology. Again, we observe an excellent agreement between analytical and simulation results.

The stacked histograms in Fig. 11 show the contributions to the overall packet loss probability of the four categories of losses (CO,IA,NH,FH) considered in the model (viz., coordinated, information asymmetry, near-hidden and far-hidden terminals, respectively). We observe that losses between coordinated stations (CO), which are the only ones considered by most analytical works appearing in the literature, contribute only to a small fraction of the overall packet loss probability, confirming that it is of crucial importance to consider the different type of losses described in Section III-C.

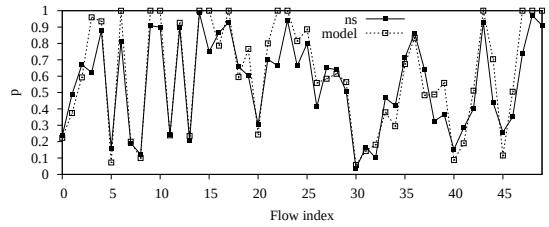


Fig. 10. The conditional packet loss probability of each node in the 200/200 topology, according to model and simulation

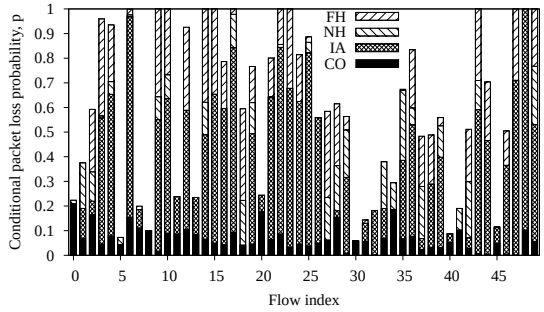


Fig. 11. The contribution of the various type of losses to the overall packet loss probabilities of each nodes in the 200/200 topology

In contrast, the authors of [6] suggest to compute the packet loss probability $p(i)$ of a node i using the simple formula

$$p(i) = 1 - \prod_{k \in V_i \bigcup V_r} [1 - \tau(k)] \qquad (11)$$

where V_i and V_r are the set of nodes within range of i and within range of its receiver r, respectively. Moreover, they linearize the expression of τ as a function of p (equation 2) so as to solve efficiently the interaction among nodes. In Fig. 12 we compare the packet loss probabilities of the individual nodes as obtained in *ns*, and as predicted by the technique proposed in [6] (named CARV04 model).

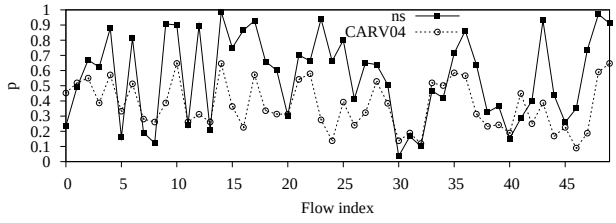


Fig. 12. Packet loss probabilities of nodes in the 200/200 topology according to simulation and based on formula (11)

We observe that the estimates produced by the CARV04 model deviate significantly from simulation, especially for nodes experiencing a very large loss probability. This is due to the fact that (11) implicitly assumes that all transmissions are slot-synchronized, whereas this is not the case in a CSMA multi-hop wireless network.

Next we consider the same topology of Fig. 7, in which we increase the sensing range of nodes to 400 m, while keeping the transmission range equal to 200 m. This configuration,

referred to as 200/400 topology, produces a much higher connectivity, as shown in Fig. 13.

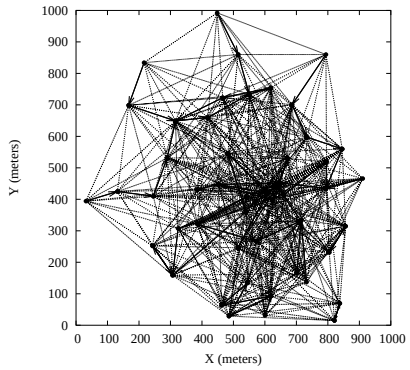


Fig. 13. 50-node topology 200/400

Fig. 14 compares analytical and simulation results for the individual throughput of the 50 flows in the new configuration. We observe that a few flows capture most of the available bandwidth, while the vast majority of flows get negligible throughput. The model identifies correctly the set of dominating flows as well the extent of starvation. The average relative error in the prediction of individual flow throughputs with respect to the maximum throughput on a link is 2.7%.

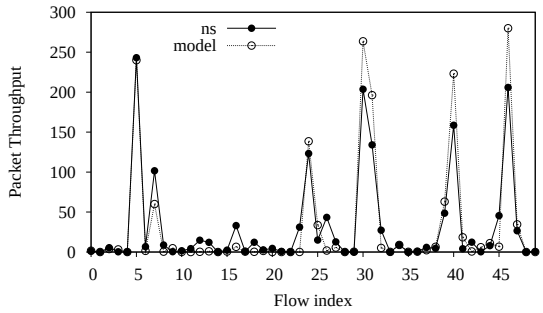


Fig. 14. Throughput of each individual flow in the 50-node topology 200/400

In this topology it is particularly interesting to understand the origin of the severe starvation resulting from the interaction among flows. The few dominating flows are located at the border of the topology, and they are mostly isolated from each other. Poor flows in the center of the network are completely starved because of a superposition of FIM effects due to the rich flows. Thus, in this topology starvation is simply due to the use of carrier sense. We have verified this fact running the model under the (unrealistic) assumption that $p = 0$ for all stations: the set of dominating flows, as well the extent of starvation, are almost unchanged. Moreover, similar results were obtained when RTS/CTS is turned off (not shown here due to lack of space).

We have also investigated how much the observed starvation is related to our simplifying assumption at the physical layer. Fig. 15 depicts the throughput of the flows in the 200/400 topology, sorted from highest to lowest, under different models of the physical layer. The curve labelled "ideal channel" corresponds to the error-free channel with zero capture considered in this paper. The other two curves incorporate power capture effects (with a capture ratio of 10) and independent channel fading on each link according to the Ricean model. In particular, we employ [16] to simulate short time-scale fading (different from link to link) and consider two different values of the Ricean factor K related to the line of sight component, and a path loss exponent of 4. Channel errors and power capture effects indeed mitigate the imbalance among the largest link rates (they also reduce the aggregate network throughput), but do not significantly affect the starvation of the flows in the tail of the distribution.

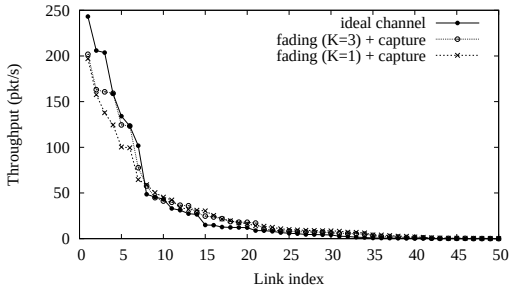


Fig. 15. Link throughput under different models of the physical layer, in the 200/400 topology

V. Measuring starvation

Our model accounts for starvation and the high imbalance in throughput distribution induced by CSMA-based random access protocols. Quantifying this imbalance through a set of metrics is crucial in the development of solutions that effectively address the starvation problem.

We first introduce a reference system where starvation effects due to CSMA MAC are eliminated. We then provide metrics and graphical tools that can accurately capture the effects of starvation and distinguish between imbalance due to topology structure and starvation due to the coordination problems of CSMA MAC protocol.

A. **Reference System**

We consider a random access system in which the root causes of starvation are structurally eliminated by the access mechanism. In particular, we consider a system in which all nodes are synchronized to slots of fixed length equal to the packet size. Each node uses a general random backoff mechanism, and can perform only a single action per slot, transmit or listen, determined at the beginning of each slot. An example of such a system is Slotted Aloha, in which nodes attempt to transmit in each slot with fixed probability. Since in a slotted system all transmissions occur within a slot, and have to start at well defined slot boundaries common to all stations, the starvation effects due to both FIM and IA problems are inherently eliminated.

Without carrier sense, the slotted system must rely only on appropriate setting of per-slot flow transmission probabilities. Fortunately, solutions exist for setting such probabilities to satisfy a desired optimality objective. In particular, it is

possible to compute transmission probabilities that yield max-min fair rates [22] or proportional fair rates [13], with the latter computable using only information within the immediate neighborhood of each flow. In our experiments, we employ the latter approach and set the Slotted Aloha transmission probabilities to achieve proportional fairness in order to balance fairness and throughput objectives.

B. Capturing imbalances in throughput distribution

Fig. 16 depicts the time fractions[2] flows transmit successfully under 802.11 and the slotted system, for the 200/400 50-node topology. The 50 flows have been sorted in decreasing order of the fraction obtained in 802.11. The two systems yield not only significantly different flow time shares, but flows are also ranked differently within each system. Moreover, the *aggregate* time fraction achieved by all links in 802.11 is significantly greater than the slotted system. For example, the aggregate time fraction achieved by 802.11 is 2.875 vs. 0.748 for the slotted system (a 73.9% decrease). However, according to Fig. 16, the slotted system always gives better treatment to the starving 802.11 flows by decreasing the bandwidth of the dominating flows. That is, no flows receive zero throughput and starve in the slotted system.

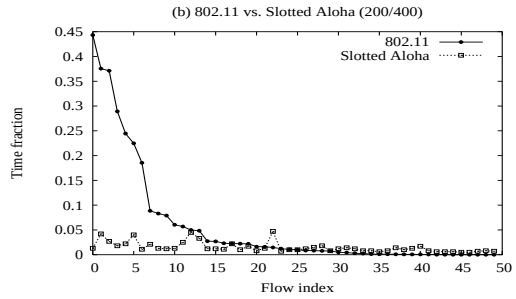


Fig. 16. 802.11 vs. slotted system: throughput curve (200/400)

A performance measure that incorporates both fairness and total system throughput is aggregate network utility. For such a measure, we consider the sum of the logarithms of the fractions of time occupied by the flows. In the 200/400 topology, the aggregate utility of 802.11 is $-\infty$ because some flows achieve exactly zero throughput. In the 200/200 configuration, the aggregate utility for the slotted system is -162.65, whereas for 802.11 it is -168.57. These values are quite close, yet correspond to two systems that behave in a very different manner. Consequently, this metric provides limited insights and we turn to alternate performance measures that can capture the imbalances in the throughput distribution.

The Gini index is used in economics to quantify inequality of resource shares. It is derived from the Lorenz curve, which plots the cumulative share of aggregate throughput achieved by flows ranked from largest to smallest. The Lorenz curves for the two systems in both topology configurations are depicted in Fig. 17. If all flow throughputs were equal, the Lorenz curve would follow the diagonal. As the degree of inequality increases, so does the curvature of the Lorenz curve and thus the area between the curve and the diagonal becomes larger. The Gini index is the ratio of the area between the Lorenz curve and the diagonal over the whole area of the upper half triangle. In the 200/400 configuration, the Gini index for the 802.11 system is 0.76 whereas for the slotted system it is 0.35; this index therefore captures the severely unbalanced throughput distribution in 802.11 as compared to the slotted system. Interestingly, in the 200/200 configuration imbalance among flows reduces in 802.11, whereas it increases in the slotted system (see the right plot of Fig. 17). In this case the Gini index for the 802.11 system is 0.623 whereas for the slotted system it is 0.53. From the Lorenz curve and the Gini index, it is tempting to conclude that in the 200/200 configuration the disparity among flows is almost the same for the two systems, thus in this case 802.11 does not incur significant problems. However, this is an inaccurate conclusion, because it neglects how individual flows are treated differently in the two systems.

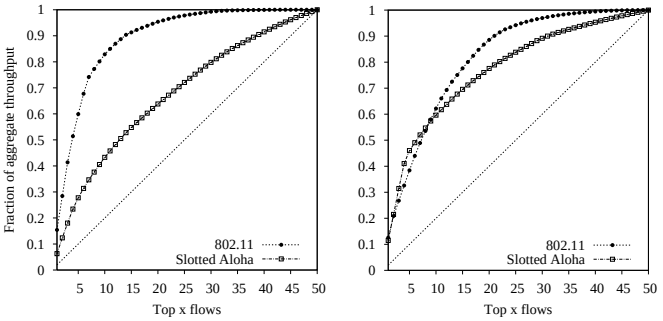


Fig. 17. Lorenz curves of 802.11 vs. slotted system in 200/400 (left) and 200/200 (right)

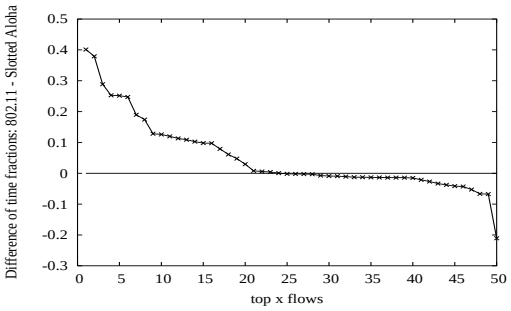


Fig. 18. Flow preference graph (200/200)

C. Isolating MAC starvation from topology-induced imbalance

While the Gini index and the Lorenz curves capture throughput imbalance, they represent the combined effect of *both* imbalance due to the asymmetric network topology (i.e., different number of contenders in different portions of the network) and starvation due to the MAC protocol (e.g., FIM and IA effects). Thus, we need performance measures that *isolate* the effects due to the MAC protocol.

[2]The time fraction of a flow equals its throughput in packets/sec multiplied by the duration of a successful transmission of a packet, which is assumed to be the same in 802.11 and the slotted system.

The "flow preference graph" reports for each individual flow the difference in throughput achieved in the two systems. Such a view allows evaluation of the number of flows that are better off in one system with respect to the other, as well as the extent of the difference. For example, Fig. 18 illustrates the flow preference graph of the 802.11 system over the slotted system in the 200/200 topology, sorting the differences from largest to smallest. The positive and negative regions indicate flows receiving better or worse treatment with the magnitude indicating the per-flow difference of the fractions of time spent in successful transmissions. We observe that flow preferences for 802.11 versus the slotted system are split nearly in half (i.e., half of the flows are better off with 802.11 and half with Slotted Aloha). However, the positive area on the left is significantly greater than the negative area on the right, indicating that a flow receives higher bandwidth under 802.11 on average, and many flows have a strong preference for 802.11.

Intuitively, an efficient solution to MAC-induced starvation should yield a distribution where 1) all flow throughputs are greater than the corresponding throughputs achieved in the slotted system and 2) the throughput ratio of any two flows is close to the corresponding throughput ratio in the slotted system, where MAC-based starvation is not present. We quantify these two desirable features using the *poverty index* P and the *disproportionality index* D, respectively.

Given the flow preference graph, we define the poverty index as the fraction of flows receiving less throughput in 802.11 than in the slotted system. Results reported in Table 2 indicate that the poverty index is close to 0.5 for both configurations 200/400 and 200/200. That is, for 802.11, about half of the flows are worse off than they would be in a Slotted Aloha system, where MAC starvation effects have been eliminated.

The disproportionality index is defined as follows: Let θ denote the angle between the N-dimensional vectors $\vec{x}$ and $\vec{y}$ of the throughput achieved by the flows in 802.11 and in the slotted system, respectively (the same index in either vectors refers to the same flow). Then, we define the disproportionality index D as:

$$D = 1 - cos\theta = 1 - \frac{\sum_{i=1}^{N} x_i y_i}{\sqrt{\sum_{i=1}^{N} x_i^2}\sqrt{\sum_{i=1}^{N} y_i^2}} \qquad (12)$$

This metric is not biased by the fact that the two systems achieve different aggregate throughputs because it depends only on the *angle* between the two vectors. Notice that since the throughputs are non-negative, D takes values between zero and one. A value of one indicates maximum disparity with respect to the slotted system; a value of zero indicates that any two flows have the same throughput ratio in both systems. For example, according to Table 2, in the 200/400 configuration $D = 0.39$, whereas in the 200/200 configuration $D = 0.26$. Hence, starvation is more pronounced in the 200/400 configuration, but it is still present significantly in the 200/200 configuration, although in this case the Gini indices corresponding to the two systems are similar.

	200/400		200/200	
Metric	802.11	Slotted	802.11	Slotted
Min	0	0.005	0.0018	0.0130
Max	0.443	0.047	0.714	0.360
Avg	0.057	0.015	0.113	0.063
Sum	2.875	0.748	5.684	3.134
Gini	0.76	0.33	0.623	0.530
SumLog	$-\infty$	-219	-169	-163
P	0.50	0	0.52	0
D	0.39	0	0.26	0

TABLE II

802.11 VS. SLOTTED SYSTEM: SUMMARY METRICS FOR 200/400 AND 200/200 TOPOLOGY CONFIGURATIONS.

D. Discussion

We conclude that when a protocol mechanism is applied to a multi-hop wireless network to address starvation, the throughput distribution can radically change. In this case, a single measure is not enough to compare the performance of the two systems, and multiple measures, both absolute and relative, are required. Metrics that do not account for the difference in performance of the individual flows (such as the aggregate network utility) do not provide sufficient differentiation of the systems. Instead, it is preferable to use metrics that depend on the specific throughputs achieved by the individual flows in the network. In addition to scalar metrics, the graphical devices we introduced (throughput curves, Lorenz curves and preference graphs) provide additional insights.

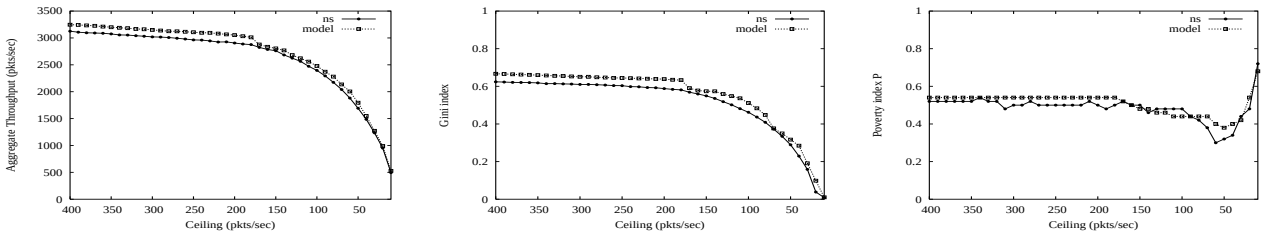
VI. APPLICATION: A SIMPLE RATE-LIMITING POLICY

In this section we present an application of our model and metrics to the study of a simple rate-limiting algorithm aimed at mitigating starvation. The idea is to throttle the dominating flows so that more 'air-time' (i.e. transmission opportunities) is left for the starving flows. Our model can be used to efficiently compute the throughput distribution resulting from any set of input rates, as explained in Section III-D; given a throughput distribution, the metrics defined in Section V are used to quantify the algorithm's ability to combat starvation.

The algorithm is iterative and uses a 'declining ceiling' policy: during each iteration, a ceiling rate is used to limit the input rate of the dominating flows while keeping the starving flows backlogged. A flow is permanently marked dominating if at any iteration it achieves throughput greater than the threshold. Once all flows receive throughput less than or equal to the threshold, the threshold is lowered by a small step. By gradually lowering the threshold, more air time is given to the starving flows at the potential expense of total network utilization.

We consider the 200/200 configuration and study the performance of our rate limiting policy, both when guided by the model and by *ns* simulations. Fig. 19 depicts the evolution of the aggregate throughput, the Gini index and the Poverty index as the threshold starts from 400 packets/sec and decreases to 10 packets/sec.

All graphs indicate an excellent match between model and simulation. Fig. 19(a) depicts how the aggregate throughput



(a) Aggregate Throughput (b) Gini index (c) Poverty index

Fig. 19. Performance of the rate limiting policy for 200/200 configuration.

decreases as the threshold decreases indicating that the starving flows cannot fully exploit the air time taken away by the dominating flows. Furthermore, the increasing slope of the curve at lower thresholds indicates a greater amount of uncompensated aggregate throughput. Fig. 19(b) depicts the decrease of the Gini index as the threshold decreases, with the index eventually reaching a value close to zero– at this point all flows have equal but very low throughput. Fig. 19(c) depicts the poverty index and best illustrates setting of the rates to yield throughputs better than in the Slotted Aloha system: the curve's initial decrease reflects an increasing fraction of flows relieved from starvation problems with a minimum achieved at a threshold of 50 packets/sec. Below this point, the poverty index increases sharply as flows experience low throughput due to rate limiting instead of MAC starvation.

VII. Conclusions

In this paper we developed an analytical model that can accurately predict each flow's throughput in a wireless network of arbitrary topology employing CSMA-based MAC protocol. Using the model, we explained the fundamental origins of starvation that manifests in such a system and proposed metrics and graphical devices that can quantify MAC-based starvation and distinguish it from natural throughput imbalances induced by the difference in number of contenders per geographical locality. Our work provides a deeper understanding of the behavior of CSMA protocols in multi-hop networks and yields new insights that can guide the analysis and design of distributed fair MAC schemes.

References

[1] N. Abramson. The aloha system: another alternative for computer communications. *AFIPS Conference Proceedings FJCC*, 37:281–285, 1970.

[2] V. Bharghavan, A. Demers, S. Shenker, and L. Zhang. MACAW: A Media Access Protocol for Wireless LANs. In *Proc. ACM SIGCOMM*, London, UK, 1994.

[3] G. Bianchi. Performance analysis of the IEEE 802.11 distributed coordination function. *IEEE Journal on Selected Areas in Communications*, 18(3):535–547, March 2000.

[4] R. Boorstyn, A. Kershenbaum, B. Maglaris, and V. Sahin. Throughput Analysis in Multihop CSMA Packet Radio Networks. *IEEE Trans. on Communications*, 35(3):267–274, March 1987.

[5] F. Cali, M. Conti, and E. Gregori. Dynamic tuning of the IEEE 802.11 protocol to achieve a theoretical throughput limit. *IEEE Transactions on Networking*, 8(2):785–799, Dec 2000.

[6] M. Carvalho and J. J. Garcia-Luna-Aceves. A Scalable Model for Channel Access Protocols in Multihop Ad Hoc Networks. In *Proc. ACM MobiCom*, Philadelphia, PA, USA, September 2004.

[7] C. Chaudet, I. Gurin Lassous, and B. Gaujal. Study of the Impact of Asymmetry and Carrier Sense Mechanism in IEEE 802.11 Multi-hops Networks through a Basic Case . In *In Proc. PE-WASUN '04*, Venice, Italy, 2004.

[8] H.S. Chaya and S. Gupta. Performance modeling of asynchronous data transfer methods of IEEE 802.11 MAC protocol. *ACM Wireless Networks*, 3(3):217–234, 1997.

[9] M.S. Chen and R. Boorstyn. Throughput Analysis of CDMA packet radio networks in the presence of noise. In *Proc. IEEE INFOCOM*, Washington DC, USA, March 1985.

[10] I. Gitman. On the capacity of slotted ALOHA networks and some design problems. *IEEE Trans. on Communications*, 23(3):305–317, 1975.

[11] R. Gupta and J. Walrand. Approximating Maximal Cliques in Ad-Hoc Networks. In *Proc. IEEE Personal, Indoor, and Mobile Radio Communications Conference (PIMRC)*, Barcelona, Spain, Sept. 2004.

[12] V. Kanodia, C. Li, A. Sabharwal, B. Sadeghi, and E. Knightly. Ordered Packet Scheduling in Wireless Ad Hoc Networks: Mechanisms and Performance Analysis. In *Proc. ACM MOBIHOC*, Lausanne, Switzerland, June 2002.

[13] K. Kar, S. Sarkar, and L. Tassiulas. Achieving proportionally fair rates using local information in Aloha networks. *IEEE Transactions on Automatic Control*, 49(10):1858–1862, October 2004.

[14] L. Kleinrock and F. A. Tobagi. Packet switching in radio channels: Part I - carrier sense multiple-access modes and their throughput-delay characteristics. *IEEE Trans. on Comm.*, 23(12):1400–1416, 1975.

[15] A. Kumar, E. Altman, D. Miorandi, and M. Goyal. New insights from a fixed point analysis of single cell IEEE 802.11 WLANs. In *Proc. IEEE INFOCOM*, Miami, FL, March 2005.

[16] R. Punnoose, P. Nikitin, and D. Stancil. Efficient simulation of Ricean fading within a packet simulator. In *IEEE Vehicular Technology Conference*, pages 764–767, 2000.

[17] S. Ray, D. Starobinski, and J. Carruthers. Performance of Wireless Networks with Hidden Nodes: A Queueing-Theoretic Analysis . *Journal of Computer Communications (Special Issue on the Performance of Wireless LANs, PANs, and Ad-Hoc Networks)*, in press.

[18] J. Shi, M. Garetto, and E. Knightly. Modeling media access in embedded two-flow topologies of multi-hop wireless networks. In *Proceedings of ACM MobiCom '05*, Cologne, Germany, September 2005.

[19] O. Tickoo and B. Sikdar. Queueing analysis and delay mitigation in ieee 802.11 random access mac based wireless networks. In *Proc. IEEE INFOCOM*, Hong Kong, China, 2004.

[20] F. A. Tobagi. Analysis of a two-hop centralized packet radio network part I: Slotted ALOHA . *IEEE Trans. on Communications*, 2(2):196–207, Feb 1980.

[21] F. A. Tobagi and J. M. Brazio. Throughput analysis of multihop packet radio network under various channel access schemes. In *Proc. IEEE INFOCOM*, San Diego, CA, USA, April 1983.

[22] X. Wang and K. Kar. Distributed Algorithms for Max-min Fair Rate Allocation in Aloha Networks. In *Proc. Annual Allerton Conference On Communication, Control and Computing*, Urbana-Champaign, IL, USA, October 2004.

[23] X. Wang and K. Kar. Throughput Modelling and Fairness Issues in CSMA/CA Based Ad-Hoc Networks. In *Proc. IEEE INFOCOM*, Miami,FL,USA, 2005.