The 17th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'06)

ANALYSIS OF THE RTS/CTS MULTIPLE ACCESS SCHEME WITH CAPTURE EFFECT

Chin Keong Ho †§
†Eindhoven University of Technology
Eindhoven, The Netherlands

Jean-Paul M. G. Linnartz §†
§Philips Research Laboratories
Eindhoven, The Netherlands

ABSTRACT

The Request-to-Send (RTS) / Clear-to-Send (CTS) scheme uses a four-way handshake to ensure proper spatial and temporal reservation of the multiple-access wireless LAN channel. By considering the (interdependence of) successive capture probabilities for a cycle of an RTS, CTS, data packet and acknowledgement, we analyze the throughput of the RTS/CTS scheme in a new mathematical framework. We also propose closed-form bounds which are useful for rate optimization.

I. INTRODUCTION

The performance of today's wireless LANs is increasingly limited by mutual interference. Several mechanisms are used to allow co-existence, such as *carrier sensing* and *virtual sensing*. Carrier sensing is a multiple access scheme that inhibits transmission when an ongoing transmission is detected at the transmitter [1]. However, in some situations users are inhibited unnecessarily [2]. This could be solved by virtual sensing in which the source first asks the destination whether the channel is idle by a Request-to-Send (RTS) packet, and the destination confirms this with a Clear-to-Send (CTS) packet [2]. Remote stations that recover the RTS or CTS packet are inhibited to transmit during some specified time, hence increasing the probability of a successful transmission.

Two types of model for packet recovery have been assumed when considering the RTS/CTS scheme. In the collision model, a packet is correctly received, i.e., *recovered*, if and only if there is no other concurrent transmission, e.g. [3]. A improved model is based on the power capture effect: a packet is recovered (even if there are interfering signals) if and only if the signal-to-noise plus interference ratio (SINR) is greater than a capture ratio, e.g. [4, 5]. In the literature [3, 4, 5], the RTS/CTS scheme has been modeled by assuming that if the RTS is recovered by the destination, then the CTS, payload (PAY) and acknowledgement (ACK) will be recovered too. This is in fact an approximation that violates the principle of the capture effect. By removing this approximation, the causal effect of the RTS, CTS, PAY and ACK in time is analyzed in [6].

This paper extends the results of [6] by calculating the throughput achieved using the RTS/CTS scheme. By choosing appropriate RTS, CTS and PAY transmission rates, we illustrate by numerical examples how throughput maximization can be carried out. Moreover, by deriving closed-form throughput lower bounds, it is shown quantitatively that the RTS/CTS scheme is superior to the ALOHA scheme when the PAY length is sufficiently long (such that that the overheads are small) and when the RTS and CTS rates are sufficiently small (such that the interference are low).

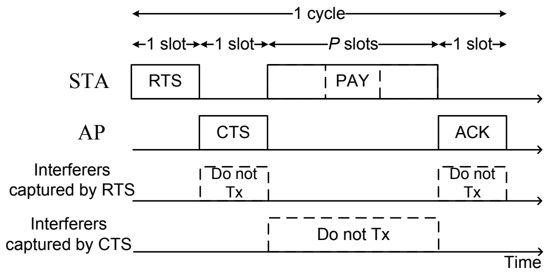

Figure 1: RTS/CTS multiple access scheme.

II. MODEL

A. RTS/CTS Scheme

The test station (STA) will be referred to simply as STA and the other STAs as interferers. We assume that STA uses RTS/CTS for channel access to the intended destination which we assume to be the access point, AP. On the other hand, the interferers use slotted ALOHA for multiple access.

The detailed RTS/CTS scheme considered here is shown in Fig. 1. A transmission cycle consists of an RTS packet, a CTS packet, a PAY (data) packet and an ACK packet; for conciseness the word "packet" is subsequently omitted when referring to the different packet types. For ease of analysis, we assume the system to be time slotted. Each RTS, CTS and ACK uses one slot, and the PAY uses P slots. As illustrated, if the interferers recover the RTS (CTS), they are inhibited to transmit during the CTS and ACK periods (PAY period, respectivly). In effect, the inhibition period is kept to a minimum; this is in contrast to the practice of IEEE 802.11, where any inhibition is carried out continuously until the cycle completes. Furthermore, we assume:

1) *No carrier sensing:* In contrast to common practice in IEEE 802.11, we explicitly do not consider that carrier sensing is carried out, but only virtual sensing via the RTS/CTS mechanism. This eliminates the *exposed node problem* associated with carrier sensing.

2) *Only* STA *is fully RTS/CTS capable:* For analytical tractability, we assume that only STA uses RTS/CTS for multiple access. The interferers uses slotted ALOHA. Yet, they obey the RTS/CTS rules, that is, they temporarily refrain from transmission after they recover any RTS or CTS.

3) *Slotted Data Detection:* Data is (transmitted and) detected during PAY on a slot by slot basis, independently of other slots. As a result, the ACK comprises of multiple acknowledgement bits, one bit for each slot of data. This assumption allows the capture model to be applied for each slot independently.

B. Capture, Channel and Network Models

The instantaneous SINR can be represented as

$$\mathsf{SINR} = \frac{p_0}{N_o + \sum_{i=1}^{N} p_i}, \quad (1)$$

where p_0, p_i is the instantaneous signal power of STA and the i^{th} interferer, respectively, while N_o is the noise power. From information theory, when the interference and the data is independent Gaussian distributed, an achievable instantaneous data rate, in bits/symbol, is given by the mutual information $\mathcal{I}(\mathsf{SINR}) = \log(1 + \mathsf{SINR})$. An *information outage* is said to occur if the instantaneous mutual information is smaller than R, i.e., when $\mathsf{SINR} < 2^R - 1$.

The capture effect, channel and network are modeled as follows.

1) *Capture Effect:* The capture effect is modeled by assuming that the packet is correctly decoded if and only if there is no information outage, hence relating the capture ratio to the rate as $z(R) \triangleq 2^R - 1$.

2) *Path Loss:* All transmitters use the same power. However, each signal experiences path loss depending on the propagation distance, a. We model the local mean power according to the path loss law $\bar{p} = a^{-\beta}$ where $\beta = 4$ is used in this paper.

3) *Channel:* Each channel exhibits quasi-static flat Rayleigh fading during one RTS, CTS, PAY or ACK transmission, but independent for different transmission links and slots. Hence, the probability density function (pdf) of the power p is $f_p(p) = \frac{1}{\bar{p}} \exp\left(-\frac{p}{\bar{p}}\right)$.

4) *Interferers' Positions:* The x- and y-coordinates of the STA and AP are denoted by position vectors $\mathbf{a}_{\mathsf{STA}}, \mathbf{a}_{\mathsf{AP}}$, respectively. When the RTS/CTS scheme is not implemented, interfering packets are transmitted according to a homogenous (spatial) Poisson process with intensity $\mathcal{G}(\mathbf{a}) = G_o$ packets per time slot per unit area, where $0 < G_o < \infty$ and $\mathbf{a}$ is in an operating region $\mathcal{A}$. The total traffic rate is $G_t = \int_{\mathbf{a}\in\mathcal{A}} \mathcal{G}(\mathbf{a})\, d\mathbf{a}$. In this paper, we let the region $\mathcal{A}$ grows infinitely large.

III. Capture Probability

Consider a source at $\mathbf{a}_{\text{source}}$ transmitting a packet of length L and rate R to a destination at $\mathbf{a}_{\text{dest}}$. The traffic intensity is $\mathcal{G}(\mathbf{a})$. The event that the packet is recovered, i.e., that the packet *captures* the receiver, is denoted as $\mathcal{E}_{\text{cap}}(\mathbf{a}_{\text{source}}, \mathbf{a}_{\text{dest}}, R, L, \mathcal{G})$. The arguments of z and $\mathcal{E}_{\text{cap}}$ will be dropped when there is no ambiguity.

A. General Capture Probability

Let f_{p_i} denote the (exponential) pdf of the power of the i^{th} interferer and $\mathcal{L}_f(s)$ the Laplace transform of function f evaluated at s. Define

$$W(2^R-1, |\mathbf{a}_{\text{source}}-\mathbf{a}_{\text{dest}}|, |\mathbf{a}_i-\mathbf{a}_{\text{dest}}|) = 1-\mathcal{L}_{f_{p_i}}((2^R-1)/\bar{p}).$$

Then, the *capture probability* can be simplified as [7]

$$\Pr\{\mathcal{E}_{\text{cap}}\} = \exp\left\{-\frac{zN_o}{\bar{p}} - \int_{\mathbf{a}_i\in\mathcal{A}} \mathcal{J}(\mathbf{a}_i)\, d\mathbf{a}_i\right\}, \quad (2)$$

$$\mathcal{J}(\mathbf{a}_i) \triangleq W(z, |\mathbf{a}_{\text{source}} - \mathbf{a}_{\text{dest}}|, |\mathbf{a}_i - \mathbf{a}_{\text{dest}}|)\mathcal{G}(\mathbf{a}_i). \quad (3)$$

B. Data Capture Probability

Let $\mathcal{E}_{\text{R}}, \mathcal{E}_{\text{C}}, \mathcal{E}_{\text{P}}$ and $\mathcal{E}_{\text{A}}$ respectively denote the capture events that a slot of the RTS, CTS, PAY and ACK are recovered. Note that for $P > 1$, $\mathcal{E}_{\text{P}}$ is the capture event of a given slot during PAY, but is statistically the same for any slot due to the assumption that the channel is independent for every slot. The event that the data in a slot carried by the PAY is considered to be transported to the destination, denoted as $\mathcal{E}_{\text{data}}$, occurs if all the above four capture events occurs, since the ACK contains acknowledgement bits for all PAY slots. Hence, the probability that the data in a PAY slot is recovered by the AP, i.e., the *data capture probability*, can be expressed using the chain rule as

$$\begin{aligned}\Pr(\mathcal{E}_{\text{data}}) &= \Pr(\mathcal{E}_{\text{R}}, \mathcal{E}_{\text{C}}, \mathcal{E}_{\text{P}}, \mathcal{E}_{\text{A}}) \\ &= \Pr(\mathcal{E}_{\text{R}}) \Pr(\mathcal{E}_{\text{C}}|\mathcal{E}_{\text{R}}) \Pr(\mathcal{E}_{\text{P}}|\mathcal{E}_{\text{R}}, \mathcal{E}_{\text{C}}) \Pr(\mathcal{E}_{\text{A}}|\mathcal{E}_{\text{R}}, \mathcal{E}_{\text{C}}, \mathcal{E}_{\text{P}}). \quad (4)\end{aligned}$$

Unlike the RTS and CTS which inhibits other users, reducing the ACK rate does not incur any penalty on the network. Often, the ACK is transmitted at the lowest possible rate to ensure reliable communication. To make the analysis concise, we approximate R_{A} as (effectively) zero and so $\Pr(\mathcal{E}_{\text{A}}|\mathcal{E}_{\text{R}}, \mathcal{E}_{\text{C}}, \mathcal{E}_{\text{P}}) = 1$. Hence, to compute $\Pr(\mathcal{E}_{\text{data}})$ then only requires computing the first three factors of (4).

We use $\Phi, \Theta \in \{\text{R}, \text{C}, \text{P}, \text{A}\}$ to denote a generic packet type. Furthermore, $\mathcal{E}$ refers to any or some of $\{\mathcal{E}_{\text{R}}, \mathcal{E}_{\text{C}}, \mathcal{E}_{\text{P}}, \mathcal{E}_{\text{A}}, \emptyset\}$, where $\emptyset$ denotes a null set. We also use the following notations:

1) R_Φ is the rate used for transmitting Φ;

2) $z_\Phi \triangleq 2^{R_\Phi} - 1$ is the capture ratio;

3) $\mathcal{G}_{\Phi|\mathcal{E}}(\mathbf{a})$ is the spatial traffic intensity at $\mathbf{a}$ when Φ is transmitted conditioned on event $\mathcal{E}$; and

4) $\mathcal{P}_{\Phi|\mathcal{E}}(\mathbf{a})$ is the probability that Φ captures a receiver at $\mathbf{a}$ conditioned on event $\mathcal{E}$.

Here, the position vector $\mathbf{a}$ is the position of an interferer for the traffic intensity or of a receiver for the capture probability.

C. Relationship of Traffic Intensity and Capture Probability

Computing any capture probability can be carried out using (2) if the traffic intensity is known. Hence, the problem of finding the capture probability is reduced to finding the corresponding traffic intensity. Based on the RTS/CTS scheme described in Sect. A., the capture probability for different packet types at $\mathbf{a}$ is

$$\mathcal{P}_{\text{R}|\mathcal{E}}(\mathbf{a}) = \Pr\left\{\mathcal{E}_{\text{cap}}(\mathbf{a}_{\mathsf{STA}}, \mathbf{a}, R_{\text{R}}, \mathcal{G}_{\text{R}|\mathcal{E}})\right\}, \quad (5)$$

$$\mathcal{P}_{\text{C}|\mathcal{E}}(\mathbf{a}) = \Pr\left\{\mathcal{E}_{\text{cap}}(\mathbf{a}_{\mathsf{AP}}, \mathbf{a}, R_{\text{C}}, \mathcal{G}_{\text{C}|\mathcal{E}})\right\}, \quad (6)$$

$$\mathcal{P}_{\text{P}|\mathcal{E}}(\mathbf{a}) = \Pr\left\{\mathcal{E}_{\text{cap}}(\mathbf{a}_{\mathsf{STA}}, \mathbf{a}, R_{\text{P}}, \mathcal{G}_{\text{P}|\mathcal{E}})\right\}. \quad (7)$$

The individual capture probability of (4) can then be expressed concisely as special cases of (5), (6), (7), respectively:

$$\Pr(\mathcal{E}_{\text{R}}) = \mathcal{P}_{\text{R}}(\mathbf{a}_{\mathsf{AP}}), \quad (8)$$

$$\Pr(\mathcal{E}_{\text{C}}|\mathcal{E}_{\text{R}}) = \mathcal{P}_{\text{C}|\mathcal{E}_{\text{R}}}(\mathbf{a}_{\mathsf{STA}}), \quad (9)$$

$$\Pr(\mathcal{E}_{\text{P}}|\mathcal{E}_{\text{R}}, \mathcal{E}_{\text{C}}) = \mathcal{P}_{\text{P}|\mathcal{E}_{\text{R}},\mathcal{E}_{\text{C}}}(\mathbf{a}_{\mathsf{AP}}). \quad (10)$$

D. Capture Probabilities $\Pr(\mathcal{E}_{\text{R}}), \Pr(\mathcal{E}_{\text{C}}|\mathcal{E}_{\text{R}}), \Pr(\mathcal{E}_{\text{P}}|\mathcal{E}_{\text{R}}, \mathcal{E}_{\text{C}})$

We summarized the computations required to calculate $\Pr(\mathcal{E}_{\text{R}}), \Pr(\mathcal{E}_{\text{C}}|\mathcal{E}_{\text{R}}), \Pr(\mathcal{E}_{\text{P}}|\mathcal{E}_{\text{R}}, \mathcal{E}_{\text{C}})$; details are provided in

$$\stackrel{(12)}{\mapsto} \mathcal{G}_{\mathrm{R}|\mathcal{E}_\mathrm{R}}(\mathbf{a}) \stackrel{(5)}{\longrightarrow} \mathcal{P}_{\mathrm{R}|\mathcal{E}_\mathrm{R}}(\mathbf{a}) \stackrel{(13)}{\longrightarrow} \mathcal{G}_{\mathrm{C}|\mathcal{E}_\mathrm{R}}(\mathbf{a}) \stackrel{(6)}{\longrightarrow} \mathcal{P}_{\mathrm{C}|\mathcal{E}_\mathrm{R}}(\mathbf{a}) \stackrel{(9)}{\longrightarrow} \Pr(\mathcal{E}_\mathrm{C}|\mathcal{E}_\mathrm{R})$$

Figure 2: Relationship of capture probabilities and traffic intensities over time and space in calculating $\Pr(\mathcal{E}_\mathrm{C}|\mathcal{E}_\mathrm{R})$.

$$\cdots \rightarrow \mathcal{G}_{\mathrm{C}|\mathcal{E}_\mathrm{R}}(\mathbf{a}) \stackrel{(14)}{\longrightarrow} \mathcal{G}_{\mathrm{C}|\mathcal{E}_\mathrm{R},\mathcal{E}_\mathrm{C}}(\mathbf{a}) \stackrel{(6)}{\longrightarrow} \mathcal{P}_{\mathrm{C}|\mathcal{E}_\mathrm{R},\mathcal{E}_\mathrm{C}}(\mathbf{a}) \stackrel{(15)}{\longrightarrow} \mathcal{G}_{\mathrm{P}|\mathcal{E}_\mathrm{R},\mathcal{E}_\mathrm{C}}(\mathbf{a}) \stackrel{(7)}{\longrightarrow} \mathcal{P}_{\mathrm{P}|\mathcal{E}_\mathrm{R},\mathcal{E}_\mathrm{C}}(\mathbf{a}) \stackrel{(10)}{\longrightarrow} \Pr(\mathcal{E}_\mathrm{P}|\mathcal{E}_\mathrm{R},\mathcal{E}_\mathrm{C})$$

Figure 3: Relationship of capture probabilities and traffic intensities over time and space in calculating $\Pr(\mathcal{E}_\mathrm{P}|\mathcal{E}_\mathrm{R},\mathcal{E}_\mathrm{C})$.

[6]. To emphasis the effects of interference, we consider the case when the noise is negligible, i.e., $N_o = 0$.

The RTS capture probability can be derived in closed-form as

$$\Pr(\mathcal{E}_\mathrm{R}) = \exp\left\{-a_s^2\pi^2 G_o\sqrt{z_\mathrm{R}}/2\right\} \quad (11)$$

where $a_s = |\mathbf{a}_\mathsf{STA} - \mathbf{a}_\mathsf{AP}|$ is the distance between STA and AP. On the other hand, the conditional CTS capture probability is fairly difficult to compute, given by the sequence of computation (as time progresses in the RTS/CTS cycle) in Fig. 2. To complete the computation, the conditional traffic intensities can be shown to be

$$\mathcal{G}_{\mathrm{R}|\mathcal{E}_\mathrm{R}}(\mathbf{a}) = G_o(1 - W(z_R, a_s, |\mathbf{a} - \mathbf{a}_\mathsf{AP}|)), \quad (12)$$
$$\mathcal{G}_{\mathrm{C}|\mathcal{E}_\mathrm{R}}(\mathbf{a}) = G_o\left(1 - \mathcal{P}_{\mathrm{R}|\mathcal{E}_\mathrm{R}}(\mathbf{a})\right). \quad (13)$$

Similarly, the conditional PAY capture probability is computed as shown in Fig. 3 using the following conditional traffic intensities

$$\mathcal{G}_{\mathrm{C}|\mathcal{E}_\mathrm{R},\mathcal{E}_\mathrm{C}}(\mathbf{a}) = (1 - W(z_\mathrm{C}, a_s, |\mathbf{a} - \mathbf{a}_\mathsf{STA}|))\mathcal{G}_{\mathrm{C}|\mathcal{E}_\mathrm{R}}(\mathbf{a}), (14)$$
$$\mathcal{G}_{\mathrm{P}|\mathcal{E}_\mathrm{R},\mathcal{E}_\mathrm{C}}(\mathbf{a}) = G_o\left(1 - \mathcal{P}_{\mathrm{C}|\mathcal{E}_\mathrm{R},\mathcal{E}_\mathrm{C}}(\mathbf{a})\right). \quad (15)$$

Note that $\mathcal{G}_{\mathrm{C}|\mathcal{E}_\mathrm{R}}$ is required to start the computation and is available from Fig. 2.

IV. Lower Bound Capture Probabilities

Closed-form expressions ease optimization studies, specifically for throughput maximization. Although $\Pr(\mathcal{E}_\mathrm{R})$ is available in closed form, *exact* closed forms for the conditional CTS and PAY capture probabilities are not available. Instead, *lower bound* closed forms are derived here.

A general approach to lower bound the capture probability during Φ is to upper bound the corresponding traffic intensity during Φ. We show that this upper bound traffic intensity can be obtained by using a *uniform* upper bound traffic intensity during Θ where Θ occurs before Φ. By choosing Θ appropriately, a closed-form for the capture probability can then be obtained. All lower bound capture probabilities are denoted by capping the exact ones with tildes; all *upper bound* traffic intensities are denoted similarly.

A. Lower Bound CTS Capture Probability

We proceed in the forward time sense as indicated by the arrows in Fig. 2. From (12), a valid upper bound of $\mathcal{G}_{\mathrm{R}|\mathcal{E}_\mathrm{R}}(\mathbf{a})$ is G_o since it can be shown that the W function is always positive. This gives us a lower bound to $\mathcal{P}_{\mathrm{R}|\mathcal{E}_\mathrm{R}}(\mathbf{a})$, which can be appreciated by observing the relationship of the capture probability and the traffic intensity in (2). Since the upper bound traffic is uniform, the lower bound capture probability is the same as the RTS capture probability, therefore

$$\widetilde{\mathcal{P}}_{\mathrm{R}|\mathcal{E}_\mathrm{R}}(\mathbf{a}) = \exp\left\{-b^2\pi^2 G_o\sqrt{z_\mathrm{R}}/2\right\} \quad (16)$$

where $b = |\mathbf{a}_\mathsf{STA} - \mathbf{a}|$. From (16) and (13), we then obtain an upper bound to $\mathcal{G}_{\mathrm{C}|\mathcal{E}_\mathrm{R}}(\mathbf{a})$ as

$$\widetilde{\mathcal{G}}_{\mathrm{C}|\mathcal{E}_\mathrm{R}}(\mathbf{a}) = G_o\left(1 - \widetilde{\mathcal{P}}_{\mathrm{R}|\mathcal{E}_\mathrm{R}}(\mathbf{a})\right) \quad (17)$$

which can be alternatively written as function of b, i.e., as $\widetilde{\mathcal{G}}_{\mathrm{C}|\mathcal{E}_\mathrm{R}}(b)$. Using (2) with transformation of the random variable from $\mathbf{a}$ to b, the desired lower bound is

$$\widetilde{\Pr}(\mathcal{E}_\mathrm{C}|\mathcal{E}_\mathrm{R}) = \exp\left\{-\int_0^\infty 2\pi b W(z_\mathrm{C}, a_s, b)\widetilde{\mathcal{G}}_{\mathrm{C}|\mathcal{E}_\mathrm{R}}(b)\,\mathrm{d}b\right\}, \quad (18)$$

where noise is assumed to be absent. After some derivations (omitted due to lack of space), we obtain

$$\widetilde{\Pr}(\mathcal{E}_\mathrm{C}|\mathcal{E}_\mathrm{R}) = \exp\{a_s^2\pi G_o\sqrt{z_\mathrm{C}} \times g(G_o\pi^2 a_s^2\sqrt{z_\mathrm{C} z_\mathrm{R}}/2) - a_s^2\pi^2 G_o\sqrt{z_\mathrm{C}}/2\} \quad (19)$$

where $g(s) \triangleq \cos(s)\left[\frac{\pi}{2} - \mathrm{Si}(s)\right] + \sin(s)\mathrm{Ci}(s)$, $\mathrm{Si}(s) = \frac{\pi}{2} - \int_s^\infty \frac{\sin y}{y}\,\mathrm{d}y$ is the sine integral and $\mathrm{Ci}(s) = -\int_s^\infty \frac{\cos y}{y}\,\mathrm{d}y$ the cosine integral.

B. Lower Bound PAY Capture Probability

The lower bound for PAY follows the same derivation as that for CTS. We proceed according to Fig. 3. First we bound $\mathcal{G}_{\mathrm{C}|\mathcal{E}_\mathrm{R},\mathcal{E}_\mathrm{C}}(\mathbf{a})$ by G_o. This is valid as seen from (14) and (13). Carrying on the arguments in a similar way, we arrive at

$$\widetilde{\mathcal{P}}_{\mathrm{C}|\mathcal{E}_\mathrm{R},\mathcal{E}_\mathrm{C}}(\mathbf{a}) = \exp\left\{-b^2\pi^2 G_o\sqrt{z_\mathrm{C}}/2\right\}, \quad (20)$$
$$\widetilde{\mathcal{G}}_{\mathrm{P}|\mathcal{E}_\mathrm{R},\mathcal{E}_\mathrm{C}}(\mathbf{a}) = G_o\left(1 - \widetilde{\mathcal{P}}_{\mathrm{C}|\mathcal{E}_\mathrm{R},\mathcal{E}_\mathrm{C}}(\mathbf{a})\right), \quad (21)$$

where $b = |\mathbf{a}_\mathsf{AP} - \mathbf{a}|$. Therefore, similar to the derivation of (19), the capture probability for the PAY is lower bounded by

$$\widetilde{\Pr}(\mathcal{E}_\mathrm{P}|\mathcal{E}_\mathrm{R},\mathcal{E}_\mathrm{C}) = \exp\{a_s^2\pi G_o\sqrt{z_\mathrm{P}} \cdot g(G_o\pi^2 a_s^2\sqrt{z_\mathrm{P} z_\mathrm{C}}/2) - a_s^2\pi^2 G_o\sqrt{z_\mathrm{P}}/2\}. \quad (22)$$

V. Throughput

A. Analysis

For the purpose of analysis, we assume that the RTS/CTS cycles are spaced sufficiently far apart so that the inhibition period of one cycle do not overlap with the next . We consider the throughput of the RTS/CTS scheme averaged over the time when STA accesses the channel via RTS/CTS.

Consider the possible cases when a cycle is terminated from the STA's perspective. There are only two, either when (i) STA does not recover the CTS (including the case that the CTS may not be sent by the AP in the first place because the RTS is not recovered), or when (ii) STA has received the ACK (but may not necessarily recover it). These cases are indicated respectively as $\{\mathrm{R}, \mathrm{C}\}, \{\mathrm{R}, \mathrm{C}, \mathrm{P}, \mathrm{A}\}$. The first case occurs if the RTS *or* the CTS is not recovered with probability

$$\Pr(\bar{\mathcal{E}}_{\mathrm{R}}) + \Pr(\mathcal{E}_{\mathrm{R}}, \bar{\mathcal{E}}_{\mathrm{C}}) = 1 - \Pr(\mathcal{E}_{\mathrm{R}}, \mathcal{E}_{\mathrm{C}}), \tag{23}$$

where $\bar{\mathcal{E}}$ indicates the complement of event $\mathcal{E}$. The second case, therefore, occurs with probability $\Pr(\mathcal{E}_{\mathrm{R}}, \mathcal{E}_{\mathrm{C}})$.

The duration of each cycle of the RTS/CTS transmission, i.e., the *cycle time*, is i.i.d. As a result, the cycle time is a renewal process. Let $s_t \in \{0, L_{\mathrm{sym}} \times R_{\mathrm{P}}\}$ be the amount of bits recovered in slot t, where L_{sym} is the number of symbols sent per slot. Note that s_t can be greater than zero only during PAY period, otherwise it is zero since only overheads are sent. By using the renewal-reward theorem [8], the time average of the throughput in bit/symbol over an infinitely large period is

$$\begin{aligned}\bar{s}(a_s, R_{\mathrm{R}}, R_{\mathrm{C}}, R_{\mathrm{P}}, P) &\triangleq \lim_{N\to\infty} \frac{1}{L_{\mathrm{sym}}} \frac{1}{N} \sum_{t=1}^{N} s_t \\ &= \frac{1}{L_{\mathrm{sym}}} \frac{\mathbb{E}[\mathcal{R}]}{\mathbb{E}[\mathcal{T}]}\end{aligned} \tag{24}$$

with probability 1. Here, the expectation $\mathbb{E}$ is carried out over the events in one cycle, $\mathcal{R}$ is the *reward* in the form of the sum of s_t in a cycle, and $\mathcal{T}$ is the cycle time in slots. As explicitly denoted, the time averaged throughput is a function of many parameters. We take a_s to be fixed, and consider how the rates and PAY length P affect the throughput. For notational convenience, the arguments of $\bar{s}$ are dropped.

On average, $P \Pr(\mathcal{E}_{\mathrm{data}})$ slots are successful during the PAY period. Hence, the expected reward is

$$\mathbb{E}[\mathcal{R}] = L_{\mathrm{sym}} R_{\mathrm{P}} P \Pr(\mathcal{E}_{\mathrm{data}}) \tag{25}$$

Taking into account that the cases $\{\mathrm{R}, \mathrm{C}\}, \{\mathrm{R}, \mathrm{C}, \mathrm{P}, \mathrm{A}\}$ use 2 and $P + 3$ slots, respectively, the expected cycle time is

$$\begin{aligned}\mathbb{E}[\mathcal{T}] &= 2(1 - \Pr(\mathcal{E}_{\mathrm{R}}, \mathcal{E}_{\mathrm{C}})) + (P+3)\Pr(\mathcal{E}_{\mathrm{R}}, \mathcal{E}_{\mathrm{C}}) \\ &= 2 + (P+1)\Pr(\mathcal{E}_{\mathrm{R}}, \mathcal{E}_{\mathrm{C}}).\end{aligned} \tag{26}$$

This result can be explained as follows. From (26), it is seen that at least two slots are used for the cycle. This overhead corresponds to the RTS and CTS slot since STA has to wait for the CTS to arrive before deciding to transmit PAY or to defer the transmission. Subsequently, if the RTS and CTS are recovered (with probability $\Pr(\mathcal{E}_{\mathrm{R}}, \mathcal{E}_{\mathrm{C}})$), P slots are used for PAY and one slot for a potential ACK. The potential ACK requires one slot since STA has to spend time receiving the ACK regardless of whether it is actually sent by the AP.

Using (25) and (26), (24) becomes

$$\begin{aligned}\bar{s} &= \frac{R_{\mathrm{P}} P \Pr(\mathcal{E}_{\mathrm{data}})}{2 + (P+1)\Pr(\mathcal{E}_{\mathrm{R}}, \mathcal{E}_{\mathrm{C}})} \\ &= R_{\mathrm{P}} \Pr(\mathcal{E}_{\mathrm{P}} | \mathcal{E}_{\mathrm{R}}, \mathcal{E}_{\mathrm{C}}) \times \Omega(P),\end{aligned} \tag{27}$$

$$\text{where} \quad \Omega(P) \triangleq \frac{P \Pr(\mathcal{E}_{\mathrm{R}}, \mathcal{E}_{\mathrm{C}})}{2 + (P+1)\Pr(\mathcal{E}_{\mathrm{R}}, \mathcal{E}_{\mathrm{C}})}. \tag{28}$$

The last line of (27) is obtained by using (4) and assuming $\Pr(\mathcal{E}_{\mathrm{A}} | \mathcal{E}_{\mathrm{R}}, \mathcal{E}_{\mathrm{C}}, \mathcal{E}_{\mathrm{P}}) = 1$.

Property 1 *For a given $R_{\mathrm{R}}, R_{\mathrm{C}}, R_{\mathrm{P}}$, the throughput $\bar{s}$ is an increasing function of P. Hence, an upper bound is given by the asymptotic throughput*

$$\lim_{P\to\infty} \bar{s} = R_{\mathrm{P}} \Pr(\mathcal{E}_{\mathrm{P}} | \mathcal{E}_{\mathrm{R}}, \mathcal{E}_{\mathrm{C}}). \tag{29}$$

Proof 1 *Note that $R_{\mathrm{P}} \Pr(\mathcal{E}_{\mathrm{P}} | \mathcal{E}_{\mathrm{R}}, \mathcal{E}_{\mathrm{C}})$ in (27) is independent of P, while as P increases, $\Omega(P)$ increases (hence proving the first statement of the proposition) and approaches 1 (hence proving the second).*

Property 2 *For any given $R_{\mathrm{R}}, R_{\mathrm{C}}, P$, the optimal R_{P} that maximizes $\bar{s}$ is the same.*

Proof 2 *When $R_{\mathrm{R}}, R_{\mathrm{C}}, P$ are given, $\Omega(P)$ is independent of R_{P} while R_{P} and $\Pr(\mathcal{E}_{\mathrm{P}} | \mathcal{E}_{\mathrm{R}}, \mathcal{E}_{\mathrm{C}})$ are functions of R_{P}. Hence, maximizing $\bar{s}$ is equivalent to maximizing (29), and is independent of the actual values of $R_{\mathrm{R}}, R_{\mathrm{C}}, P$.*

From Properties 1, 2, it follows that for any given $R_{\mathrm{R}}, R_{\mathrm{C}}$, the maximum throughput $\bar{s}$ obtained by optimizing R_{P} increases with P. This is because increasing P reduces the overhead and does not have any negative effects, which follows intuitively from the fact that the data are detected independently on a slot by slot basis.

To improve the reliability of the data, the lowest practical rate for $R_{\mathrm{R}}, R_{\mathrm{C}}$ may be chosen to ensure the largest reservation effect over space. However, it may be necessary to reduce the amount of inhibition, in which case moderate values of rates may be chosen. After $R_{\mathrm{R}}, R_{\mathrm{C}}$ are fixed, as a consequence of Properties 1, 2, we can choose P as large as practically possible and also independently optimize R_{P} for the given $R_{\mathrm{R}}, R_{\mathrm{C}}$.

Property 3 *A lower bound of $\bar{s}$ is obtained by substituting the capture probabilities in (27) with their lower bounds (19), (22).*

Proof 3 *By rewriting $\Omega(P) = \frac{P}{2/\Pr(\mathcal{E}_{\mathrm{R}}, \mathcal{E}_{\mathrm{C}}) + (P+1)}$, we obtain*

$$\bar{s} \geq \tilde{s} \triangleq R_{\mathrm{P}} \widetilde{\Pr}(\mathcal{E}_{\mathrm{P}} | \mathcal{E}_{\mathrm{R}}, \mathcal{E}_{\mathrm{C}}) \frac{P}{2/\widetilde{\Pr}(\mathcal{E}_{\mathrm{R}}, \mathcal{E}_{\mathrm{C}}) + (P+1)}. \tag{30}$$

B. *Numerical Results*

The throughput $\bar{s}$ vs R_{P} is plotted in Fig. 4. The following system parameters are fixed: $R_{\mathrm{R}} = R_{\mathrm{C}} = 0.5, a_s = 0.5, G_o = 1/\pi$, while P are varied. For a fixed P, as R_{P} increases, $\bar{s}$ increases initially, peaking at some R_{P}, say $R_{\mathrm{P}}^{\mathrm{o}}(P)$, then decreases. This is because transmitting at a low R_{P} limits the throughput, while transmitting at a high R_{P} reduces the throughput as the PAY becomes unreliable. Furthermore, as P increases, the throughput increases for any fixed R_{P}, as stated in Property 1. From Fig. 4, Property 2 is also apparent, since the $R_{\mathrm{P}}^{\mathrm{o}}(P) = 3.6$ for all P.

In addition, the throughput when STA transmits to AP using the slotted ALOHA is plotted in Fig. 4. We assumed for simplicity that ACK is not required for ALOHA and hence

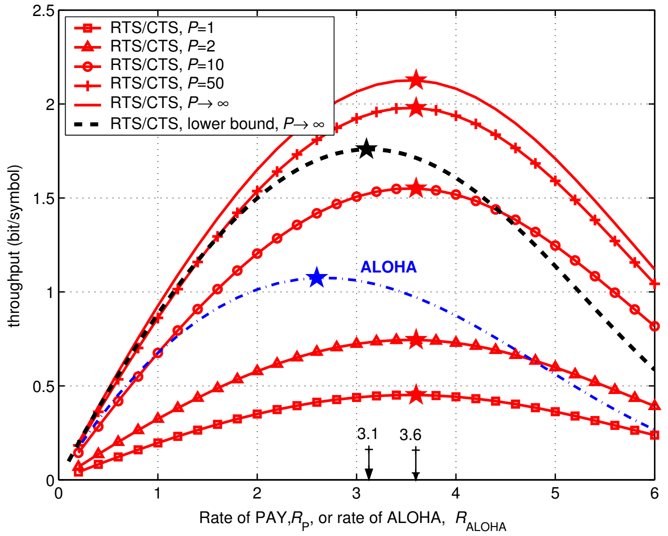

Figure 4: Throughput $\bar{s}$ vs rate of PAY slots R_{P} for various packet length P. The analytical lower bound when $P \to \infty$, and the throughput for ALOHA transmission are also plotted. The stars indicates where the maximum throughput is achieved. Parameters: $R_{\mathrm{R}} = R_{\mathrm{C}} = 0.5, a_s = 0.5, G_o = 1/\pi$.

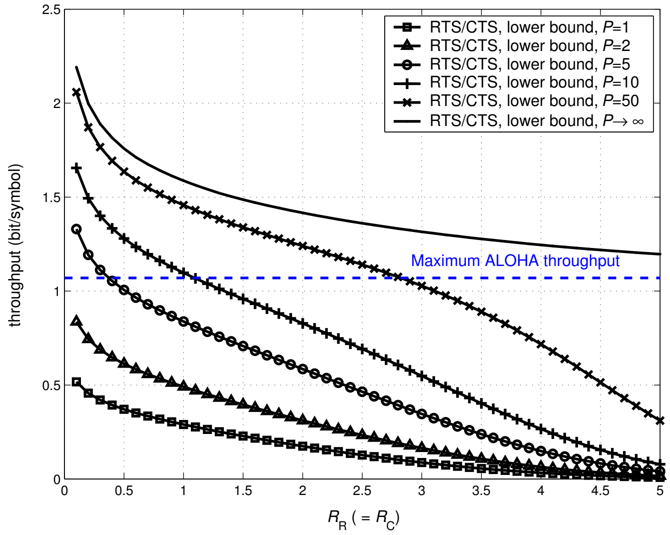

Figure 5: Lower bound RTS/CTS throughput maximized over R_{P}, plotted against $R_{\mathrm{R}} = R_{\mathrm{C}}$. For sufficiently small R_{R} and sufficiently large P, the RTS/CTS throughput is larger than the ALOHA throughput. Parameters: $a_s = 0.5, G_o = 1/\pi$.

no overhead is incurred. The ALOHA throughput is then given by the data rate, say R_{ALOHA}, multiplied by the (non-inhibiting) capture probability, i.e., (11) with z_{R} replaced by $z_{\mathrm{ALOHA}} = 2^{R_{\mathrm{ALOHA}}} - 1$. It is seen that the maximum throughput of the RTS/CTS scheme can be significantly larger than that of the ALOHA scheme. For example, when $P = 50$ and $R_{\mathrm{P}} \geq 0.1$, the RTS/CTS scheme is always better than the ALOHA scheme for $R_{\mathrm{P}} = R_{\mathrm{ALOHA}}$.

Consider the throughput lower bound $\tilde{s}$ when $P \to \infty$, also plotted in Fig. 4. It is possible to approximate R_{P}^o from the lower bound, hence giving a simple method for performing rate adaption for the PAY slots. Moreover, using the optimal rate obtained from the lower bound at $R_{\mathrm{P}} = 3.1$, a throughput that is fairly close to the maximum possible is achieved. For example, for $P \to \infty$, a throughput of 2.085 bit/symbol is achieved at $R_{\mathrm{P}} = 3.1$. There is a slight loss of 0.041 bit/symbol compared to the optimal throughput of 2.126 bit/symbol achieved at $R_{\mathrm{P}}^o = 3.6$.

Lastly, we obtain a conservative estimate of the maximum RTS/CTS throughput by maximizing R_{P} over the analytical lower bound $\tilde{s}$. For simplicity, we set $R_{\mathrm{R}} = R_{\mathrm{C}}$ although further optimization of the two rates could yield larger throughput. From Fig. 5, increasing $R_{\mathrm{R}} = R_{\mathrm{C}}$ reduces the RTS/CTS throughput since the amount of interference increases correspondingly, while increasing P increases the throughput since the overhead used becomes negligible. The ALOHA throughput, optimized over R_{ALOHA}, is also plotted. The ALOHA throughput is a constant since it does not change with R_{R} or R_{C}. It is observed that the lower bound RTS/CTS throughput is larger than the ALOHA throughput of 1.1 when R_{R} is sufficiently small (to ensure adequate reservation) and P is sufficiently large (to compensate for the overhead).

VI. Conclusion

It is shown how the capture probabilities can be used to perform throughput optimization via rate adaptation and the RTS/CTS scheme attains higher throughput than ALOHA when the RTS and CTS rates are sufficiently small and when the payload packet is sufficiently large. Due to the large degree of freedom offered by rate adaptations, other system level optimizations can also be conducted, such as to improve other quality of service.

References

[1] L. Kleinrock and F. Tobagi, "Packet switching in radio channels: Part I–carrier sense multiple-access modes and their throughput-delay characteristics," *IEEE Trans. Commun.*, vol. 23, no. 12, pp. 1400–1416, Dec. 1975.

[2] P. Karn, "MACA- a new channel access method for packet radio," in *ARRL/CRRL Amateur Radio 9th Computer Networking Conference*, Sept. 1990, pp. 134–140.

[3] G. Bianchi, "IEEE 802.11-saturation throughput analysis," *IEEE Commun. Lett.*, vol. 2, no. 12, pp. 318–320, Dec. 1998.

[4] Z. Hadzi-Velkov and B. Spasenovski, "Capture effect in IEEE 802.11 basic service area under influence of Rayleigh fading and near/far effect," in *Proc. 13th IEEE Personal, Indoor and Mobile Radio Communications*, vol. 1, Lisbon, Portugal, Sept. 2002, pp. 172–176.

[5] J. H. Kim and J. K. Lee, "Capture effects of wireless CSMA/CA protocols in Rayleigh and shadow fading channels," *IEEE Trans. Veh. Technol.*, vol. 48, no. 4, pp. 1277–1286, July 1999.

[6] C. K. Ho and J. P. M. G. Linnartz, "Calculation of the spatial reservation area for the rts/cts multiple access scheme," in *WIC Twenty-seventh Symposium on Information Theory in the Benelux*, Noordwijk, The Netherlands, June 2006, pp. 181–188.

[7] J. P. M. G. Linnartz, "Slotted ALOHA land-mobile radio networks with site diversity," *IEE Proc.-I*, vol. 139, no. 1, pp. 58–70, Feb. 1992.

[8] R. G. Gallager, *Discrete Stochastic Processes*. Boston: Kluwer Academic Publishers, 1996.