BEWARE: Background Traffic-Aware Rate Adaptation for IEEE 802.11

Shao-Cheng Wang and Ahmed Helmy
Department of Computer and Information Science and Engineering
University of Florida, Gainesville, FL., U.S.A.
Email: shaochew@ufl.edu, helmy@ufl.edu

Abstract

IEEE 802.11-based devices employ rate adaptation algorithms to dynamically switch data rates to accommodate the fluctuating wireless channel conditions. Many studies observed that, when there are other stations transmitting in the network, existing rate adaptation performance degrades significantly due to the inability of differentiating losses between wireless noise and contention collisions. Previous studies proposed to exploit optional RTS frames to isolate the wireless losses from collision losses, and thus improve rate adaptation performance. In this paper, we conduct a systematic evaluation on the effectiveness of various existing rate adaptation algorithms and related proposals for loss differentiations, with multiple stations transmitting background traffic in the network. Our main contributions are two-fold. Firstly, we observe that existing RTS-based loss differentiation schemes do not perform well in all background traffic scenarios. In addition, our study reveals that RTS-based loss differentiation schemes can mislead the rate adaptation algorithms to persist on using similar data rate combinations regardless of background traffic level, thus result in performance penalty in certain scenarios. The fundamental challenge is that a good rate adaptation algorithm must dynamically adjust the rate selection decision objectives with respect to different background traffic levels. Secondly, we design a new Background traffic aware rate adaptation algorithm (BEWARE) that addresses the above challenge. BEWARE uses a mathematical model to calculate on-the-fly the expected packet transmission time based on current wireless channel and background traffic conditions. Our simulation results show that BEWARE outperforms other rate adaptation algorithms without RTS loss differentiation by up to 250% and with RTS by up to 25% in throughput.

1. Introduction

With the large-scale deployments of wireless local area networks (WLANs) in homes, offices, and public areas, the IEEE 802.11 standard has become the dominant technology in providing low-cost high-bandwidth wireless connections. A large part of the success of IEEE 802.11 protocol can be attributed to the implementation of several simple yet fully distributed algorithms in dealing with the fundamental challenges for wireless communications. For example, the IEEE 802.11 standard employs multiple data rates with different levels of complexity and redundancy in signal modulation and coding schemes to combat the volatile nature of wireless channel. IEEE 802.11-based stations then implement rate adaptation algorithm (RAA) to dynamically select the best transmission rate that yields the highest performance in the given wireless channel conditions.

The key challenges are that RAA must not only accurately estimate the channel condition in order to infer the most suitable data rate, but also be very responsive to the rapidly fluctuating wireless channel dynamics. Several approaches have been proposed [1]-[8] to use various metrics such as received signal strength, local Acks, and packet statistics to design a RAA in addressing the above challenges. The effectiveness of RAAs has been extensively evaluated under various wireless channel conditions, when there is only one station in the network. On the other hand, in multiple-user environment, several studies [9][10] reported that the performance of some types of RAAs, e.g. Automatic Rate Fallback (ARF)[1], degrades drastically because the RAA mistakenly lowers its data rate when the consecutive frame losses are caused by collision losses not by wireless losses.

There have been a few studies attempting to aid rate adaptation algorithms in dealing with the collision effects in multi-user environment. Their key idea is to provide RAAs the ability to "differentiate" between wireless losses and collision losses. For example, by assuming the only cause for the data frame transmission failure after a successful Request-To-Send/Clear-To-Send (RTS/CTS) exchange is due to channel error not collision, [9][10] use RTS/CTS to filter out collision losses from rate decision process. On the other hand, [10][11] suggest to add extra frames and fields to explicitly notify the sending station whether the transmission failure is due to collision or channel errors.

978-1-4244-2100-8/08/$25.00 ©2008 IEEE

While these proposals provide significant improvements compared to RAAs without loss differentiation capability, it is unclear whether loss differentiation is good enough to deal with all kinds of mixed wireless and collision loss scenarios. The fundamental problem is, as we will show later in this paper, that *background traffic from other contending stations changes the throughput ranking of the operating data rates*. In other words, under the same wireless condition, the data rate yielding the highest throughput in no background traffic scenarios is not necessarily the best one when background traffic exists. This is particularly problematic for existing loss differentiation schemes as they filter out all collision losses for RAA, the RAAs become insensitive to the throughput ranking changes caused jointly by wireless losses and collision losses, thus resulting in performance degradation.

In this paper, we design a new Background traffic aware Rate Adaptation Algorithm (BEWARE) that explicitly addresses the mixed effects from wireless and collision losses. Our contributions of this paper are: ***i***. we systematically evaluate the performance of RTS-based loss differentiation in different mixed wireless and collision losses scenarios. We identify when and why RTS-based loss differentiation does not work in certain scenarios, ***ii***. we use the insight in these systematic evaluations to identify a novel metric – *the expected packet transmission time* – to explicitly address the mixed effects from wireless and collision losses on all available data rates. We propose an online algorithm to estimate this parameter for all data rates and embed this information into the RAA design as the key rate decision maker, and ***iii***. we compare the performance of BEWARE with other RAAs with and without loss differentiation, and observe up to 250% and 25% performance improvement, respectively.

The rest of the paper is organized as follows. Section 2 reviews the existing RAAs and related loss differentiation approaches. Section 3 evaluates the performance of existing RAAs and loss differentiation schemes. Section 4 presents the design of our background traffic aware rate adaptation algorithm, and Section 5 evaluates its performance under various background traffic scenarios. Section 6 concludes.

2. Related work

In this section, we briefly review the existing rate adaptation algorithms (RAAs) and related loss differentiation schemes that help RAAs deal with collisions in multiple-user environment.

2.1 Existing Rate Adaptation Algorithms

There have been quite a few RAAs proposed by academia and industry. They can be broadly classified into three categories based on the information they collect for rate selection decisions: 1) statistics based RAAs, 2) received signal strength (RSS) based RAAs, and 3) hybrid RAAs.

1) Statistics-based RAAs: Based on the statistics the RAA uses for rate decisions, we can further categorize this class of RAAs into three different approaches. i) *Retry-based rate adaptation* [1][2] uses consecutive transmission successes/losses (e.g. 10 and 2 in ARF [1], respectively) as the indicator of good/bad wireless condition. However, despite its easy design, previous study [4] has shown that, due to randomness of the wireless loss behavior, there is very weak correlation between past consecutive transmission successes/losses and future channel condition. Thus, this approach tends to yield pessimistic rate estimations. ii) *Frame-Error-Rate(FER)-based rate adaptation* [3][4] calculates FER by the ratio of the number of received ACK frames to the number of transmitted frames. The RAA decreases and increases the operation data rate if FER exceeds some pre-determined thresholds. However, as wireless channels are so vulnerable to many factors such as multipath, channel fading, and obstructions, the major drawback for FER-based RAA is the inadequacy in using one set of pre-determined FER thresholds in all circumstances. iii) *Throughput-based rate adaptation* [5] calculates each data-rate's throughput based on the packet length, bit-rate, and the number of retries collected during a predefined decision window (~1 sec). The major drawback of this approach is that, as the decision window has to be large enough to collect meaningful statistics, it causes the rate adaptation algorithm to be less responsive to sudden wireless condition changes.

2) Signal-strength-based RAAs: This class of RAAs [6][7] relies on wireless signal strength information, such as Received Signal Strength Indicator (RSSI) or Signal-to-Noise Ratio (SNR), to make the rate adjustment decisions. The RAAs pick the data rate by a pre-determined mapping between the received signal strength and throughput. However, in reality, such mapping is highly variable and a model established before-hand may not be applicable to any environments later. Meanwhile, signal-strength-based RAAs have to overcome the communication issue of piggybacking the signal strength measurement taken at the receiver side to sender so that sender can adjust the data rate accordingly. One has to either use explicit signaling [7], which is incompatible to the IEEE 802.11 standard, or assume the channel is symmetric

[6], which is clearly not the case in real-world scenarios.

3) Hybrid RAAs: In this approach, RAA [8] collects both frame transmission statistics and received signal strength, and use statistics-based controller as the core rate adaptation engine. The rate decision can be overridden by signal strength based controller if it detects sudden changes in received wireless signal strength. As hybrid RAA design still assumes symmetric wireless channel and pre-established RSSI-to-rate thresholds, this approach is not immune from the drawbacks we discussed in signal-strength-based RAAs section.

In summary, all types of RAAs strive to obtain accurate channel estimations from different kinds of loss characteristics and decide when to decrease or increase the rate. However, in multiple-user environment, packet collisions incur new sources of frame losses. None of these RAAs explicitly address this issue. In the next section, we review several proposals that try to aid RAAs in dealing with collision effects.

2.2 Loss differentiation for rate adaptation

Previous studies reported that, because ARF treats collision losses no different than wireless losses, ARF excessively decreases its rate upon contention collisions, even when wireless channel is close to perfect. As this effect causes severe performance degradation for ARF when background traffic exists, we refer this effect as "rate poisoning". There have been two approaches to aid rate adaptation algorithms in differentiating wireless losses from collision losses. i) *Loss differentiation by RTS/CTS*: With RTS/CTS exchanges preceding data transmissions, [9] and [10] assume that the only cause for the data frame transmission failure after a successful RTS/CTS exchange is due to channel error not collision. Therefore, filtered by RTS/CTS, RAA rate decision process reacts only on wireless losses and is no longer affected by the collision effect. Kim et al. [9] further propose Collision-Aware Rate Adaptation (CARA) to reduce the extra RTS/CTS overhead by selectively turning on RTS/CTS after data frame transmissions fail at least once without RTS/CTS. The data rate is increased as the consecutive success count reaches 10, similar to ARF.[1] ii) *Loss differentiation by explicit notification:* [10] and [11] propose to add extra frames and fields to explicitly notify the sending station of the source of losses. However, both proposals require incompatible changes to the IEEE 802.11 standard and thus are not favorable for real-world deployments.

In summary, loss differentiation is the dominating approach for RAAs in dealing with collision effects when there are other stations transmitting traffic in the network. However, it is not clear whether loss differentiation is sufficient to guide RAAs to perform well in various multiple-user environments with mixed wireless and contention conditions. As we will show later in the paper, while RTS-based loss differentiation works in certain circumstances, we also find other scenarios that RTS-based loss differentiation performs poorly.

3. Performance of rate adaptation algorithms with background traffic

In this section, we first explain briefly how IEEE 802.11 rate adaptation works. In particular, we analyze how rate selection objective varies with the level of background traffic. Furthermore, we systematically evaluate the performance of various RAAs with RTS loss differentiation schemes under different scenarios, including varying number of stations in the network and the distance between stations and access point. As we will show in this section, it is critical to examine how and why these RAAs do not perform well with background traffic. By such an investigation, we not only better understand the necessity for a RAA that does take background traffic into consideration, but also gain insight into how to design such a RAA.

3.1. IEEE 802.11 rate adaptation with different level of background traffic

To visualize the throughput-distance tradeoff among multiple data rates employed by IEEE 802.11 standard, in Fig. 1, we use ns-2 [12] to simulate an 802.11a station's maximum throughput as it moves away from the access point (AP) in Ricean fading environment [13][2]. As seen in Fig. 1, among the 8 data rates available in IEEE 802.11a, higher data rates can achieve higher throughput, but their transmission ranges are shorter. The crossing points of two adjacent data rates indicate that, at a given location, the error rate of the high data rate is becoming too high that it's actually more favorable to use the next lower data rate to benefit from the lower error rate. Clearly, rate adaptation mechanism should try to follow such transitions as close as possible to select the best data rate according to the current wireless channel condition experienced by the link. Ideally, if a rate adaptation mechanism has perfect knowledge of the current network condition, its data rate selections follow

[1] In this study, we only consider the adaptive RTS/CTS mechanism (called CARA-1 in [9]), as the optional Channel Collision Assessment (CCA) detection, which is called CARA-2 in [9], only provides marginal performance gain over CARA-1.

[2] Refer to Sec. 5.1 for detailed simulation parameters.

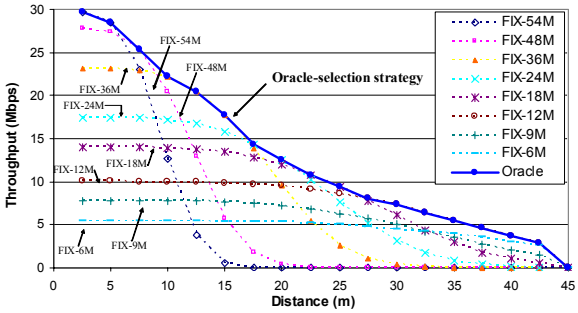


Figure 1. Throughput versus distance for IEEE 802.11a data

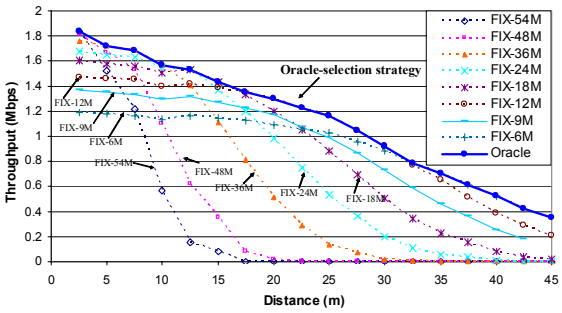


Figure 2. Throughput versus distance for IEEE 802.11a data rates, with 12 background traffic stations

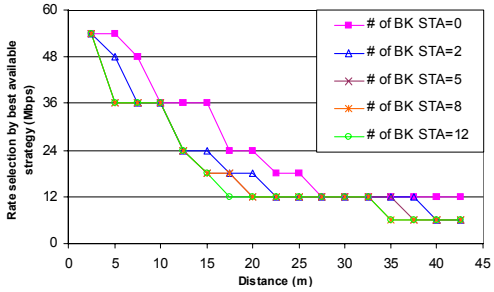


Figure 3. Oracle-selected data rate under different number of background traffic stations

closely with the outer envelope (plotted as thick solid line) of Fig 1. In this way, the throughput yielded by the rate adaptation mechanism is always maximized given a particular channel condition. We will refer to this outer envelope concept as the "*oracle-selection strategy"* and its performance as *maximum throughput* throughout the paper.

On the other hand, Fig. 2 plots the performance of the same data rate set under the same wireless channel condition, but with 12 other stations continuously transmitting background traffic in the network. We can see that, not only the shape of staircase like throughput-distance curves changes, but the rates selected by the *oracle-selection strategy* also change for the same location. It is because the data frames transmitted by any data rate are subject to not only wireless losses but also collision losses caused by medium contentions with other stations. In other words, the extra backoff time spent in medium contentions and collisions change the crossing points of two adjacent data rates, and thus the rate switching strategy. This combined effect also changes the performance ranking of data rates for a given location. Fig. 3 further illustrates this effect by plotting the rate selections by the *oracle-selection strategy* when operating with different number of saturated background traffic stations. The rate selected by the *oracle-selection strategy* varies widely with background traffic intensity. In other words, the rate adaptation strategy that works well in one background traffic scenario may not work in other background traffic scenarios, hence the rate adaptation mechanism needs to explicitly address this phenomenon. With this observation, we argue that ***it is very critical for rate adaptation designs to be aware of such changes and adjust their rate selection strategies to accommodate such changes; otherwise they will suffer from serious performance degradation***.

3.2. Performance of RAAs with RTS loss differentiation

Previous studies have reported superior performance of ARF with RTS on over that with RTS off. However, those studies did not provide systematic investigations into whether RAA with RTS differentiation really achieves the optimal throughput in all situations and why it does or does not. In this subsection, we compare the performance of RAAs with RTS differentiation to the *oracle-selection strategy*.

With the same simulation settings in the previous section, we first place all stations at 2.5m away from the AP and turn on RTS for all stations. We isolate the effects of RTS loss differentiation in performance comparisons by enabling only one station with RAA on, and all other background traffic stations with fixed data rate. In Fig. 4, we plot the RAA enabled station's throughput (normalized against the maximum throughput) as it moves away from the AP. We can observe from Fig. 4 that, when the RAA-enabled station is close to the access point, ARF with RTS (ARF-RTS) performs almost the same as the *oracle-selection strategy*, regardless of how many stations transmitting background traffic in the network. However, as the RAA-enabled station moves away from the access point, we can see that ARF-RTS starts to lose track of the best available rate. In addition, with more background traffic stations transmitting in the network, ARF-RTS station's performance deviates further from the maximum throughput. Finally, ARF-RTS station's performance approaches back to ~90%

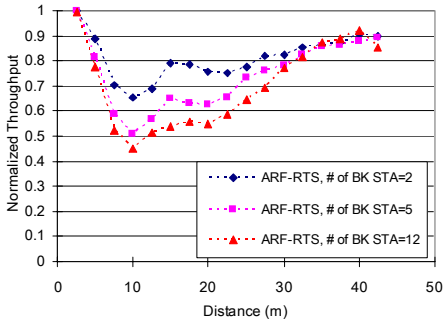


Figure 4. Normalized throughput for ARF-RTS, with various number of background traffic stations in RTS access mode

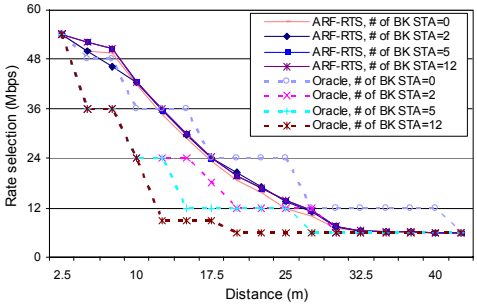


Figure 5. Data rate selection for ARF-RTS and *oracle-selection strategy*, with various number of background traffic stations in RTS access mode

of maximum throughput when the station is reaching the transmission edge (>35m) of the AP.

To further explain such scenario, we plot Fig. 5 to illustrate rate selection breakdowns of ARF-RTS along with that of *oracle-selection strategy*, as distance to access point increases. We can see that the rate selections of ARF-RTS remain almost the same regardless the number of background traffic stations in the network, as opposed to the rate selections of *oracle-selection strategy* that vary widely with background traffic level as we discussed above. This is because RTS frames isolate the wireless losses from collision losses and make the rate decisions solely on wireless losses. As a result, RAAs become insensitive to the throughput ranking changes caused jointly by wireless losses and collision losses, and persist in using the rate selections that is only suitable in no background traffic scenarios. On the other hand, when the RAA-enabled station is far away from AP, the only operable data rate is the lowest rate. In this case, ARF-RTS's rate selection coincides with the *oracle-selection strategy* and thus performs close to 100% of the maximum throughput. We further examine the rate selections of other statistics-based RAAs (e.g. ONOE [14], Sample-Rate [5], and RRAA-basic [4]) with RTS-on in background traffic scenarios, and find the same phenomenon exists. It follows that ***turning on RTS misleads RAAs into using rates only suitable for no-background-traffic in scenarios with background traffic, where these rates are not always suitable.*** As a result, RTS loss differentiation only works well when the rate selections are similar for all other background traffic scenarios.

In summary, we have made the following two important observations in this section: *i) The oracle-selection strategy varies significantly with the level of background traffic. We argue that any rate adaptation mechanism should be aware of such change at the presence of background traffic, or it will suffer from serious performance degradation. ii) We show that none of the existing RAAs we have investigated perform well in every background traffic scenario. We see that even RTS loss differentiation can hurt the performance in some situations.*

4. BEWARE Design

From the lessons we learn from previous section, we know that the key for RAA algorithm to perform well in background traffic scenarios is to incorporate not only wireless channel statistics but also background traffic conditions as the indicators in accessing the effectiveness of each available data rate. As a result, in this section, we present the design of BEWARE, a Background traffic aWaAre RatE adaptation algorithm for IEEE 802.11-based MAC. The center part to this design is to use a mathematical model to calculate the expected packet transmission time of each data rate that attributes the combined costs of wireless channel errors and background traffic contentions as we discussed in Sec. 3. The rate selection engine then uses this metric to find the data rate that yields the highest throughput in the given wireless channel and background traffic condition. The goals to design such rate selection strategy are two-fold: it has to be robust against any degree of background traffic; meanwhile, it is also responsive to random and even drastic wireless channel changes.

Although using the *expected-packet-transmission-time* as rate selection metric may seem at first as straightforward, it became clear only after our thorough and systematic investigations (in the previous section) on how and why various existing RAAs do not perform well with background traffic. In addition, this concept is novel as no existing studies, to the best of our knowledge, have used such a rigorous metric in RAA design. Next, we describe the mathematical model for expected packet transmission time calculations in Sec. 4.1, then the rate selection engine in Sec. 4.2.

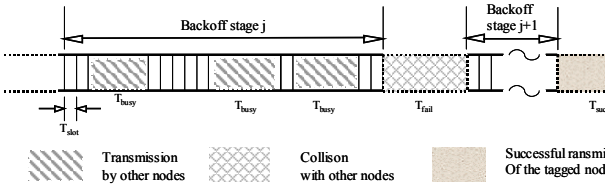


Figure 6. Packet transmission and collision events during IEEE 802.11 MAC backoff

4.1 Packet transmission time estimation

The core of BEWARE design is the estimation for expected packet transmission time of each data rate, with the consideration of mixed effects from wireless channel condition and collisions. In CSMA/CA-based 802.11 MAC, the overall time duration required to complete a packet transmission is dictated by the backoff procedure. Therefore, as shown in Fig. 6, we calculate expected packet transmission time by carefully analyzing the duration and occurring probability of different events take place at backoff stages, as follows.

a) When the backoff timer decrements, the time slot is either sensed as idle (for T_{slot}, the length of one time slot) or as busy occupied by background traffic transmission (for T_{busy}, the average medium occupation time used by background traffic transmissions). We define P_{busy} be the probability that, at a given time slot, the backoff timer is frozen due to busy medium in carrier sensing. It follows that the occurring probability of idle slot and busy slot is (1- P_{busy}) and P_{busy}, respectively.

b) When the backoff timer expires (i.e. decrements to zero), the attempt of packet transmission either fails (after T_{fail}) or succeeds (after T_{succ}). We define P_{fail} to be the frame error probability. It follows that the occurring probability of packet failure and success is P_{fail} and (1- P_{fail}), respectively. Note that P_{fail} represents the transmission failure events due to various packet failure sources such as channel fading, interference, and hidden terminals.

Once these parameters are collected, we can construct a mathematical model calculating the occurring probability for combinations of all different backoff events throughout all backoff stages. We first define the occurring probability $F^{j}_{k,n-k}$ that, in any single backoff stage j with backoff timer selected from 0 to W_j (maximum number of backoff slots in stage j), there are exactly k busy time slots and (n-k) idle slots:

$$F^{j}_{k,n-k} = \frac{1}{W_j} C^{i+k}_{k} P_{busy}{}^{k} (1-P_{busy})^{n-k} \quad , \quad 0 \le k \le n \le W_j. \qquad (1)$$

Moreover, we know that any combination of number of busy and idle slots can be a cumulative effect from successive backoff stages. Therefore, we then define $S^{j}_{k,n-k}$ for probability of backoff counter being frozen (k-j) times and idle (n-k) times that up to back off stage j (which implies packet transmission failed j times),

$$S^{j}_{k,n-k} = \begin{cases} (1-P_{fail})F^{0}_{k,n-k} \\ \quad , for \quad j=0, \qquad 0 \le k \le n \le W_0 - 1. \\ (1-P_{fail})P_{fail} \sum_{l=0}^{k} \left(\sum_{i=0}^{n-k} S^{j-1}_{l,i} \times F^{j}_{k-l,n-k-i} \right) \\ \quad , for \quad 1 \le j \le m-2, \quad j \le k \le n \le \sum_{i=0}^{j} W_i - 1. \\ P_{fail} \sum_{l=0}^{k} \left(\sum_{i=0}^{n-k} S^{j-1}_{l,i} \times F^{j}_{k-l,n-k-i} \right) \\ \quad , for \quad j=m-1, \quad m-1 \le k \le n \le \sum_{i=0}^{m-1} W_i - 1. \end{cases} \qquad (2)$$

Note that m in this equation is the maximum number of retries specified in the standard.

$S^{j}_{k,n-k}$ includes all possible cases, from combination of previous stage(s) to the current stage, which result in (n-k) idle slots, (k-j) busy slots, and j failed transmission periods. In addition, the time points when such combinations happen can be characterized by,

$$T^{j}_{k,n-k} = (k-j)*T_{busy} + j*T_{fail} + (n-k)*T_{slot} + T_{succ}$$
$$for \quad 0 \le j \le m-1, \quad j \le k \le n \le \sum_{i=0}^{j} W_i - 1 \qquad (3)$$

As a result, the expected packet transmission time can be expressed by

$$T_{avg} = \sum_{k=0}^{N} \sum_{n=k}^{N} \sum_{j=0}^{m-1} (S^{j}_{k,n-k} * T^{j}_{k,n-k})$$
$$, \quad where \quad N = \sum_{i=0}^{m-1} (W_i - 1). \qquad (4)$$

Once the expected packet transmission time is obtained, it is sent to the rate selection module for rate selection decisions as we describe in the next subsection. We note that the detailed derivations and evaluations of the accuracy of such model can be found in our previous work [17]. In addition, we know that the throughput of the tagged node is a direct function of the expected packet transmission time. Therefore, we can expect the above mechanism to be a good metric for rate selection decisions.

On the other hand, we can see that the average packet transmission time is a function of several parameters from the environment, i.e. P_{busy} P_{fail}, and T_{busy}. In the following, we show how our model captures the mixed effects from background traffic and

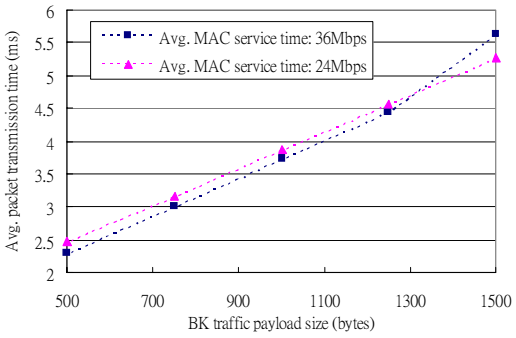


Figure 7. Average packet transmission time of two adjacent data rates when changing T_{busy}.

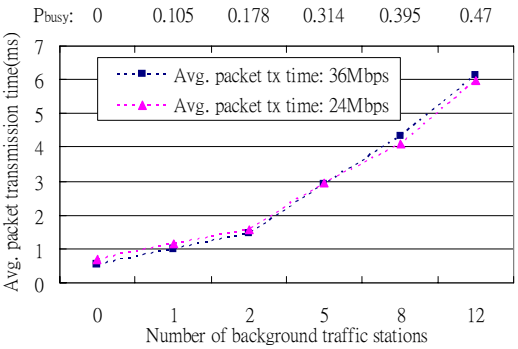


Figure 8. Average packet transmission time of two adjacent data rates when changing P_{busy}.

wireless channel losses and make rate switching decisions accordingly between two adjacent data rates.

In Fig. 7, we plot the average packet transmission time of two adjacent data rates in IEEE 802.11a standard, 36Mbps and 24Mbps, with changing the background traffic stations payload sizes but keeping all other parameters fixed. We can see that average packet transmission time of 36Mbps becomes larger than that of 24Mbps after background traffic payload size increases to more than 1300bytes. Recall from the derivations above, change in background traffic payload size corresponds to change in T_{busy}, and in turn the length of busy medium slots in backoff stages. In other words, as T_{busy} increases, busy medium slots become longer in each backoff stage, and consequently the backoff stage length is longer. As a result, as the higher data rates are more vulnerable to have more backoff stages due to high wireless loss rates (P_{fail}), longer backoff stages cause the expected packet transmission time of the higher data rate to grow faster than that of the lower data rate. Therefore, there is a crossing point where the expected packet transmission time of 36Mbps becomes higher than that of 24Mbps as T_{busy} increases. It follows that beyond such point, the

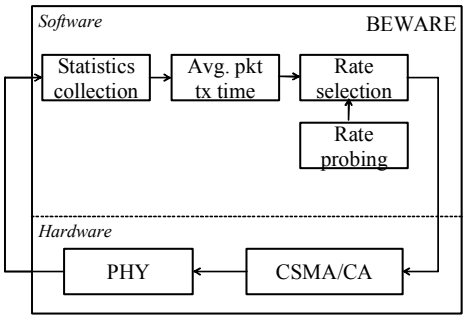


Figure 9. Structure of BEWARE design

performance for 24Mbps packets is better than 36Mbps packets even the wireless conditions for both data rates remain unchanged. It is essential to note that, for RAAs that only consider wireless loss effects in rate decisions, they can not capture the above performance crossing point caused by background traffic changes and make rate switching decisions accordingly.

Similarly, Fig. 8 plots the average packet transmission time of 36Mbps and 24Mbps with the scenarios that there are different numbers of background traffic stations in the network. Note that all other parameters including wireless loss conditions remain unchanged in these scenarios. We can also see a crossing point where average packet transmission time of 36Mbps becomes larger than that of 24Mbps when there are more than 5 background traffic stations in the network. The reason is similar to what we discuss above. More background traffic stations correspond to larger P_{busy}, which cause more number of busy medium slots in each backoff stage, and in turn longer overall backoff duration. This also explains the data rate performance ranking changes in different background traffic scenarios that we observe in Sec. 3.

4.2. Rate selection engine

In this section, we describe the high-level design of how BEWARE makes rate selection decisions. As shown in Fig. 9, the BEWARE design can be broken down into the following tasks:

1) Statistics collection/processing: After the packet transmission completes, transmission environment statistics, including T_{busy}, P_{busy}, and P_{fail}, are collected and processed by exponentially weighted moving average (EWMA) to smooth out the biases to the sudden changes in current wireless channel and collision conditions. In addition, BEWARE keeps track other statistics such as number of successful/failed packets of different data rates.

2) Expected packet transmission time calculation: With the environmental parameters collected in the above module, this module use the mathematical model described in previous subsection

to calculate the expected packet transmission time. The resultant expected packet transmission time are updated with recent history values by EWMA and fed into rate selection module for processing.

3) Rate probing: Periodically, BEWARE sends packets at a data rate other than the current one to update the expected transmission time of other data rates. In order to avoid the common rate-probing pitfalls reported in [4], BEWARE adopts various measures to ensure probing other data rates is not done very often and the cost is not too high. That is, BEWARE limits the frequency of packet probing to a fraction (~5%) of the total transmission time. BEWARE also limits the number of retries allowed for probing packets to 2 to save costly waiting time for unsuccessful probing. In addition, BEWARE does not probe data rates that suffer from excessive failures for most recent packet attempts (4 recent successive packets have been unacknowledged), and those whose expected transmission time with no background traffic already exceed the expected transmission time of current operating data rate.

4) Rate selection decisions: The rate selection module constantly compares the expected packet transmission time of current data rate and that of others, and decides to change operating data rate whenever it finds a data rate yields the shorter transmission time (and thus highest throughput). BEWARE also implements a short-term frame loss reaction mechanism in case wireless channel conditions change too rapidly. That is, the rate selection module forces data rate to decrease one level when the packets exhaust all retries for three times consecutively.

4.3. Discussion

We note that most of the parameters required for calculating expected transmission time can be directly obtained from passive channel activity monitoring, which does not incur any extra overhead. To be more specific, we can determine P_{fail} by counting the ratio of failed packet transmission attempts and total packet transmission attempts. We also obtain P_{busy}/T_{busy} by keeping track of the number/duration of experienced busy medium slots, respectively. On the other hand, T_{fail} and T_{succ} are directly determined by the operating data rate and T_{slot} is specified in different version of IEEE 802.11 standard. In practice, it may be difficult to obtain some of these parameters accurately due to implementation complexity in real devices. We can consider alternative approaches [15][16] by using number of consecutive idle slots between two busy slots to estimate P_{busy} and P_{fail}.

On the other hand, one may argue that the expected packet transmission time can be directly obtained by keeping track of the medium access time of every packet without involving time-slot level channel monitoring. While this approach has been proposed in multi-hop wireless mesh network routing studies [18][19], it may not be suitable for MAC layer rate adaptation decisions for the sampling granularity it provides. In other words, the statistics averaged through potentially dozens of packets in the past may not be able to provide the most up-to-date channel information for the rate adaptation decisions that are made in per-packet basis. In addition, one may argue that collecting time-slot level statistics (i.e. P_{busy} and T_{busy}) might prevent the stations from going into sleep mode, which is critical for energy savings. We can optimize the energy consumptions by collecting the statistics only when the station has packets to send. Exploring the tradeoff between energy savings and collecting most up-to-date statistics is one of the topics in our future work.

5. Performance evaluation

In this section, we use ns-2 [12] to evaluate the performance of BEWARE and other RTS-based loss differentiation RAAs, including ARF with RTS/CTS (as referred to ARF-RTS) and CARA-1 under various mixed wireless and background traffic scenarios.

5.1. Simulation setup

We enhance the ns-2 simulator to support 802.11a Physical layer (PHY) and port various RAAs from previous studies or the real-world driver implementations [14]. We simulate scenarios in an infrastructure-based network, which contains one Access Point (AP) and a number of static wireless stations spreading in the network. We consider realistic wireless channel conditions by using and Ricean fading model with parameter K=6, and environment maximum velocity v=10m/s. The traffic sources are UDP flows unless stated otherwise.

5.2. Performance of single station with varying distance

We first focus on RAAs' performance with varying distance under background traffic scenarios. We place 2~12 stations on a circle around the AP within 2.5 meter radius, and all stations transmit UDP background traffic with RTS access mode. The transmission data rate of background traffic stations is locked at 54Mbps because of their proximity to AP. We then add one RAA-enable station in the network and measure the RAA's performance by varying the distance between RAA-enable station and AP. We show results with 12

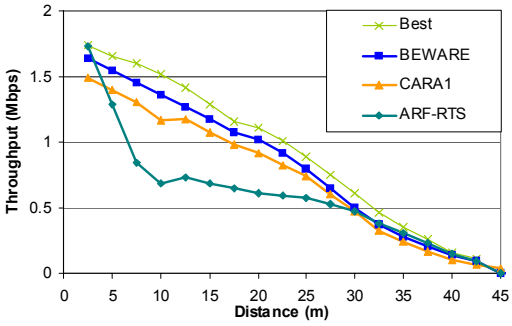


Figure 10. Throughput comparison for Best (oracle-selection strategy), BEWARE, CARA1, and ARF with RTS/CTS, with 12 background traffic stations in RTS access mode

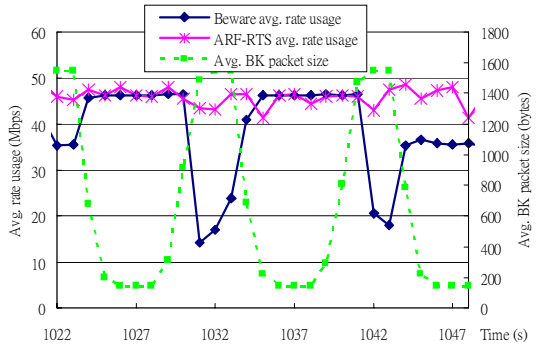


Figure 11. Data rate selection for BEWARE and ARF-RTS with dynamically changing background traffic payload size

stations transmitting background traffic as an example in Fig. 10, while results with other number of background traffic stations show similar trend. In all cases, the performance of BEWARE follows closely to what is offered by the *oracle-selection strategy* by only 10% less in throughput, and the performance of CARA-1 trails behind BEWARE by another 10%-15%. On the other hand, similar to what we discuss in Section 3.2., the performance of ARF-RTS significantly derails from the *oracle-selection strategy* when the distance from station to AP is close-by to intermediate (5m~35m). It is because, in this range, the rate selections for no background traffic deviate significantly from the rate selections for this background traffic scenario. As we discussed in Section 3, ARF with RTS loss differentiation suffers from performance degradations by continuing to use the rate selections only suitable for no background traffic.

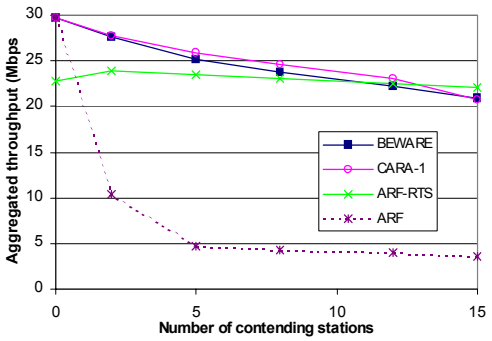


Figure 12. Aggregate throughput comparison for BEWARE, CARA1, ARF-RTS, and ARF in close-by topology with various number of contending stations

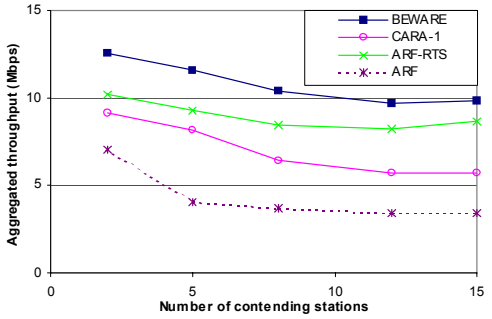


Figure 13. Aggregate throughput comparison for BEWARE, CARA1, ARF-RTS, and ARF in random topology with various number of contending stations

5.3. Performance of single station with dynamically changing background traffic

In this subsection, we further investigate how different RAAs adapt with dynamically changing background traffic levels. We place 12 background traffic stations randomly scattered in the network and always use the lowest transmission rate (6Mbps) to guarantee high packet delivery rate. We synchronize the traffic patterns of the background traffic stations so that, for every 3~5 seconds, they all change packet payload size around the same time. We then add one RAA-enabled station in the network and measure the RAA's performance. We can see from Fig. 11, as the average packet size of background traffic changes, the average data rate used by ARF-RTS does not show notable changes. On the other hand, we can see that BEWARE tries to adapt its rate selections as background traffic packet size changes. Recall from the discussions in Sec. 4, as the higher data rates are more vulnerable to have more backoff stages due to high wireless loss rates (P_{fail}), the longer backoff stages caused by increased background traffic payload sizes make the higher data rates less favorable to operate in situations with large background traffic payload size.

Therefore, we can see that BEWARE tries to adapt to the lower data rates when it senses such changes. On the other hand, when the background traffic payload sizes are small, BEWARE tries to use the highest possible data rate for optimal performance. We observe that BEWARE outperforms ARF-RTS for ~20% in throughput in this dynamically changing background traffic environment. We further investigate the effects of other background traffic changing patterns, such as increasing the frequency of background traffic payload size fluctuations and changing the number of simultaneously transmitting stations, and we observe that BEWARE consistently outperforms ARF-RTS for 25%-50% in various dynamic changing background traffic scenarios.

5.4. Aggregated performance with different topology

We now evaluate aggregate performance when all stations turn on RAA and operate with the same RAA homogeneously. We first simulate a topology with minimum wireless losses, in which various numbers of stations are uniformly placed at 2.5m away from AP and each station transmits fixed size 1500-byte long UDP traffic. As shown in Fig. 12, ARF's aggregate per-formance degrades severely due to the "rate poisoning" effect we discussed in Sec. 3. On the other hand, with the help from RTS loss differentiation, ARF-RTS performs well for any number of contending stations. Furthermore, BEWARE and CARA-1 perform closely and both outperform ARF-RTS in most cases, thanks to the overhead reduction design in CARA-1 and accurate background traffic effect estimation in BEWARE.

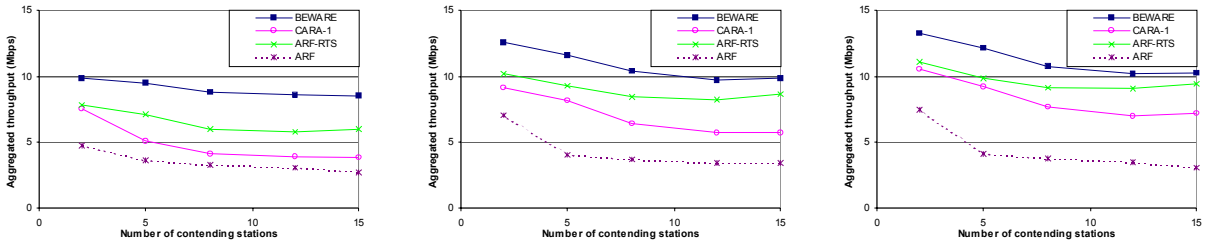
Secondly, we simulate a random topology with various numbers of stations randomly scattered in the network with maximum distance 45m away from AP to guarantee no hidden terminals. Each station transmits UDP traffic with random size. As shown in Fig. 13, the performance ranking differs from what we observe in Fig. 12. While ARF still suffers from rate poisoning and performs the worst, CARA-1 no longer outperforms ARF-RTS and ranks second from the worst. It is because, as nodes spreading at different distance to AP, both wireless loss and contention losses affect the first without-RTS data frame transmissions specified in CARA-1 algorithm, which cause CARA-1 stations de-crease data rate over aggressively. On the other hand, BEWARE still performs the best in random topology. On average, BEWARE outperforms ARF by 200%-250% and ARF-RTS, the best proposed by previous studies, by 20%-25% in aggregate performance.

5.5. Aggregated performance under various channel fading condition

We now compare the performance of different RAAs under various channel fading conditions. We vary the Ricean parameter K and Doppler spread f_m. Note that, as K increases, the line-of-sight component is stronger and the overall channel SNR increases. On the other hand, as f_m increases, the channel condition changes more rapidly. Fig 14 plots the aggregate performance of different RAAs under different K in a random topology similar to what we used in previous sub-section. We can see that, as K increases, the overall throughput of all RAAs increases as expected. However, the ranking of RAA performance remains unchanged. BEWARE outperforms ARF-RTS, CARA-1, and ARF under all different K parameters we studied. We then plot Fig 15 with the aggregate performance of different RAAs under different Doppler spread. We can see that, as f_m decreases, BEWARE still outperforms ARF-RTS in most cases, but the performance gap between BEWARE and ARF-RTS closes. To be more specific, as BEWARE outperforms ARF-RTS by 25% when f_m =17Hz, this advantage decreases to 5% when f_m decreases to 3.5Hz. Previous studies [11][20] reported that, as ARF is designed to increase its rate after several consecutive packet successes, ARF-based RAA tends to yield higher throughput by taking advantage of the slower changing channel environment. However, the performance of ARF degrades when the wireless channel condition changes rapidly. On the other hand, we can see that, as BEWARE yields comparable performance in different f_m environments, BEWARE is robust to both fast-changing and slow-changing wireless channel conditions.

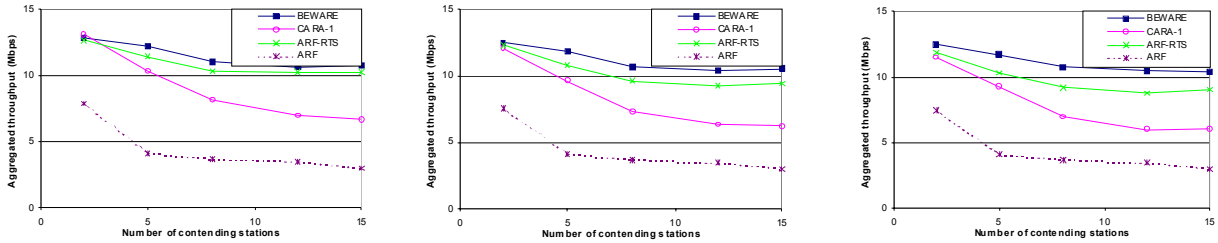
5.6. Performance with heterogeneous RAA deployments

As rate adaptation is an option that is left open for wireless card vendors to implement, it is not uncommon that there are stations equipped with different RAAs in real world scenarios. Therefore, it is essential to evaluate the performance of different RAAs in heterogeneous scenarios. In this experiment, we evaluate how different RAAs improve the individual and aggregate performance with a gradual upgrade deployment. We consider a network with 12 stations randomly placed within the transmission range of the AP, and transmit UDP traffic with random size. We start with the baseline scenario where all stations operate with ARF without RTS/CTS, in which all stations operate at the lowest data rate due to "rate poisoning" problem. We then gradually upgrade a number of stations with BEWARE or ARF-RTS, and



(a) *K=0* (b) *K=6* (c) *K=12*

Figure 14. Aggregate throughput comparison for BEWARE, CARA1, ARF-RTS, and ARF in random topology under different Ricean Parameter *K*



(a) f_m=3.5HZ (v= 0.2 m/s) (b) f_m=10HZ (v= 0.6 m/s) (c) f_m=17HZ (v= 1 m/s)

Figure 15. Aggregate throughput comparison for BEWARE, CARA1, ARF-RTS, and ARF in random topology under different Doppler Spread f_m

evaluate the aggregate performance improvement over baseline scenario and individual performance improvement of the same station after upgrade. We can see from Fig. 16 that, as the aggregate performance of ARF-RTS improves when upgraded stations added to the network, the individual performance of ARF-RTS actually decreases when less than half of the stations in the network are upgraded. When there are just a few stations upgraded with ARF-RTS, individual performance of upgraded stations decrease due to excessive use of higher data rates as we discuss in Sec. 3.2. Meanwhile, aggregate performance increases as other stations take advantage of the excess loss transmission opportunities incurred by upgraded stations. On the other hand, when there are more and more stations upgraded with ARF-RTS, ARF-RTS stations mutually take advantage of other upgraded stations' loss transmission opportunities, and collectively result in higher aggregate throughput even the rate selections made by these stations are not the most suitable ones for the corresponding scenario. By contrast, both individual and aggregate performance of BEWARE start to improve when just 1 station is upgraded. In addition, as the stations upgraded with BEWARE start to use data rates that is appropriate for the given wireless and collision conditions, other stations benefit from the extra free transmission time spared by BEWARE stations, and thus yields higher throughput even they are not upgraded with BEWARE. Note that this is an essential feature that, when incorporating any new algorithm to interoperate with other existing algorithms, the new algorithm should not hurt the performance of other existing algorithms.

In summary, with the homogeneous and heterogeneous background traffic scenarios we evaluate in this section, we observe that, while the effectiveness of RTS-based loss differentiation RAAs differ in different scenarios, BEWARE always yields the best performance for most cases. In addition, even with only one station equipped with BEWARE in the network, both individual performance of BEWARE and aggregate network performance improve over the rate-poisoned all-ARF network.

6. Conclusion

In this paper, we first identify that data rate selection strategies of 802.11-based stations should accommodate the different rate selection criterions in

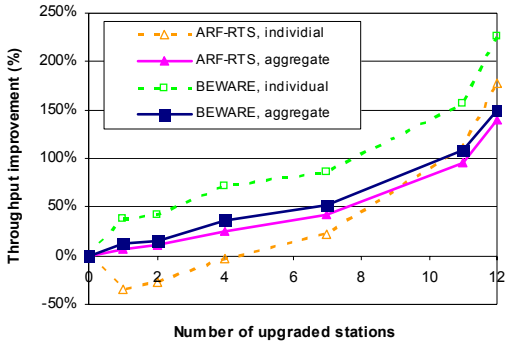


Figure 16. Individual and Aggregate throughput improvement of BEWARE and ARF-RTS with various number of contending stations in heterogeneous deployments

different background traffic scenarios. This observation further helps us explain why RTS-based loss differentiation schemes, which are proposed by previous studies to aid rate adaptation algorithms in dealing with collision effects, do not perform well in certain scenarios. In particular, RTS-based loss differentiation hurts the performance by persistently using the same rate selections regardless of background traffic level. Therefore, these observations motivate us to design a rate adaptation algorithm that explicitly addresses wireless and contention factors in its design.

We propose a novel background traffic-aware rate adaptation, BEWARE, that uses an accurate mathematical model to estimate the effectiveness of the data rates in given wireless and contention conditions. We show that the rate selections of BEWARE are close to what are selected by the *oracle-selection strategy* that has global knowledge of network conditions. We also show that, compared to other RTS-based loss differentiation schemes, BEWARE yields the best performance in the scenarios we have investigated in the paper.

As a work-in-progress, we are working on implementing BEWARE into the real 802.11a wireless card driver. Meanwhile, we also plan to investigate the interactions between rate adaptation algorithms and upper-layer protocols such as TCP. We believe that, as the design of BEWARE fully addresses the wireless and contention factors in MAC layer, it should render the best performance when integrated with upper-layer protocols.

7. References

[1] A. Kamerman, L. Monteban, "WaveLAN-II: A High-Performance Wireless LAN for the Unlicensed Band," Bell Labs Technical Journal, pp. 118–133, Summer '97

[2] M. Lacage, M. H. Manshaei, T. Turletti, "IEEE 802.11 rate adaptation: a practical approach," ACM MSWiM '04

[3] B.E. Braswell, J.C. McEachen, "Modeling Data Rate Agility in the IEEE 802.11a WLAN Protocol," OPNETWORK '01.

[4] S. Wong, H. Yang, S. Lu, V. Bharghavan, "Robust Rate Adaptation in 802.11 Wireless Networks," ACM MOBICOM '06.

[5] J. Bicket. Bit-rate Selection in Wireless Networks. MIT Master's Thesis, '05.

[6] J. P. Pavon and S. Choi, "Link Adaptation Strategy for IEEE 802.11 WLAN via Received Signal Strength Measurement," IEEE ICC'03

[7] Gavin Holland, Nitin Vaidya, and Paramvir Bahl, "A Rate-Adaptive MAC Protocol for Multi-Hop Wireless Networks," ACM MobiCom '01

[8] I. Hatatcherev, K. Langendoen, R. Lagendijk, H. Sips, "Hybrid rate control for IEEE 802.11," ACM MobiWac '04

[9] J.S. Kim, S.K. Kim, S.H. Choi, D. Qiao, "CARA: collision-aware rate adaptation for IEEE 802.11 WLANs," IEEE INFOCOM '06.

[10] Q. Pang, V.C.M. Leung, and S.C. Liew, "A Rate Adaptation Algorithm for IEEE 802.11 WLANs Based on MAC-Layer Loss Differentiation," IEEE BRAODNETS '05.

[11] J.H. Yun and S.W. Seo, "Novel collision detection scheme and its applications for IEEE 802.11 wireless LANs," Computer Communications Volume 30, Issue 6, March '07.

[12] "The Network Simulator – ns-2," http://www.isi.edu/nsnam/ns/

[13] R. J. Punnoose, P. V. Nikitin, and D. D. Stancil, "Efficient Simulation of Ricean Fading within a Packet Simulator," IEEE VTC'00-Fall.

[14] MADWIFI, http://sourceforge.net/projects/madwifi/

[15] M. Heusse, F. Rousseau, R. Guillier, and A. Duda, "Idle Sense: An Optimal Access Method for High Throughput and Fairness in Rate Diverse Wireless LANs," ACM SIGCOMM '05

[16] C. Hu and J. Hou, "A Novel Approach to Contention Control in IEEE 802.11e-Operated WLANs," IEEE INFOCOM '07.

[17] S.C. Wang and A. Helmy, "Performance Limits and Analysis of Contention-based IEEE 802.11 MAC," IEEE LCN '06.

[18] R. Draves, J. Padhye, and B. Zill. "Routing in Multi-Radio, Multi-Hop, Wireless Mesh Networks," ACM MobiCom '04.

[19] T. Salonidis, M. Garetto, A. Saha, and E. Knightly, "Identifying High Throughput Paths in 802.11 Mesh Networks: a Model-based Approach," IEEE ICNP '06.

[20] D. Qiao and S. Choi, "Fast-Responsive Link Adaptation for IEEE 802.11 WLANs," IEEE ICC'05.