A Protocol-Independent Approach for Analyzing the Optimal Operation Point of CSMA/CA Protocols

Yu Cheng
Electrical and Computer Engineering
Illinois Institute of Technology
cheng@iit.edu

Xinhua Ling
BlackBerry System Architecture
Research In Motion
xling@rim.com

Weihua Zhuang
Electrical and Computer Engineering
University of Waterloo
wzhuang@uwaterloo.ca

Abstract—This paper presents a protocol-independent approach to reveal a new insight into the performance of carrier sense multiple access with collision avoidance (CSMA/CA) protocols: the family of CSMA/CA protocols, independent of implementation details, share the same optimal operation point where the maximum protocol capacity is achieved. The protocol-independent analysis is inspired by the concept of virtual time slot. At the timescale of virtual-slot, all the CSMA/CA protocols show the same behavior pattern and, therefore, a generic virtual-slot based S-G (VS S-G) analysis is developed to compute the optimal operation point. The accuracy of the VS S-G analysis is benchmarked against the precise protocol-specific analysis, in particular, for the 802.11 distributed coordination function (DCF) and the 802.15.4 contention access period (CAP). Furthermore, this paper discusses how to integrate the network-layer queueing analysis with the VS S-G analysis at the medium access control (MAC) layer to form a generic cross-layer framework for call-level network capacity analysis.

I. INTRODUCTION

The medium access control (MAC) protocol plays a critical role in determining the capacity of a wireless network. The family of carrier sense multiple access with collision avoidance (CSMA/CA) protocols has been adopted as the main wireless MAC solution. Different flavors of CSMA/CA protocols are extensively applied in wireless sensor networks and personal area networks (WPANs), e.g., the IEEE 802.15.4 [1]; in single-hop wireless local area networks (WLANs) and multi-hop wireless mesh networks, e.g., the family of IEEE 802.11 [2]; and in vehicular ad hoc networks, e.g., the IEEE 802.11p [3].

A fundamental issue in wireless network performance analysis is to compute the maximum MAC protocol capacity, normally defined as the maximum channel utilization or maximum throughput over the wireless links. A significant inconvenience in applying the existing methods for MAC capacity analysis, either the Markov chain based modeling [4]–[6] or the tagged user analysis [7]–[9], is that the protocol details have to be thoroughly investigated to establish and solve the MAC model. Moreover, protocol-specific approaches have limited applicability, since new physical-layer techniques and MAC-protocols keep emerging rapidly. In this paper, we aim at developing a novel *protocol-independent* approach, which only depends on the CSMA/CA nature but independent of

the MAC implementation details, to obtain the MAC protocol capacity.

We focus on the single-hop wireless networks. It has been demonstrated for 802.11 based WLAN that maintaining the MAC at an *optimal operation point* is essential for achieving the maximum protocol capacity [10]. The simulation and analytical studies in [10], [11] show that the *channel busyness ratio* (CBRO) is an appropriate indicator of the optimal operation point, whose value under the 802.11 distributed co-ordination function (DCF) MAC maintains stable independent of the number of mobile nodes. In this paper, we reveal that the optimal operation point (i.e., the optimal CBRO), as an upper bound defining the operation region of the MAC, is determined by the CSMA/CA nature rather than by the MAC implementation details. A protocol-independent approach is developed to accurately compute the optimal operation point.

The protocol independent analysis is inspired by the fact that a slotted CSMA/CA protocol can be described by the concept of *virtual time slot*, which is defined according to the type of event that happens at the beginning of that slot [4], [12], [13]. Observed at the timescale of virtual slot, all the CSMA/CA protocols show the same behavior pattern: the slots alternate among the statuses of idling, collision, and successful transmission. At the steady state, a virtual-slot protocol in fact behaves similarly to a slotted ALOHA protocol, given the probabilities associated with the three types of virtual slots. It is well-known that the S-G analysis [14], [15] can be applied to obtain the maximum MAC capacity of slotted ALOHA. As a counterpart, we develop the *virtual-slot based S-G (VS S-G) analysis* as a generic approach to compute the optimal operation point and the corresponding maximum MAC capacity for all the CSMA/CA protocols.

The accuracy of the VS S-G analysis is to be validated by comparing the analytical results with those obtained from protocol-specific analysis. We consider two typical CSMA/CA protocols, the 802.11 DCF and the 802.15.4 contention access period (CAP). The optimal CBRO for the 802.11 DCF is derived in [10]; the optimal CBRO analysis for 802.15.4 CAP is a new contribution in this paper, which by itself should be of interest to the research community. The computer simulations are used to ensure the preciseness of the protocol specific analysis, serving as benchmarks to evaluate the VS S-G analysis. Moreover, the details of the CBRO analysis for

802.11 DCF and 802.15.4 CAP illustrate the complexity of protocol-specific MAC modeling. The VS S-G analysis can be readily applied in both cases, given the MAC parameters, and the results match the protocol-specific results closely.

One important application of the optimal operation point is to achieve the maximum admission region under the call admission control (CAC). It has been demonstrated for 802.11 DCF that CAC based on the optimal CBRO can lead to the maximum quality of service (QoS) capable admission region [10], [11]. In [16], we integrate the optimal CBRO obtained in [10] with network-layer queueing analysis to obtain the maximum number of on/off voice flows supportable over a WLAN, under the QoS constraint of a delay bound. In this paper, we will discuss how to extend the admission control model of [16] into a generic *cross-layer analytic framework* by incorporating the protocol-independent optimal CBRO. Such a framework can significantly facilitate the *call-level capacity* analysis, defined as the maximum number of QoS-guaranteed flows that can be supported by the wireless network.

The remainder of this paper is organized as follows. Section II reviews more related work. Section III develops the protocol-independent VS S-G analysis. The protocol-specific analysis for 802.11 DCF and 802.15.4 CAP are presented in Section IV and Section V, respectively. Section VI presents the numerical results to demonstrate the accuracy of the VS S-G analysis. Section VII discusses applying the optimal CBRO for call-level capacity analysis through the network/MAC cross-layer modeling. Section VIII gives the concluding remarks.

II. RELATED WORK

The seminal work of Bianchi [4] injected a strong impetus to the Markov chain based MAC modeling. In principle, the Markov modeling could be applied to any CSMA/CA protocol, with either saturated traffic [4]–[6] or unsaturated bursty traffic [17], [18]. However, the size of the state space will increase exponentially with the number of users and the user buffer capacity, which tends to make the Markov analysis intractable.

In more recent studies, the *tagged user analysis* (TUA) has been proposed as a more tractable yet accurate model for distributed MAC modeling [7]–[9]. With TUA, each user is regarded as a G/G/1 queueing system and its influence on others due to shared channel is taken into consideration. The TUA modeling has various advantages, including capability of per-node queueing analysis [7], [19], convenience in analyzing multiclass MAC supporting service differentiation [9], [16], flexibility in network/MAC cross-layer analysis [16], and extendability to multi-hop networks [20]. In this paper, the protocol-specific approach to compute the optimal operation point for 802.15.4 CAP is based on the TUA modeling.

It is noteworthy that all the existing CSMA/CA MAC analytical model exploits, either implicitly or explicitly, the *fixed-point* technique to facilitate the analysis [4], [21]. A fixed point is a parameter, based on which the stochastic MAC behavior of a certain user can be completely characterized, independent of other users. The fixed point technique has been applied both to obtain the transition probabilities for Markov

TABLE I
SUMMARY OF THE MAIN NOTATIONS

Notation	Definition
G	the Poisson arrival rate in a slot
N	the number of nodes in the network
p_d	the probability of seeing an idle slot
p_s	the probability of seeing an successful transmission slot
p_c	the probability of seeing an collision slot
p	the probability that a given transmission is collided
σ	the length of a idle slot
T_s	the length of a successful transmission slot
T_c	the length of a collision slot
R_s	the channel utilization or MAC throughput
R_b	the channel busyness ratio
τ	the channel access probability
φ	the probability that a node starts channel sensing in a given slot, under the 802.15.4 CAP
θ	the probability that a node sense the channel busy, under the 802.15.4 CAP
η_s	the per-node throughput under the 802.15.4 CAP
M	the maximum number of backoff stages, under the 802.15.4 CAP

modeling [4]–[6], [21] and to derive the user-specific queueing system with TUA modeling [7]–[9], [16], [22] .

One aspect of the complexity associated with the protocol-specific analysis is to properly determine the fixed point. For 802.11 DCF analysis, the channel access probability or the packet collision probability has been widely adopted as a fixed point. However, there is no theoretical proof that such fixed-point selection is generally applicable. In fact, there have been evidences showing that the channel access probability is not a valid fixed-point for 802.15.4 analysis [23], [24]. The proposed VS S-G analysis circumvents such complexity of fixed-point selection through protocol-independent traffic modeling. For convenience, the main mathematical notations used in this paper are summarized in Table I.

III. VIRTUAL-SLOT BASED S-G ANALYSIS

A virtual time slot is defined according to the type of event happening in that slot [4], [12], [13]. The length of a virtual time slot equals to one physical time slot σ if the channel is idle, or the packet transmission time T_s if the channel is experiencing a successful transmission, or the packet collision time T_c if a collision happens. The duration of T_s and T_c includes both the data transmission time and the extra time for MAC operations [4], [12], [16], [22].

Observed at the timescale of virtual slot, all the CSMA/CA protocols show the same behavior that the channel alternates among the statuses of idling, successful transmission, and collision. At the steady state, a typical slot resides at one of the three statuses with a certain probability. Let p_d, p_s, and p_c denote the probabilities of seeing an idle slot, a successful transmission slot, and a collision slot, respectively. With the size variation of different types of slots being ignored, the virtual-slot system behaves stochastically similar to a slotted ALOHA [14] protocol. Therefore, the well-known S-G analysis [14], [15] for computing the maximum throughput can be extended to analyzing the virtual-slot system.

A. Poisson Arrival in a Slot

With slotted ALOHA, each backlogged node independently accesses the channel with probability τ in each slot. Given the number of nodes N, the total number of packet arrivals in a slot follows a binomial distribution. The essence of S-G analysis is that when the number of nodes N goes to infinity with τ scaling to maintain $N\tau = G$, the number of packet arrivals in a slot converges to the Poisson random variable with parameter G. The Poisson arrival rate significantly facilitates computing the maximum throughput. In the following, we reveal that the packet arrivals in a virtual slot also converges to a Poisson random variable under the infinite-node scenario.

In a virtual-slot system, if it can be justified that the channel access activities from different backlogged nodes are independent, Poisson modeling of the arrival rate under the infinite-node scenario is also valid based on the reasoning similar to that used in the S-G analysis. The independent channel access is intuitively true for the 802.11 DCF MAC. With 802.11 DCF, each node's channel access activity is controlled by its local backoff mechanism. Moreover, the backoff operation is active only when the channel is idle; therefore, the channel accesses from different nodes are independent when the effect of backoff mechanism is summarized into a channel access probability τ [4], [16], [22]. Although there is no rigorous theoretical proof of the independence in channel access, such an assumption has been extensively used in 802.11 DCF analysis, either Markov chain based modeling [4]–[6] or TUA based modeling [7]–[9]. The accuracy of those analytical results, validated by simulations or experiments, indirectly justifies the independence in channel access.

The DCF MAC has been enhanced to provision multi-class service differentiation, e.g., in [25], [26] and the references therein. In the context of multi-class DCF MAC, the Poisson arrival rate under the infinite-node scenario is valid too. The main MAC service differentiation scheme is to differentiate the backoff parameters, which results in different channel access probabilities associated with different service classes. Assuming S classes provisioned, the backoff scheme associated with class-i leads to an access probability of τ_i and there are N_i class-i nodes. If N_i goes to infinity with $G_i = N\tau_i$ fixed for $i = 1, \cdots, S$, the number of packet arrivals associated with class-i in a slot converges to a Poisson random variable with parameter G_i, and the total arrival rate converges to the Poisson random variable with parameter $G = \sum_{i=1}^{S} G_i$.

The condition of independent channel access may not always hold. With the variety of CSMA/CA implementations, the backoff scheme may be different from that used in 802.11 DCF, and the service differentiation mechanisms may include differentiating arbitration interframe space (AIFS) and transmission opportunity (TXOP) [26], in addition to differentiating the backoff contention window. All these implementation details may induce correlation into the channel access processes from different nodes. Nevertheless, the packet arrivals in a slot can still be approximated by a Poisson random variable when the correlation in channel access is weak, according

to the *Central Limit Theorem* [27]. Specifically, we have the following theorem.

Theorem 1: For a CSMA/CA MAC protocol, when the packet arrival processes of different nodes are weakly correlated, the aggregate packet arrivals in a slot can be conservatively approximated by a Poisson random variable under the infinite-node scenario. The approximation is in the sense that both the actual aggregate arrivals and the Poisson approximation converges to Gaussian random variables with the same mean, but the Poisson approximation has a smaller variance.

Proof: Refer to Appendix A. ∎

B. Virtual-Slot S-G Analysis

By modeling a CSMA/CA protocol as a virtual-slot system, let G denote the Poisson traffic load in a slot. The probability of seeing an idle (a successful transmission) slot p_d (p_s) equals the probability that zero (a single) transmission trail happens at a slot. That is, $p_d = e^{-G}$, $p_s = Ge^{-G}$, and the probability of seeing a collision slot $p_c = 1 - Ge^{-G} - e^{-G}$. Thus, the steady-state channel utilization R_s can be computed as

$$R_s = \frac{p_s T_s}{p_d \sigma + p_s T_s + p_c T_c} \tag{1}$$

$$= \frac{Ge^{-G} T_s}{e^{-G}\sigma + Ge^{-G} T_s + (1 - Ge^{-G} - e^{-G})T_c}. \tag{2}$$

Following the S-G analysis [14], in order to obtain the maximum channel utilization (equivalently, the maximum MAC protocol capacity), we can determine the optimal workload G^* according to

$$\frac{d}{dG}R_s|_{G=G*} = 0 \tag{3}$$

which then gives

$$e^{G^*}(1 - G^*) =$$
$$\frac{G^* [T_s(1 - G^*) + G^*T_c - \sigma]}{e^{-G^*}\sigma + G^*e^{-G^*}T_s + (1 - G^*e^{-G^*} - e^{-G^*})T_c} \tag{4}$$

after some manipulation. The value of G^* can be numerically solved from (4) and then substituted into (2) to obtain the maximum channel utilization R_s^*.

It is recommended in [10] to use the channel busyness ratio rather than the channel utilization to indicate the operation point. The *channel busyness ratio* is defined as the proportion of time that the channel is in the status of successful transmission or collision. It is very convenient for a node to monitor the channel busyness ratio as a CSMA-based MAC protocol works on physical and virtual carrier sensing mechanisms. However, the channel utilization is hard to monitor, since it is difficult for a node to discriminate successful transmission from collision when other nodes are using the channel. Therefore, we define the *optimal channel busyness ratio* R_b^* as

$$R_b^* = \frac{G^*e^{-G^*} T_s + (1 - G^*e^{-G^*} - e^{-G^*})T_c}{e^{-G^*}\sigma + G^*e^{-G^*}T_s + (1 - G^*e^{-G^*} - e^{-G^*})T_c}. \tag{5}$$

The optimal CBRO indicates the optimal operation point where maximum protocol capacity is achieved.

It is noteworthy that the optimal CBRO is different from the maximum CBRO. When the traffic load is so high to overload the channel, most of the transmission will be collided. The channel in this scenario is very busy but the throughput (channel utilization) is very low. The experiences from studying the 802.11 DCF [10], [11] indicate that when the channel busyness ratio exceeds the optimal CBRO, the network becomes congested quickly; it is important to control the MAC operation region within the optimal CBRO to obtain high throughput, small delay, and small delay variance. The protocol-independent analysis based on (2) and (4) makes the simple yet efficient CBRO control readily available to the whole family of CSMA/CA protocols.

IV. PROTOCOL-SPECIFIC ANALYSIS OF 802.11 DCF

A protocol-specific approach for analyzing the optimal operation point of 802.11 DCF has been presented in [10], which is summarized in this section for completeness. More importantly, with comparison to the 802.15.4 CAP analysis presented in next section, the complexity of protocol-specific approaches can be clearly demonstrated.

Although the 802.11 DCF can be well described by a virtual-slot system, the protocol-specific analysis computes the optimal operation point through modeling the channel access details rather than exploiting the Poisson nature in the case of large number of nodes. Consider that the total number of nodes is N, the probabilities for observing an idle, successful, and colliding slot can be expressed respectively as

$$\begin{cases} p_d = (1-\tau)^N \\ p_s = N\tau(1-\tau)^{N-1} \\ p_c = 1 - p_d - p_s. \end{cases} \tag{6}$$

Correspondingly, the channel idleness ratio R_d, channel busyness ratio R_b, and channel utilization R_s can be obtained as

$$\begin{cases} R_d = \frac{p_d\sigma}{p_d\sigma + p_s T_s + p_c T_c} \\ R_b = 1 - R_d \\ R_s = \frac{p_s T_s}{p_d\sigma + p_s T_s + p_c T_c}. \end{cases} \tag{7}$$

Without loss of generality, we consider that all the MAC layer packets have the same length.[1] The virtual slot lengths, T_s for a successful transmission and T_c for a collision, can be determined from the 802.11 DCF standard. Another important metric indicating the efficiency of the MAC protocol is *normalized throughput V*, which is the net channel utilization after deducting the MAC protocol overhead due to interframe spacing and acknowledgment [2]. We have

$$V = \frac{p_s \times T_{\text{data}}}{p_d\sigma + p_s T_s + p_c T_c} = \frac{R_s \times T_{\text{data}}}{T_s}. \tag{8}$$

where T_{data} is the portion within T_s representing the effective transmission time.

---

[1]In the case that variable length packets are allowed, using the average length for MAC modeling can result in accurate steady-state performance analysis [4], [10], [16], [22].

In order to compute the optimal operation point, the *packet collision probability p* is further introduced, which is the probability that one node encounters collisions when it transmits. It is not difficult to see that p can be linked to τ as

$$p = 1 - (1-\tau)^{N-1}. \tag{9}$$

Using the relationship of (9) in (6), we can express p_d, p_s, and p_c as functions of p. Further applying $p_d(p)$, $p_s(p)$, and $p_c(p)$ into (7), the ratios R_d, R_b, and R_s can then be written as functions of p. We can therefore determine the optimal congestion probability $p*$ leading to the maximum channel utilization by

$$\frac{d}{dp}R_s|_{p=p*} = 0. \tag{10}$$

Applying p^* to (7) and (8), we can get the maximum channel utilization R_s^*, the optimal channel busyness ratio R_b^*, and the maximum normalized throughput V^*. Note that the optimal ratios here depend on the number of nodes N, while those obtained from VG S-G analysis do not.

V. PROTOCOL-SPECIFIC ANALYSIS OF 802.15.4 CAP

The IEEE 802.15.4 standard defines the protocol stack for low-power low-rate wireless networks. In the beacon-enabled operating mode, the contention access period within a superframe implements a non-persistent slotted CSMA/CA with binary exponential backoff MAC [1]. A salient difference between 80.15.4 CAP and 802.11 DCF is that a node can transmit only after two consecutive times of sensing an idle channel in the former, while just one channel sensing is required in the latter. In addition, it differs from the 802.11 DCF in that the backoff counter does not freeze when the channel is busy, but keeps on decreasing until reaching zero.

A. Correlated Channel Access

It has been discussed that Poisson modeling is accurate for slotted ALOHA or a virtual-slot MAC with independent channel access (e.g., 802.11 DCF): each node with a certain probability independently accesses the channel, after sensing an idle channel. Such independent channel access is broken in the 802.15.4 CAP, as the backoff operation disregards the channel status. The channel access effort from a node may be blocked by the on-going transmission due to a previous successful channel access from another node. The results presented in [23], [24] demonstrate that the channel access probability is not a valid fixed-point for 802.15.4 analysis.

For modeling 802.15.4 CAP, it has been suggested to use the *channel sensing probability* as the fixed point [23], [24], [28], which is the probability that a node starts channel sensing in a randomly chosen slot. We have developed an accurate analytical model for 802.15.4 CAP around the channel sensing probability [29], applicable in both saturated and nonsaturated scenarios, based on which we discuss how to compute the optimal operation point in the following.

B. Channel Utilization Analysis

As shown in Fig. 1, the 802.15.4 CAP activities define a 3-level renewal process. At the smallest time scale, the level-1 renew cycle Z_t is formed in accordance with the reset of the backoff parameters. The reset could be due to the backoff stage being exhausted, defining a Z_t^1 type cycle; or due to a transmission trial, defining a Z_t^2 type cycle. Over a larger time scale, the end of each transmission trial represents a renewal point of the level-2 renewal process. A level-2 cycle Z_b can be further classified into Z_b^1 and Z_b^2, depending on whether the transmission trial is successful or collided. At the largest time scale, the level-3 renewal cycle Z_s is the period from the end of a successful transmission to the end of next successful transmission.

1) The Single-Sensing Case: For an easier presentation, we first analyze the optimal operation point for the single-sensing case, where the transmission is activated after a single channel clear assessment (CCA). The result is then extended to the practical case of two CCAs.

Consider a single-hop 802.15.4 network with N nodes. Let J denote the number of sensing attempts in a Z_t cycle, and φ the sensing attempt rate at a slot. According to the renewal award theorem [30], $\varphi = \overline{J}/\overline{Z_t}$, which can also be interpreted as the steady-state probability that a node starts sensing at a certain slot, termed as the *channel sensing probability*. Use θ to denote the *channel sensing failure probability*, defined as the probability of sensing a busy channel. In the case of single CCA, the average number of sensing attempts for one node in a level-l renewal cycle can be expressed as

$$\overline{J} = (1-\theta) + 2\theta(1-\theta) + \cdots + (M-1)\theta^{M-2}(1-\theta)$$
$$+ M\theta^{M-1} = \sum_{m=0}^{M-1} \theta^m = \frac{1-\theta^M}{1-\theta} \qquad (11)$$

where M denotes the maximum number of backoff stages allowed. We can then have

$$\varphi = \frac{\overline{J}}{\overline{Z_t}} = \frac{1-\theta^M}{\overline{Z_t}(1-\theta)}. \qquad (12)$$

Furthermore, by analyzing the relationship between sensing failure events and the channel status, we have obtained the relationship [29]

$$\theta = \frac{T_s(1-(1-\varphi)^N)}{1+T_s(1-(1-\varphi)^N)} \qquad (13)$$

where T_s denotes the MAC-layer packet transmission time, which equals the collision time T_c in 802.15.4 CAP.

Given the sensing failure probability θ, according to Fig. 1, each level-1 renewal cycle Z_t contains a transmission with probability $p_x = (1-\theta^M)$. Thus, the number of level-1 cycles contained in a level-2 cycle Z_b follows a geometric distribution with parameter p_x, and the average length of a Z_b cycle is

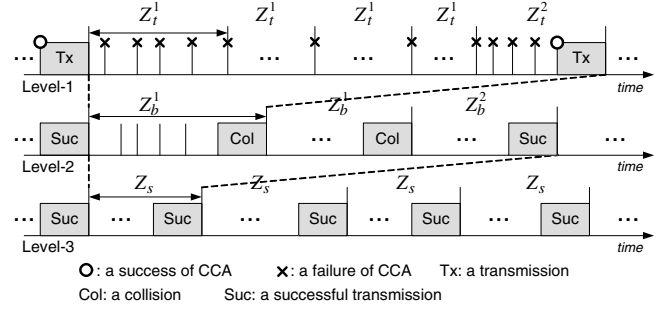$$\overline{Z_b} = \frac{1}{p_x}\overline{Z_t} = \frac{\overline{Z_t}}{1-\theta^M}. \qquad (14)$$

Fig. 1. The 3-level renewal process, with 5 backoff stages allowed.

In particular, the transmission trail from a tagged node will be successful when none of the other $N-1$ nodes starts sensing the channel in the same slot as the tagged node does, which determines the conditional successful transmission probability for a tagged node as $p_{cs} = (1-\varphi)^{N-1}$. Similar to the relationship between level-1 and level-2 cycles, the number of level-2 cycles contained in a level-3 cycle Z_s follows a geometric distribution with parameter P_{cs}, which gives the average length of a Z_s cycle as

$$\overline{Z_s} = \frac{1}{p_{cs}}\overline{Z_b} = \frac{\overline{Z_b}}{(1-\varphi)^{N-1}}. \qquad (15)$$

Combining (12), (14), and (15), we obtain

$$\overline{Z_s} = \frac{\overline{Z_t}}{(1-\theta^M)(1-\varphi)^{N-1}} = \frac{1}{\varphi(1-\theta)(1-\varphi)^{N-1}} \qquad (16)$$

By further applying the result in (13), $\overline{Z_s}$ can be expressed as a function of φ

$$\overline{Z_s} = \frac{1+T_s(1-(1-\varphi)^N)}{\varphi(1-\varphi)^{N-1}}. \qquad (17)$$

Considering that there is only one successful transmission in a Z_s cycle, the throughput of an individual node is

$$\eta_s = \frac{T_s}{\overline{Z_s}} = \frac{T_s\varphi(1-\varphi)^{N-1}}{1+T_s(1-(1-\varphi)^N)} \qquad (18)$$

The total MAC throughput (i.e., the channel utilization) can be computed as

$$R_s = N\eta_s = \frac{NT_s\varphi(1-\varphi)^{N-1}}{1+T_s(1-(1-\varphi)^N)} \qquad (19)$$

According to the renewal process given in Fig.1, the channel busyness ratio can be obtained as

$$R_b = \frac{NT_s}{\overline{Z_b}} = \frac{R_s}{p_{cs}} = \frac{NT_s\varphi}{1+T_s(1-(1-\varphi)^N)} \qquad (20)$$

2) The Double-Sensing Case: Now consider the channel access mechanism with two CCAs, which is specified in the standard [1]. The analysis for the double-sensing case can be conveniently achieved by appropriately modifying those equations regarding θ and φ derived in the single-sensing case. For differentiation, we add the superscript " ′ " to all the relevant variables for the double-sensing case.

Let p_1 denote the probability of sensing a busy channel in the first CCA, and p_2 the probability of sensing a busy

channel in the second CCA given that the channel is idle in the first CCA. Since a sensing failure event will happen when the channel is sensed busy in either of the two CCAs, the sensing failure probability can then be expressed as

$$\theta' = p_1 + (1 - p_1)p_2. \qquad (21)$$

It is not difficult to understand that the relationship between the number of sensing attempts and the sensing failure probability maintain the same as that in the single-sensing case, that is, $\overline{J'} = \sum_{m=0}^{M-1}(\theta')^m$. The channel sensing probability in the double-sensing case, which indicates the probability that the first CCA starts from a given slot, maintains the expression as $\varphi' = \overline{J'}/\overline{Z'_b}$.

With θ' and φ' in place, we have demonstrated in [29] that the relationship of (13) still applies in the double-sensing case. Based on the concept of 3-layer renewal process, it can be seen that the expressions (14) to (17) are valid also in the double-sensing case. Due to all these relationships inherited from the single-sensing case, the expressions of R_s in (19) and R_b in (20) also apply to the double-sensing case. That is, both the single and double sensing cases have the same expressions for R_s and R_b, respectively, so we drop the superscript "'" in the optimal operation point analysis given below.

C. Optimal Operation Point Analysis

With R_s being expressed as a function of φ in (19), we can compute the optimal φ^* to maximize R_s by

$$\frac{d}{d\varphi}R_s|_{\varphi=\varphi^*} = 0 \qquad (22)$$

which then gives

$$N\varphi^*(1 + T_s) = 1 + T_s(1 - (1 - \varphi^*)^N) \qquad (23)$$

after some manipulation. The value of φ^* can be numerically solved from (23). Applying the φ^* to (19) and (20), we can then obtain the maximum channel utilization R_s^* and the optimal channel busyness ratio R_b^*.

VI. PERFORMANCE EVALUATION

In this section, we apply the VS S-G analysis to compute the optimal operation point for both 802.11 DCF and 802.15.4 CAP. The accuracy of the VS S-G analysis is validated through comparison with the protocol-specific results obtained according to Section IV and Section V. In all the examples, we use Maple [31] to obtain the numerical results. The accuracy of the protocol-specific analysis is illustrated by computer simulations.

A. VS S-G Analysis over 802.11 DCF

Consider a single-hop 802.11 DCF network configured with the parameters given in Table II of [10]. The successful transmission time T_s and collision transmission time T_c are determined according to the standard [10], for both cases with and without RTS/CTS, which will be used in both the VS S-G analysis and the protocol-specific analysis.

The maximum channel utilization R_s^* and the associated optimal channel busyness ratio R_b^* has been computed in [10]
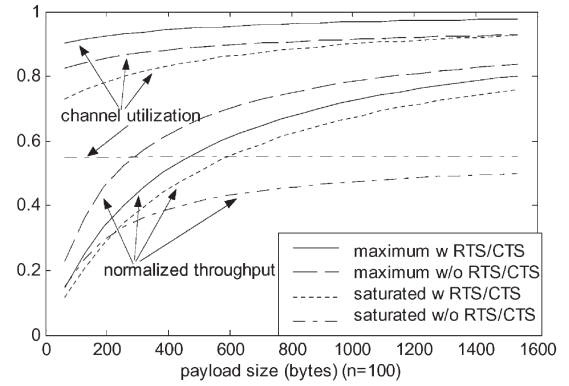

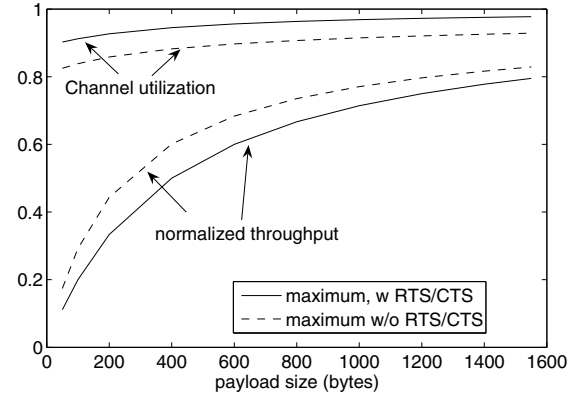Fig. 2. Optimal operation points based on 802.11 specific analysis [10].

Fig. 3. Optimal operation points based on the VS S-G analysis.

(as summarized in Section IV), where computer simulation results are also presented to validate the analysis accuracy. In [11], R_b^* is applied for admission control; the capability of identifying the maximum admission region, demonstrated through simulations, further justify the accuracy of the optimal operation point.

Fig. 2 presents the maximum channel utilization R_s^* and the maximum normalized throughput V^* obtained through protocol specific analysis with number of nodes $N = 100$, which is a replot of the Fig. 3 in [10]. As a comparison, R_s^* and V^* are computed by VS S-G analysis and presented in Fig. 3. Comparing the two figures, we can see that the VS S-G analysis almost exactly matches the 802.11 specific analysis. Such a good match is not surprising, because 802.11 DCF incorporates independent channel access, which results in a good Poisson approximation of the packet arrival process in a slot, when N is large.

B. VS S-G analysis over 802.15.4 CAP

The 802.15.4 CAP is a CSMA/CA protocol with correlated channel access, where the Poisson approximation accuracy is impacted by the amount of covariance among the different channel access processes, according to the proof of Theorem 1 given in Appendix A. Thus, the VS S-G analysis for the 802.15.4 CAP may not match the protocol-specific analysis as close as that for the 802.11 DCF.

The optimal operation point analysis presented in Section V is based on the 802.15.4 MAC model developed in [29], so its accuracy is inherently determined by the accuracy of the underpinning MAC model. We have thoroughly validated the preciseness of that 802.15.4 MAC model in [29]. Here, we focus on the comparison between the VS S-G analysis and the protocol-specific benchmark.

Fig. 4 presents the optimal operation points R_s^* and R_b^* that are computed from protocol-specific analysis and VS S-G analysis, respectively. Note that the protocol-specific optimal point can be computed, given N and T_s according to Section V-C. The value of T_s is given in terms of physical slots. From Fig. 4, we can have the following observations:

1) Over a wide range of N values, the VS S-G analysis maintains close to the protocol-specific analysis, whereas the gap is due to the correlated channel access.
2) The optimal channel busyness ratio under the protocol-specific analysis is basically independent of the N value, as demonstrated for 802.11 DCF in [10].
3) The maximum channel utilization R_s^* based on the protocol-specific analysis converges to a constant value when N becomes large, demonstrating the effectiveness of the Central Limit Theorem as described in Theorem 1.
4) Regarding R_s^*, it appears that the gap between the VS S-G analysis and the protocol-specific analysis increases versus N. For example, the gap between the two curves is $0.6493 - 0.6161 = 0.0332$ at $N = 5$, at increases to $0.6493 - 0.5918 = 0.0575$ at $N = 60$. However, if we consider the per-node gaps $0.0332/5$ and $0.0575/60$, not only is the gap negligibly small, it also reduces quickly with N. The observations again demonstrate the effectiveness of the Central Limit Theorem.
5) The R_b^* based on the VS S-G analysis is smaller than that based on the protocol-specific analysis, while the VS S-G based R_s^* is larger. A slightly smaller estimate of R_b^* is particularly meaningful for the call admission control. As to be discussed in Section VII, a smaller R_b^* will result in a smaller admission region. A slightly conservative admission region is always preferred than an aggressive one in practice. The former strategy can avoid overusing the resource for guaranteed QoS; while the latter one may result in QoS collapse for all the nodes already in the network, when a new comer drives the total traffic load exceeding the network capacity.

We also evaluate the optimal operation points versus the packet payload. For convenience, we plot the R_b^* and R_s^* versus the packet transmission time T_s in Fig. 5. We can see that over the wide range of transmission times, the VS S-G analysis stays close to and traces the protocol-specific analysis, yet with a gap determined by the channel access correlations. The R_b^* based on VG S-G analysis is an conservative estimation compared to the benchmark values in most of the scenarios, except when $T_s < 3$ physical slots. However, the violation of conservativeness under very small T_s values, which may hurt the admission control as discussed above, can be prevented
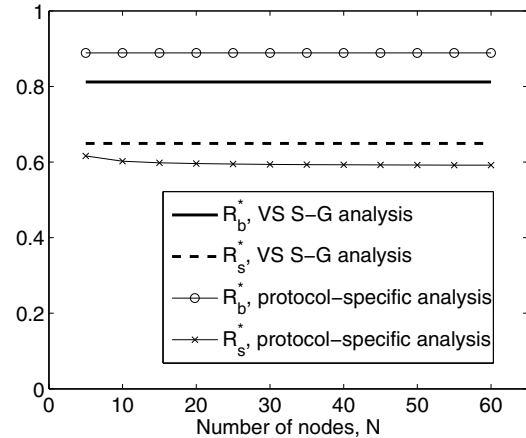

Fig. 4. The VS S-G analysis compared to the protocol-specific analysis versus number of nodes, $T_s = 8$ physical slots.
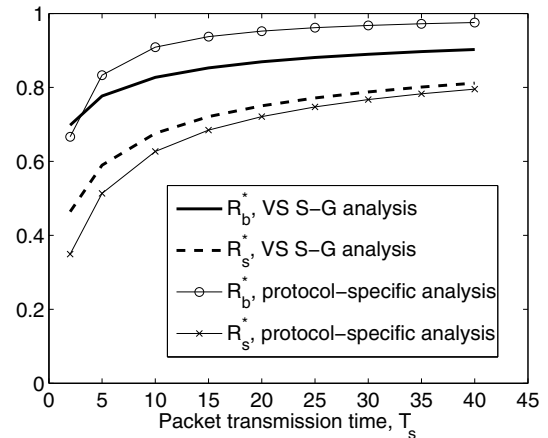

Fig. 5. The VS S-G analysis compared to the protocol-specific analysis versus packet transmission time, $N = 50$.

in practice by configuring a lower bound of the MAC-layer packet size. In addition, Fig. 5 shows that both R_b^* and R_s^* increase with the packet transmission time, equivalently, the payload size. Such a relationship in fact applies to all the CSMA/CA protocols, because in each renewal cycle including a successful transmission or a collision, the portion occupied by the idling slots reduces against the increased payload size.

VII. ADMISSION CONTROL

In [10], [11], it is shown that the maximum DCF MAC capacity and satisfying QoS performance can only be achieved when the network operation range is constrained below the optimal CBRO. Such observations motivate applying the optimal operation point in call admission control to ensure the maximum admission region [11].

Considering an IP application, each node at the network layer is normally modeled as a G/G/1 queue. It has been further demonstrated in [10], [11] that, when the optimal operation point constraint is applied, each packet sees an approximately constant service rate; therefore the G/G/1 queue
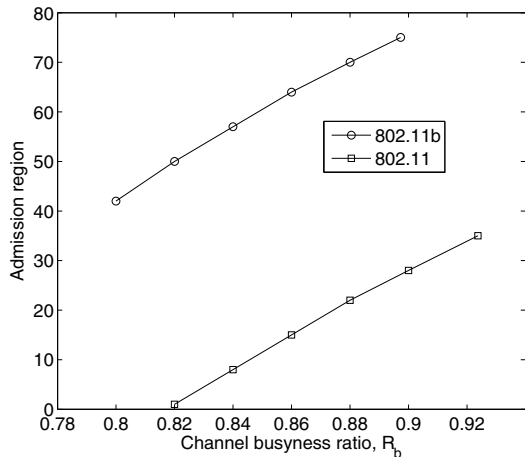
Fig. 6. The admission region versus the channel busyness ratio for on/off flows.

at each node can be well approximated by a G/D/1 queue. Inspired by the above results, we present in [16] that a *cross-layer analytical framework* can be formed by integrating three pillar components: i) the queueing analysis at the network layer to determine the effective bandwidth [32] for packet-level QoS guarantee, ii) the non-saturated MAC modeling to determine the serving capacity provisioned by the MAC layer to the network layer, and iii) an optimal operation point based admission control applied to validate the G/D/1 approximation and ensure the maximum admission region.

The main issue associated with the protocol-specific computation of the optimal operation point is that modeling MAC details is normally not an easy task, particularly when the channel access processes from different nodes are not independent as shown in Section VI. For a new CSMA/CA implementation, without computing the optimal operation point first, the cross-layer analytical framework could not be applied. The simplicity and the generality of the VS S-G analysis significantly extend the applicability of the cross-layer framework for admission control.

When the optimal CBRO R_b^* is applied to identify the admission region in accordance with the cross-layer framework, it is meaningful to evaluate the sensitivity of the admission region regarding R_b, because the VS S-G based operation point may be conservative estimation of the real value as discussed in VI-B. Since applying the cross-layer framework over the 802.15.4 MAC is still an open issue, we here use the 802.11 DCF as an example to illustrate the sensitivity issue.

The efficiency of the proposed cross-layer framework has been demonstrated in [16] for 802.11 DCF capacity analysis, with on/off voice flows considered. We follow all the traffic and MAC configurations given in [16]. In particular, the 802.11 DCF parameters are as listed in Table I of [16], and we here consider two standards, 802.11 with data rate of 2 Mbps and 802.11b with data rate of 11 Mbps. The "admission region versus R_b^*" curves are plotted in Fig. 6, which are computed by setting the channel busyness ratio incorporated in the cross-

layer framework to different values. From the figure, we can see that the admission region is sensitive to R_b, basically following a linear relationship. Such sensitivity indicates that precise computation of the optimal operation point is important for computing an accurate call-level admission region. For 802.15.4, as the correlation in channel access results in a constant gap in estimating the optimal operation point, a compensation factor may by applied to the VS S-G result for admission control, which will be investigated in our future study.

VIII. CONCLUSION

In this paper, we develop a virtual slot based S-G analysis to compute the optimal operation point for the family of CSMA/CA protocols. The optimal operating point is defined as the optimal channel busyness ratio over which the maximum channel utilization is achieved. The VS S-G approach is protocol independent by exploiting the facts that all the CSMA/CA protocols can be described by a virtual slot system, disregard of the implementation details, and the total packet arrivals in a slot can be modeled or approximated by a Poisson random variable when the number of nodes is large. The VS S-G analysis is benchmarked against precise 802.11 DCF analysis and 802.15.4 CAP analysis. Moreover, the efficiency in applying the optimal operation point to obtain the maximum admission region is discussed, with 802.11 DCF as an example. For the future work, the optimal operation point based admission control over 802.15.4 is to be thoroughly investigated.

APPENDIX A
PROOF OF THEOREM 1

We can determine the optimal operation point by considering the homogeneous MAC, where all the nodes are equipped with the same backoff configurations and thus stochastically share the same class of QoS. The service differentiation schemes applied in a heterogeneous multiclass MAC essentially define the priority of each service class to access the channel, which will not change the maximum channel utilization in the scenario with infinite user nodes.

By modeling the CSMA/CA MAC as a virtual-slot system, in steady state, the backoff mechanism can be summarized into a channel access probability, say τ for each node. Let X_i denote the number of arrivals from node i in a slot, which should be a Bernoulli random variable with mean $E[X_i] = \tau$ and variance $\text{Var}(X_i) = \tau(1 - \tau)$. When there are N nodes, the variance of the total arrivals to a slot is

$$\text{Var}\left(\sum_{i=1}^{N} X_i\right) = \text{Cov}\left(\sum_{i=1}^{N} X_i, \sum_{j=1}^{N} X_j\right) \quad (24)$$

$$= \sum_{i=1}^{N} \text{Var}(X_i) + 2\sum_{i=1}^{N}\sum_{j<i} \text{Cov}(X_i, X_j). \quad (25)$$

A formal definition of the weak correlation/dependence is to describe $\{X_i\}$ as a sequence of m-dependent random variables [33], [34], where $X_1, X_2, \cdots, X_s$ are independent of X_t,

$X_{t+1}, \cdots$, provided $t - s > m$. When N goes to infinity, if the traffic load maintains constant as $G = N\tau$ and the correlation among the random variables X_i is upbounded as

$$\lim_{N \to \infty} 2 \sum_{i=1}^{N} \sum_{j < i} \mathrm{Cov}(X_i, X_j) = K$$

$$< \lim_{N \to \infty} \sum_{i=1}^{N} \mathrm{Var}(X_i) = \lim_{N \to \infty} N\tau \left(1 - \frac{G}{N}\right) = G, \quad (26)$$

the central limit theorem for m-dependent random variables, with m fixed [34] or unbounded [35], tells that $\sum_{i=1}^{N} X_i$ converges to a Gaussian random variable with mean G and variance $G + K$.

To investigate the Poisson approximation of the aggregate arrival $\sum_{i=1}^{N} X_i$, consider a Poisson random variable with parameter G. Since $G = N\tau$, the Poisson random variable can be equivalently considered as the summation of N independently and identically distributed (i.i.d.) Poisson random variables with parameter τ. Use Y_i to denote the ith Poisson random variable with parameter τ. The central limit theorem for i.i.d. random variables tells that the distribution of $\frac{\sum_{i=1}^{N} Y_i - N\tau}{\sqrt{N\tau}}$ tends to the standard Gaussian distribution as $N \to \infty$. Thus, $\sum_{i=1}^{N} Y_i$ tends to Gaussian with mean G and variance G, considering that $G = N\tau$.

Based on the above analysis, we can see that a Poisson random variable with the same traffic load can well approximate the distribution of the actual packet arrivals in a slot, according to the central limit theorem with N going to infinity. The approximation is conservative in the sense that the variance value $G + K$ is underestimated as G. The approximation accuracy increase when K becomes smaller, i.e., when the correlation among the channel access processes become weaker.

REFERENCES

[1] *Wireless Medium Access Control (MAC) and Physical Layer (PHY) specifications for low-rate wireless personal area networks (LR-WPANs)*, IEEE Std. 802.15.4 Std., 2003.

[2] *Wireless Medium Access Control (MAC) and Physical Layer (PHY) specifications*, IEEE Std. 802.11 Std., 1999.

[3] *Standard Specification for Telecommunications and Information Exchange between Roadside and Vehicle Systems — 5GHz Band Dedicated Short Range Communications (DSRC) Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, ASTM e2213-03, Sept. 2003.

[4] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 535–547, Mar. 2000.

[5] F. Cali, M. Conti, and E. Gregori, "Dynamic tuning of the IEEE 802.11 protocol to achieve a theoretical throughput limit," *IEEE/ACM Trans. Networking*, vol. 8, no. 6, pp. 785–799, Dec. 2000.

[6] M. Heusse, F. Rousseau, G. Berger-Sabbatel, and A. Duda, "Performance anomaly of 802.11b," in *Proc. IEEE INFOCOM*, 2003, pp. 836–843.

[7] O. Tickoo and B. Sikdar, "Queueing analysis and delay mitigation in IEEE 802.11 random access MAC based wireless networks," in *Proc. IEEE INFOCOM*, 2004, pp. 1404–1413.

[8] K. Medepalli, F.A. Tobagi, "System centric and user centric queueing models for IEEE 802.11 based wireless LANs," in *Proc. IEEE Broadnets*, 2005, pp. 612-621.

[9] A. Sheikh, T. Wan, and Z. Alakhdhar, "A unified approach to analyze multiple access protocols for buffered finite users," *Journal of Network and Computer Applications*, vol. 27, no. 1, pp. 49–76, Jan. 2004.

[10] H. Zhai, X. Chen and Y. Fang, "How well can the IEEE 802.11 wireless LAN support quality of service?" *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 3084–3094, Nov. 2005.

[11] H. Zhai, J. Wang, and Y. Fang, "Providing statistical QoS guarantee for voice over IP in the IEEE 802.11 wireless LANs," *IEEE Wireless Commun.*, vol. 13, no. 1, pp. 36–43, Feb. 2006

[12] A. Abdrabou and W. Zhuang, "Service time approximation in IEEE 802.11 single-hop ad hoc networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 1, pp. 305-313, Jan. 2008.

[13] X. J. Dong and P. Varaiya, "Saturation throughput analysis of IEEE 802.11 wireless LANs for a lossy channel," *IEEE Commun. Lett.*, vol. 9, no. 2, pp. 100–102, Feb. 2005.

[14] D. Bertsekas and R. Gallenger, *Data Networks*, 2nd Edition, Prentice Hall, 1992.

[15] H. H. Tan and K. Tsai, "Packet output processes of CSMA and CSMA/CD protocols," *IEEE Trans. Commun.*, vol. 44, no. 4, pp. 464–474, Apr. 1996.

[16] Y. Cheng, X. Ling, W. Song, L. X. Cai, W. Zhuang, and X. Shen, "A cross-layer approach for WLAN voice capacity planning," *IEEE J. Select. Areas Commun.*, vol. 25, no. 4, pp. 678–688, May 2007.

[17] F. Alizadeh-Shabdiz and S. Subramaniam, "Analytical models for single-hop and multi-hop Ad Hoc networks," *Mobile Networks and Applications*, vol. 11, no. 1, pp. 75-90, Feb. 2006.

[18] G. R. Gantieni, Q. Ni, C. Barakat, and T. Turletti, "Performance analysis under finite load and improvements for multirate 802.11," *Computer Communications*, vol. 28, no. 10, pp. 1095–1109, Jun. 2005.

[19] O. Tickoo and B. Sikdar, "A queueing model for finite load IEEE 802.11 random access MAC," in *Proc. IEEE ICC*, 2004, pp. 175–179.

[20] K. Medepalli and F. A. Tobagi, "Towards performance modeling of IEEE 802.11 based wireless networks: a unified framework and its application," in *Proc. IEEE INFOCOM*, 2006.

[21] 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*, 2005, pp. 1550–1561.

[22] L. X. Cai, X. Shen, J.W. Mark, L. Cai and Y. Xiao, "Voice capacity analysis of WLAN with unbalanced traffic," *IEEE Trans. Veh. Technol.*, vol. 55, no. 3, pp. 752–761, May 2006.

[23] S. Pollin *et al.*, "Performance analysis of slotted IEEE 802.15.4 medium access layer," Interuniversity Micro-Electronics Center, Belgium, Tech. Rep., Sept. 2005.

[24] Z. Tao, S. Panwar, D. Gu, and J. Zhang, "Performance analysis and a proposed improvement for the IEEE 802.15.4 contention access period," in *Proc. IEEE WCNC*, 2006, pp. 1811–1818.

[25] Y. Xiao, "Performance analysis of priority schemes for IEEE 802.11 and IEEE 802.11e wireless LANs," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1506–1515, Jul. 2005.

[26] G. Bianchi, I. Tinnirello, and L. Scalia, "Understanding 802.11e contention-based prioritization mechanisms and their coexistence with legacy 802.11 stations," *IEEE Network*, vol. 19, no. 4, pp. 28–34, Jul./Aug. 2005

[27] K. Kim and G. Shevlyakov, "Why gaussianity?" *IEEE Signal Processing Mag.*, vol. 25, no. 2, pp. 102–113, Mar. 2008.

[28] I. Ramachandran, A.K. Das, and S. Roy, "Analysis of the contention access period of IEEE 802.15.4 MAC," University of Washington, Tech. Rep. UWEETR-2006-0003, Feb. 2006.

[29] X. Ling, Y. Cheng, J. W. Mark, and X. Shen, "A renewal theory based analytical model for the contention access period of IEEE 802.15.4 MAC," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2340-2349, Jun. 2008.

[30] R. Nelson, *Probability, Stochastic Processes, and Queueing Theory*, Springer-Verlag, 1995.

[31] A. Heck, *Introduction to Maple (3rd ed.)*. Springer-Verlag, New York, 2003.

[32] F. P. Kelly, "Notes on effective bandwidth," in *Stochastic Networks: Theory and Applications*, F.P. Kelly, S. Zachary, and I. Ziedins, Eds. Oxford, U.K.: Oxford Univ. Press, 1996, pp. 141-168.

[33] J. Davidson, *Stochastic Limit Theory — An Introduction for Econometricians*. Oxford, U.K.: Oxford Univ. Press, 1994.

[34] S. A. Orey, "Central limit theorem for m-dependent random variables," *Duke Math. J.*, vol. 25, pp. 543-546, 1958.

[35] K. N. Berk, "A central limit theorem for m-dependent random variables with unbounded m," *The Annals of Probability*, vol. 1, no. 2, pp. 352-354, 1973.