

A Fast and Accurate Performance Analysis of Beaconless IEEE 802.15.4 Multi-Hop Networks[☆]

Rachit Srivastava^a, Sanjay Motilal Ladwa^a, Abhijit Bhattacharya^a, Anurag Kumar^a

^a*Dept. of Electrical Communication Engineering
Indian Institute of Science, Bangalore, 560012, India*

Abstract

We develop an approximate analytical technique for evaluating the performance of multi-hop networks based on beaconless IEEE 802.15.4 (the “ZigBee” PHY and MAC), a popular standard for wireless sensor networks. The network comprises sensor nodes, which generate measurement packets, relay nodes which only forward packets, and a data sink (base station). We consider a detailed stochastic process at each node, and analyse this process taking into account the interaction with neighbouring nodes via certain time averaged unknown variables (e.g., channel sensing rates, collision probabilities, etc.). By coupling the analyses at various nodes, we obtain fixed point equations that can be solved numerically to obtain the unknown variables, thereby yielding approximations of time average performance measures, such as packet discard probabilities and average queueing delays. The model incorporates packet generation at the sensor nodes and queues at the sensor nodes and relay nodes. We demonstrate the accuracy of our model by an extensive comparison with simulations. As an additional assessment of the accuracy of the model, we utilize it in an algorithm for sensor network design with quality-of-service (QoS) objectives, and show that designs obtained using our model actually satisfy the QoS constraints (as validated by simulating the networks), and the predictions are accurate to well within 10% as compared to the simulation results in a regime where the packet discard probability is low.

Keywords: wireless sensor networks, performance analysis of multi-hop CSMA networks, design of multihop CSMA networks

1. Introduction

Wireless sensor networks (WSNs), a concept that originated in the mid-1990s, have now reached a stage in their evolution that we are beginning to see their actual deployment. It is not unreasonable to expect that in 10-15 years the world will be covered with wireless sensor networks with access to them via the so called Internet of Things (IoT)[1]. IEEE 802.15.4 [2] is a popular standard for the physical layer and medium access control for low-power

[☆]A preliminary version of this work, entitled *Performance Analysis of Beacon-less IEEE 802.15.4 Multi-hop Networks*, was presented in the Fourth International Conference on Communication Systems and Networks (COMSNETS), Bangalore, 2012. This work was supported by a research grant from the Department of Electronics and Information Technology (DeitY), Government of India, through the Automation Systems Technology (ASTE) program, by a DeitY-NSF funded Indo-US project on Wireless Sensor Networks for Protecting Wildlife and Humans, and by a Department of Science and Technology J.C. Bose National Fellowship.

Email addresses: rachitsri@gmail.com (Rachit Srivastava), sanjofpesit@gmail.com (Sanjay Motilal Ladwa), abhijit@ece.iisc.ernet.in (Abhijit Bhattacharya), anurag@ece.iisc.ernet.in (Anurag Kumar)

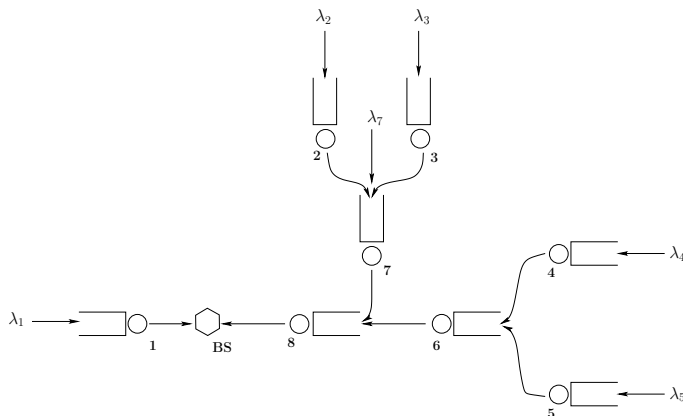


Figure 1: A WSN consisting of nodes arranged in a tree topology. Source node i has packet generation rate λ_i . Relay nodes 6, 7 and 8 forward the data of their predecessors to the base-station, BS.

wireless sensor networks. With the growing importance of wireless sensor networks in industrial applications [3], we need analysis and design techniques for multi-hop IEEE 802.15.4 networks. The standard uses CSMA/CA for medium access control, and defines two types of CSMA/CA algorithms - beamed (or slotted) and beaconless (or unslotted), respectively (see [2] for details). In this paper we develop a new approximate analytical technique for multi-hop beaconless IEEE 802.15.4 networks, and demonstrate the usefulness of such an analytical tool in designing multi-hop networks with Quality of Service (QoS) objectives.

Figure 1 depicts the queueing schematic of a network of the type that we are concerned with in this paper. There are sensor nodes, namely, 1, 2, 3, 4, 5 and 7, that generate measurement packets according to independent point processes at rate λ_i that need to be transported to the base station (BS). The sensor nodes can also serve as *relays* for the traffic of other sensor nodes, e.g., Node 7 is a relay for nodes 2 and 3. Two additional nodes, 6 and 8, serve only as relay nodes, due to the limited range of the radios associated with the sensor nodes. The nodes use unslotted CSMA/CA to contend for the wireless medium and transmit their packets. Even the stability analysis of such networks is a difficult problem (even for networks with single hop traffic; see [4] for a detailed discussion). In principle, the entire network can be modeled via a large coupled Markov chain, with the state at each node being the number of packets in its queue, and the state of the contention process for the head-of-the-line (HOL) packet. Such an approach is well known to be intractable even for a network with a single contention domain (“no hidden nodes” network) and saturated queues (see [5, 6]). Thus all researchers have taken recourse to developing approximate analyses. Table 1 summarizes some of the recent work on modeling approaches for beaconless IEEE 802.15.4 networks where we have listed the major limitations of the proposed models as well. Owing to the need for low power operation and large coverage, in general, a sensor network is multi-hop with the presence of hidden nodes. Several techniques have been proposed in the literature in the form of MAC protocols to avoid the hidden terminal problem; for example, the dual busy-tone multiple access (DBTMA) scheme proposed in [7], and the floor acquisition multiple access (FAMA) discipline proposed in [8]. See [7] and the references therein for details of these techniques. However, none of these techniques have been part of any commercial implementation of IEEE 802.15.4 CSMA/CA, the most common MAC protocol for wireless sensor

networks. Thus, it is important to be able to predict the performance of IEEE 802.15.4 CSMA/CA in the presence of multi-hopping, and hidden terminals. Further, for networks that carry measurement traffic, the queue occupancies vary with time. Unlike other models, we consider all of these attributes in our modeling approach.

Our Contributions and Comparison with Related Work

We consider a multi-hop WSN consisting of static sources and relays arranged in a tree topology operating in the beacon-less mode with acknowledgements (ACKs). Each node has an infinite buffer space and may operate in the saturated or unsaturated regime with fixed packet length. Different analysis techniques are developed for networks with hidden nodes and networks without hidden nodes due to the difference in channel activity durations perceived by a node in the presence and absence of hidden nodes. Under certain approximations, we model the stochastic process evolving at a node by incorporating the influence of the other nodes in the network by their (unknown) time averaged statistics, and then couple these individual node processes via a system of fixed point equations, which is solved using an iterative scheme to obtain the unknown variables. Although this *decoupling* (or *mean-field*) approximation is popular in such situations, our more detailed model incorporating several issues not considered together hitherto requires a careful handling of the analysis. We identify and calculate two QoS measures for each source node, namely, the packet delivery probability and end-to-end packet delay, in terms of these variables. We observe that in a multi-hop IEEE 802.15.4 network, the packet delivery probability falls sharply with increase in packet arrival rates even before the end-to-end packet delay becomes substantial.

The rest of the paper is organized as follows: In Section 2, we give an overview of the unslotted CSMA/CA mechanism. Section 3 explains the node behaviour and elaborates upon the different analysis techniques used for networks without hidden nodes and networks with hidden nodes. Section 4 is devoted to detailed derivation of the system of fixed point equations. Section 5 describes the iterative scheme used to solve the system of fixed point equations, and obtain the performance measures. In Section 6, we provide a discussion on the validity of the fixed point approach. Section 7 compares our analysis with simulations. Section 8 demonstrates how the analysis can be used for designing multi-hop networks with QoS objectives. Finally, we conclude the paper in Section 9.

2. Unslotted CSMA/CA for Beacon-less IEEE 802.15.4 Networks

A typical node behaviour under unslotted CSMA/CA is shown in Figure 2. A node with an empty queue remains idle until it generates a packet or receives one from its predecessor nodes. When a node has data to send (i.e., has a non-empty queue), it initiates a random back-off with the first back-off period being sampled uniformly from 0 to $2^{macminBE} - 1$, where *macminBE* is a parameter fixed by the standard. For each node, the back-off period is specified in terms of slots where a slot equals 20 symbol times (T_s), and a symbol time equals 16 μs .¹ The node then performs a CCA (*Clear Channel Assessment*) to determine whether the channel is idle. If the CCA succeeds, the

¹Note, however, that there is no central coordinator that maintains synchrony of the slots *across the nodes*, and hence the CSMA/CA protocol is unslotted.

Table 1: A comparison of the literature on analytical performance models of CSMA/CA based networks. The bold text in the “Network Scenario and Limitations” columns highlights the limitations of the model that we are able to handle in our model.

Authors	Network Scenario & Limitations	Modeling Approach
Kim et al. (2006)[9]	beaconless 802.15.4; rare packet arrivals, star network; no hidden nodes; no queueing at nodes ; with ACKs	2-dim discrete time Markov chain for each node with states: (backoff stage, backoff counter) while in backoff, and separate states when transmitting and idling; with unknown transition probabilities; coupling the per node models to get the unknowns via fixed point approach
Qiu et al. (2007)[10]	802.11 DCF; (un)saturated, star network ; with hidden nodes and ACKs; infinite buffer space	discrete time Markov chain with states as the set of transmitting nodes having unknown transition probabilities, together with RSS measurements between each pair of node to obtain CCA and packet failure probabilities; iterative procedure for solving unknowns
Kim et al. (2008)[11]	beaconless 802.15.4; star topology; no hidden nodes; no collisions ; infinite buffer space	unsaturated analysis of single hop network by simple mathematical model of M/G/1 queue; no collision is considered; computes QoS as packet delay and packet loss probability; computes battery life time
Singh et al. (2008)[12]	beaconed 802.15.4; (un)saturated, star network; no hidden nodes ; with ACKs; infinite buffer; equal traffic arrival rate at all nodes	Markov renewal process at each node with cycle length dependent on the number of nodes available to attempt in the first backoff period of the cycle; transition probabilities and conditional expectations using node decoupling; coupling via stationary probabilities of remaining nodes; saturation analysis results used for the unsaturated case
Buratti-Verdone (2009)[13]	beaconless 802.15.4; star topology; no hidden nodes; no ACKs and retransmissions; no queueing	bidimensional process with backoff counter and backoff stage at time t ; evaluation of sensing, transmission and success probability; computes mean energy spent by each node for a transmission
He et al. (2009)[14]	beaconed 802.15.4; saturated, star network; no hidden nodes ; with ACKs;	two Markov chains for each node: one embedded at the end of transmissions in the common channel, and the other during backoffs embedded at each slot; both having states as (backoff stage, backoff counter); chains coupled to get a fixed point approach
Martalo et al. (2009)[15]	beaconed 802.15.4; unsaturated, star network; no hidden nodes; no ACKs; finite buffer space	two semi-Markov processes: one for tagged node having unknown transition probabilities, and the other for the shared radio channel that couples all the per node models to obtain the unknown quantities; MGFs used for finding queueing delays
Goyal et al. (2009)[16]	beaconless 802.15.4; unsaturated star network; no hidden nodes; no queueing at nodes ; with ACKs	tracked the number of nodes with non-empty queues assuming that the probability of m nodes having non-empty queues at any given time is same as the probability of $m - 1$ empty nodes getting a new packet to send while a non-empty node is sending its current packet (for coupling individual node models); assumed that the CCA attempt processes of other nodes with non-empty queues are Poisson processes (decoupling)
Lauwens et al. (2009)[17]	beaconless 802.15.4; saturated, star network; no hidden nodes; no ACKs	semi-Markov model for each node with unknown and non-homogeneous transition probabilities based on the backoff stage of the node; also found the distribution of backoff intervals and iterated over these together with the stationary probabilities of the semi-Markov processes
Jindal-Psounis (2009)[18]	802.11 DCF; unsaturated, multi-hop network; with hidden nodes and ACKs; infinite buffer space; modeling involves details that are specific to 802.11 such as the RTS/CTS mechanism ;	derived per-edge collision and idle probabilities in two-edge topologies in terms of expected service times; for a general topology, derived collision and idle probabilities for each edge in terms of those of neighboring two-edge topologies; using these probabilities, set up and solved a Markov chain model for each edge (with states as the current backoff window, backoff counter, and time since the last successful/unsuccessful RTS/CTS exchange) to express the expected service times of the edges via a set of fixed point equations, which were solved iteratively
DiMarco et al. (2010)[19]	beaconless 802.15.4; unsaturated, multi-hop network; no queueing at nodes ; with hidden nodes and ACKs	3-dim discrete time Markov chain for each node having state: (backoff stage, backoff counter, retransmission counter) with CCA and collision probabilities (unknown) as state transition probabilities (decoupling); assumed independence of node processes to find expressions for the CCA and collision probabilities using stationary probabilities of different Markov chains (coupling)
Sen-De (2010)[20]	beaconless 802.15.4; unsaturated, multi-hop network; with hidden nodes; no queueing; analysis does not match simulation well	1-dim Markov model for transmitting node with states being in CCA, backoffs states and data transmission; steady state transmission probability and probability of success (throughput) are obtained through fixed point iterations
Shyam-Kumar (2010)[21]	beaconless 802.15.4; (un)saturated, multi-hop network; with hidden nodes and ACKs; infinite buffer; analysis does not match simulation well	saturated analysis using continuous time Markov chain with states as the set of transmitting nodes having unknown transition probabilities; found CCA and failure probabilities using steady state probability of the Markov chain; unsaturated network modeled as a Markov renewal process constituted by the nodes having non-empty queues, saturation analysis results used in each cycle with the set of non-empty nodes
Marbach et al. (2011)[22]	non-adaptive CSMA characterized by a fixed attempt probability on each link , and a non-zero sensing period on each link; asynchronous, unsaturated, multihop network; over-simplified model - does not conform to any existing standard	intuitively formulated fixed point equations involving node idling probabilities, and transmission attempt rates; showed uniqueness and asymptotic accuracy for large networks with small sensing period and appropriately decreasing link attempt probabilities; for a given sensing period, characterized the achievable rate region by showing that for every arrival rate in that set, there exists a policy such that the link service rates obtained from the fixed point equations exceed the link arrival rates
Park et al. (2013)[23]	beaconed 802.15.4 ; unsaturated star network; no hidden nodes; no queueing at nodes except the coordinator ; with ACKs; equal traffic arrival rate at all nodes	the contention access scheme (CAP) is modeled via a 4-dim Markov chain for each device with states as {back-off stage, state of the back-off counter, state of retransmission counter, state of deferred transmission}; the GTS allocation scheme is modeled by a finite 1-dim Markov chain with state as the number of waiting requests of the coordinator; the two chains are coupled via the distribution of the number of GTS requests per superframe, which abstracts the CSMA/CA mechanism of the CAP

node does a *Rx-to-Tx turnaround*, which is 12 symbol times, and starts transmitting on the channel. If, on the other hand, the CCA fails, a new back-off process is started with the back-off exponent raised by one, i.e., to $macminBE+1$, provided that this is less than its maximum value, $macmaxBE$. The maximum number of successive CCA failures for the same packet is governed by $macMaxCSMABackoffs$, exceeding which the packet is discarded at the MAC layer. The standard allows the inclusion of acknowledgements (ACKs) which are sent by the intended receivers on a successful packet reception. Once the packet is received, the receiver performs a *Rx-to-Tx turnaround*, which is again 12 symbol times, and sends a 22 symbol fixed size ACK packet. A successful transmission is followed by an *InterFrame Spacing*(IFS) before sending another packet.

When a transmitted packet collides or is corrupted by the PHY layer noise, the ACK packet is not generated, which is interpreted by the transmitter as failure in delivery. The node retransmits the same packet for a maximum of $aMaxFrameRetries$ times before discarding it at the MAC layer. After transmitting a packet, the node turns to the Rx-mode and waits for the ACK. The $macAckWaitDuration$ determines the maximum amount of time a node must wait for in order to receive the ACK before concluding that the packet (or the ACK) has collided. The default values of $macminBE$, $macmaxBE$, $macMaxCSMABackoffs$, and $aMaxFrameRetries$ are 3, 5, 5, and 3 respectively.

3. Description of the Modeling Approach

As shown in Figure 1, the network consists of sensors, relays and a base-station (BS). We consider time invariant links and static (immobile) nodes. The links are characterized by a predefined target packet error rate (PER) (e.g., a PER of 1% on each link). We assume that the sensors generate traffic according to independent point processes. These constitute the aggregate external arrival process for the sensor network. Each node transmits its data to the next-hop node according to the topology; throughout this paper, *we shall work with a tree topology so that each transmitter has exactly one receiver node*. The intermediate nodes along a route (from a sensor to the base station) may be relays in which case they simply forward the incoming traffic, or they may be sensors which transmit their own packets as well as the received packets. Based on the network congestion, the nodes may discard packets due to consecutive failed CCAs or frame retries. Figure 2 shows a typical sample path of the process evolution at a node.

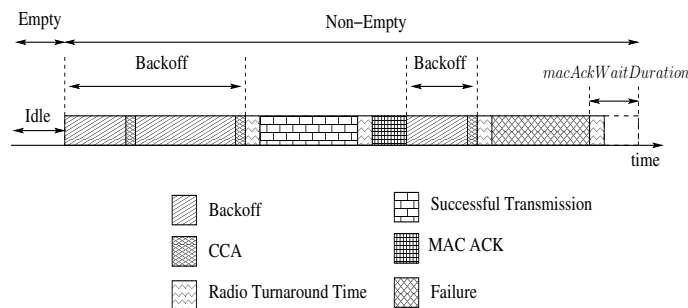


Figure 2: Process evolving at a tagged node as it goes through various states.

3.1. Modeling Simplifications

The process evolutions at various nodes in the system are coupled, and each node (source or relay) can be present in any state; further, different nodes can interact with different subsets of nodes based on their Carrier Sense (CS) range and positions.² Hence, the random process representing the system is multi-dimensional and non-homogeneous. The exact analysis of such a process is intractable. We, therefore, make several simplifications:

- (S1) All packets are of the same fixed length. *Hence, the DATA transmission duration is fixed.*
- (S2) The internode propagation delays are of the order of nanoseconds, and hence can be taken to be zero.
- (S3) Collision model (zero packet capture): If a receiver is in the communication range of two or more nodes that are transmitting simultaneously, it does not receive any of those transmitted packets, i.e, all the packets involved in a collision get corrupted unlike an SINR model, where some of the interfering packets may still be received successfully, depending on the SINR threshold. Later, in Section 7, we have performed simulations with the Capture (SINR) Model as well to show that the error incurred due to this simplification is insignificant.
- (S4) The Inter Frame Spacing (IFS) is negligible.
- (S5) By (S3), a packet is necessarily bad if any other packet is being heard by the same receiver. Even if this is not the case, a packet can be in error due to noise. The packet error probability on each link is fixed and known. *As already mentioned, the links are not time varying and we can empirically obtain the error characteristics of a link.*
- (S6) ACK packets are short and, therefore, the probability of corruption by PHY layer is noise small enough to be neglected. This simplification is based on the fact that the error probability of ACKs is an order of magnitude less than that of data packets since length of ACK packets (22 symbols) is an order of magnitude less than that of data packets. Thus an error probability of 2% for data packets implies an error probability of 0.2% for ACK packets.
- (S7) A node's CCA succeeds when there is no transmission by any node in its Carrier Sense (CS) range at the time of initiation of its CCA. *Recall that in the standard, the channel state is averaged over the 8 symbol duration.*
- (S8) The time taken by a transmitting node for the activities of successful transmission and collision are the same. We denote this time by T_{tx} . *If the transmitted data collides at the receiver, the macAckWaitDuration for the transmitting node is equal to the sum of the turnaround time of 12 T_s and the ACK duration of 22 T_s , a total of 34 T_s (see Figure 3).*

²A node j is said to be within the Carrier Sense (CS) range of a node i if any channel activity (transmission) due to node j can be detected (but not necessarily decoded) by node i .

(S9) We assume symmetry in carrier sensing, and signal reception, i.e., if a node i can detect (respectively decode) the packet transmissions of another node j , then node j can also detect (resp. decode) the packet transmissions of node i .

Our simulation results in Section 7 suggest that these simplifications do not affect the accuracy of our analysis significantly.

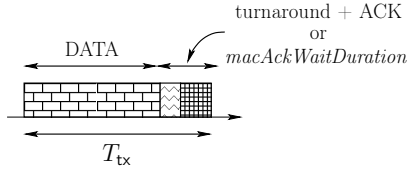


Figure 3: The transmission period T_{tx} which includes the DATA transmission time T_x , the turnaround time of $12 T_s$, and the MAC ACK duration of $22 T_s$ (or the $macAckWaitDuration$ of $34 T_s$, by (S8)).

Decoupling approximation: We employ an approximation whereby we model each node separately, incorporating the influence of the other nodes in the network by their average statistics, and as if these nodes were independent of the tagged node. This commonly used approximation is also called a “mean field approximation”, and has been widely used in the wireless networking literature [5, 6, 18].

3.2. Modeling Node Activity

By modeling simplification (S1), we assume that all packets entering node i have the same fixed length, and hence take the same amount of time when being transmitted over the medium (denoted by T_{tx}). The channel activity perceived by node i is only due to the nodes in its CS set (i.e., the set of nodes within its CS range) denoted by Ω_i . Note that Ω_i does not include node i . Each node alternates between periods during which its transmitter queue is empty and those during which the queue is non-empty. When a node’s queue is non-empty and the node is not transmitting, it is performing repeated backoffs for the head-of-the-line (HOL) packet in its queue. Each backoff ends in a CCA attempt. Figure 4 depicts the stochastic process evolution at a node *during the periods when its queue is nonempty*. We employ the decoupling approximation to analyse this process. Consider the completion of a transmission by Node i (see

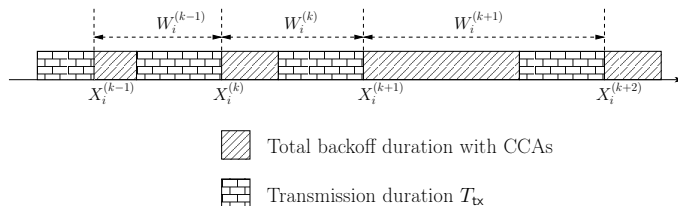


Figure 4: The alternating (contention-transmission) process obtained by observing the process at a node i after removing all the idle time periods at Node i . The transmission completion epochs are denoted by $\{X_i^{(k)}\}$ and the contention-transmission cycle lengths by $\{W_i^{(k)}\}$.

Figure 4); this could have been a success or a collision. After the completion of its transmission, Node i contends for the channel by executing successive backoffs; the nodes $j \in \Omega_i$ either have non-empty queues and are also contending, or have empty queues and are not contending. With these observations in mind, the CCA attempt process at Node i

conditioned on being in backoff periods is modeled as a Poisson process of rate β_i . For each node $j \in \Omega_i$, the CCA attempt process, conditioned on j 's queue being non-empty *and* being in backoff, or being empty, is modeled by an independent Poisson process with rate $\bar{\tau}_j^{(i)}$, $j \neq i$. This is the rate of attempts of Node j as perceived by Node i during Node i 's back-off periods. This approximation of modeling the CCA attempt process at a node conditioned on being in backoff as a Poisson process has been widely used in the literature as part of the decoupling approximation. See Appendix 9.1 for a more detailed discussion on this approximation.

Now observe from Figure 4 that as a result of the assumption of independent Poisson attempt processes at the nodes, the instants $X_i^{(k)}$ are renewal instants, and the intervals $W_i^{(k)}$ form an i.i.d.³ sequence.

Before proceeding further, let us discuss the various packet collision scenarios that can arise, and their implications in our analysis. Packet collisions in a wireless network with a CSMA MAC can occur due to the presence of hidden nodes. In addition, *there are scenarios where transmissions from two transmitters placed within the CS range of each other can overlap, known as **Simultaneous Channel Sensing and Vulnerable Window*** (also known as First and Second Collision window respectively, in [16]). These two collision scenarios are explained in Figures 5 and 6, respectively. Note that we are making use of (S2) to avoid including insignificant timing details here.

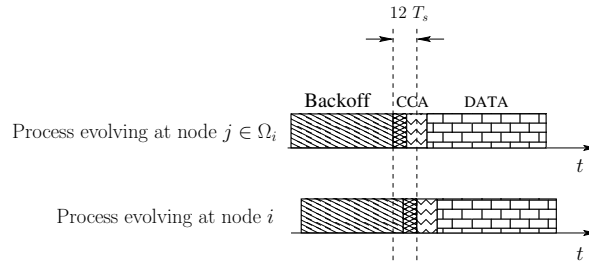


Figure 5: Node $j \in \Omega_i$ finishes its backoff, performs a CCA, finds the channel idle and starts transmitting the DATA packet. Node i finishes its backoff anywhere in the shown $12 T_s$ duration and there is no other ongoing transmission in Ω_i , its CCA succeeds and it enters the transmission duration. As a result, the DATA packets may collide at $r(i) \in \Omega_j$ and/or $r(j) \in \Omega_i$.

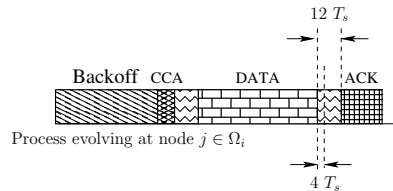


Figure 6: Vulnerable window of $4 T_s$ in the transmission period of node $j \in \Omega_i$ during which node i 's CCA attempt would be successful.

Since the vulnerable window size is small compared to the data transmission duration of a node (see Figure 6), *we shall assume that the probability of a node's CCA initiating in the vulnerable window of another node is negligible*, and therefore, neglect the resulting packet collisions. Further, since the ACK packet size is just a small fraction of DATA packet size (e.g. compare an ACK packet (22 symbols) with a DATA packet of length 260 symbols at PHY

³Independently and identically distributed

layer), we shall assume that the probability of packet collision involving ACK packets is negligible.

We now return to the renewal process observed at Node i . Figure 7 shows details of a renewal cycle of node i . The dashed transmission durations belong to nodes other than i . The cycle always ends with a transmission from node i . This last transmission in a cycle always corresponds to a non-collided packet, which, however, could be received in error.

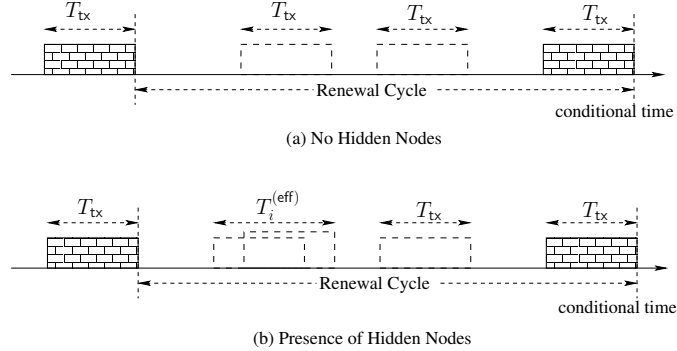


Figure 7: Renewal cycles at a tagged node, say i : (a) a renewal cycle when there are no hidden nodes in the system (all activity periods on the channel are of length T_{tx}); (b) a renewal cycle when there are hidden nodes (due to hidden node collisions, the activity periods on the channel can be longer than T_{tx} .)

Observe from Figure 7 the difference in the channel activity durations perceived by Node i in the presence and absence of hidden nodes. We now elucidate this difference in length of activity periods in a renewal cycle.

(a) Absence of Hidden Nodes: When a node transmits, all the other nodes in the network can hear it, resulting in CCA failures for other nodes that try to assess the channel in the transmission period. Also if two nodes are involved in simultaneous channel sensing, the activity period may extend from T_{tx} to a maximum of $(T_{tx} + 12 T_s)$. Since $12 T_s \ll T_{tx}$, we can assume that the activity period is only a single transmission period of duration T_{tx} .

(b) Presence of Hidden Nodes: Since a node's hearing capacity is limited, it may not perceive the activities of all the nodes in the network which may cause dilation of activity period. As shown in Figure 8, node j and k are hidden from

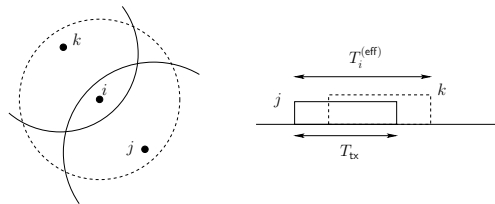


Figure 8: Dilation of transmission period as perceived by node i .

each other with respect to the receiver node i . Suppose node j starts transmitting while nodes i and k are silent. Since node k cannot hear node j , it can start its own transmission while the transmission from node j is still going on. For node i , this looks like a dilated transmission period, whose length is denoted by $T_i^{(eff)}$.

Table 2: A list of frequently used symbols. In brackets is shown the location (section number) of first use of the symbol.

Symbol	Meaning
α_i	CCA failure probability of Node i (Section 4.1.1)
β_i	CCA attempt rate of Node i when in backoff (Section 3.2)
b_i	Fraction of time Node i is in backoff when non-empty (Section 4.1)
σ_i	Mean service rate at Node i (Section 4.1.4)
$\bar{\tau}_j^{(i)}$	Node j 's attempt rate, conditioned on j 's queue being non-empty and in backoff, or being empty (Section 3.2)
Ω_i	Set of nodes in CS range of Node i , not including itself (Section 3.2)
η_i	Probability that Node i makes a CCA attempt before any node in Ω_i (Section 4.1)
c_i	Probability of simultaneous channel sensing by Node i with another node in Ω_i (Section 4.1)
l_i	Link error probability on the outgoing link from Node i (Section 4.1.3)
λ_i	Exogenous arrival rate at Node i (Section 4.1.4)
ν_i	Total arrival rate into Node i (Section 4.1.4)
T_{tx}	Packet transmission duration (Section 3.1)
$T_i^{(eff)}$	Dilated activity period perceived by Node i due to collision of nodes in Ω_i that are hidden from one another (Section 3.2)
γ_i	Probability of failure of a packet transmitted from Node i (Section 4.1.3)
n_c	Number of successive CCA failures after which a packet is discarded (Section 4.1.4)
n_t	Number of successive transmission failures after which a packet is discarded (Section 4.1.4)
r_i	Probability that the HOL packet at Node i was transmitted, and it encountered a transmission failure (Section 4.1.4)
δ_i	Packet discard probability at Node i (Section 4.1.4)
θ_i	Goodput at Node i (i.e., the rate of packet arrivals <i>from</i> Node i to its intended receiver) (Section 4.1.4)
q_i	Probability that Node i 's queue is non-empty (Section 4.1)
\bar{B}_i	Mean backoff duration at Node i during a single transmission attempt (Section 4.1.5)
\bar{h}_i	Unconditional fraction of time Node i is not transmitting (Section 4.1.3)
$r(i)$	Intended receiver of Node i (Section 3.2, Figure 5)
$C_{r(i)}^{(1)}$	Set of nodes in Ω_i that can cause interference at the intended receiver of Node i (Section 4.1.3)
$C_{r(i)}^{(2)}$	Set of nodes hidden from Node i that can cause interference at the intended receiver of Node i (Section 4.1.3)
$\alpha_j^{(-i)}$	Probability of CCA failure of Node j only due to the nodes in $\Omega_j \setminus \Omega_i$ (Section 4.1.8)
S_i	Random variable indicating service time at Node i for the HOL packet (Section 5.4)
ρ_i	Traffic intensity (i.e., the arrival rate multiplied by the mean service time required for each arrival) at Node i (Section 5.4)
$\bar{\Delta}_i$	Mean sojourn time of a packet at Node i (Section 5.4)

4. Derivation of Node Interaction Equations

In the following subsections, we consider the node processes described in Section 3, and identify certain useful steady state quantities at each node, e.g., CCA attempt rate. Then we employ the decoupling approximation (described in Section 3), and detailed stochastic analysis, to write down certain fixed point equations relating these quantities. In doing this, *we assume the network to be stable so that these steady state quantities exist*. Then in Section 5, we use an iterative scheme to solve for the unknown variables, and finally use the quantities thus obtained to evaluate the network performance, namely, end-to-end delay and packet delivery probability from each source node. The analysis will be presented for the more general case where hidden nodes may be present in the system; wherever necessary, we shall mention the changes required in the expressions for the special case where no hidden nodes are present in the system. Before proceeding further, we present in Table 2, a list of the main symbols we will be using throughout the derivation.

4.1. Derivation of Fixed Point Equations

Before delving into the details of the analysis, let us develop some intuition as to how the analysis should proceed. Eventually, we would be interested in obtaining the end-to-end delivery probability and mean delay from each source node to the sink. Let us first focus on the end-to-end delivery probability. We postpone the discussion of mean delay until Section 5.4. Since we are considering a tree topology, the end-to-end delivery probability along a path should depend on the length of the path, and the *packet discard probability* on each link in that path. Now, as we

explained earlier in Section 2, packet discard at a node⁴ occurs either due to successive CCA failures, or due to successive transmission failures. Thus, to estimate the packet discard probability, we need to estimate the CCA failure probability, and the transmission failure probability at a node. This will be the starting point of our derivation.

We begin by analyzing the process at node i shown in Figures 4 and 7. Consider an epoch at which node i starts its backoffs in order to send the first packet in its queue. At this instant, each node $j \in \Omega_i$ is either empty or performing backoffs for its HOL packet or transmitting its HOL packet. Recall that the rate at which node i completes backoffs during its backoff periods is denoted by β_i . In a network with hidden nodes, in general, $\Omega_j \setminus \Omega_i$ is nonempty. Hence, CCA attempts of node j could be blocked by the activities of nodes in $\Omega_j \setminus \Omega_i$, thus leading to the node i perceiving a CCA attempt rate from node j that is less than the CCA attempt rate when $\Omega_j \setminus \Omega_i = \phi$ (absence of hidden nodes). We have denoted this effective CCA attempt rate of node j as perceived by node i by $\bar{\tau}_j^{(i)}$ for all $j \in \Omega_i$. Modeling these conditional attempt processes as independent Poisson processes, we define:

$$\eta_i = \frac{\beta_i}{\beta_i + \sum_{j \in \Omega_i} \bar{\tau}_j^{(i)}} \quad ; \quad g_i = \frac{1}{\beta_i + \sum_{j \in \Omega_i} \bar{\tau}_j^{(i)}}$$

$$\text{and} \quad c_i = \left(1 - e^{-12 T_s \beta_i} \right)$$

Thus, η_i is the probability that node i makes a CCA attempt before any node in Ω_i , g_i is the mean time until the first CCA attempt, and c_i is the probability that after an attempt by a node in Ω_i , node i attempts within the 12 symbol vulnerable period, thus causing a “simultaneous” channel sensing collision (see Figure 5).

We introduce the following notation which will be used throughout the derivation:

$N_j^{(\text{cca})}(t)$: Total number of CCAs attempts of Node j in time $(0, t]$

$N_j^{(\text{ccaf}-i)}(t)$: Number of failed CCA attempts of Node j due to the nodes in $\Omega_j \setminus \Omega_i$ in time $(0, t]$

$I_j(t)$: The time during which node j is not transmitting in time $(0, t]$

$T_j^{(\text{bo})}(t)$: Time for which node j is in backoff in the interval $(0, t]$

$T_j^{(\text{ne})}(t)$: Queue non-empty period of node j in the interval $(0, t]$

Note that with these notation, we have

$$\beta_j = \lim_{t \rightarrow \infty} \frac{N_j^{(\text{cca})}(t)}{T_j^{(\text{bo})}(t)} \text{ a.s}$$

We further define,

$$b_j = \lim_{t \rightarrow \infty} \frac{T_j^{(\text{bo})}(t)}{T_j^{(\text{ne})}(t)} ; \quad q_j = \lim_{t \rightarrow \infty} \frac{T_j^{(\text{ne})}(t)}{t}$$

⁴In a tree topology, each node has exactly one receiver, and hence we can use the terms node and link interchangeably when referring to discard probabilities.

where b_j can be interpreted as the long-term fraction of time node j is in backoff provided it is non-empty, and q_j is the long-term fraction of time that node j is non-empty, which is also the queue non-empty probability of node j , assuming the system is ergodic.

With these notation and definitions, we are now ready to obtain expressions for the CCA failure and transmission failure probabilities at a tagged node i , which we proceed to do next.

4.1.1. Probability of CCA failure, α_i

The probability of CCA failure is the probability of the occurrence of at least one transmitting node in the CS range of node i , given that node i performs a CCA. Defining $N_i^{(\text{cca})}(t)$ and $N_i^{(\text{f})}(t)$ as the total number of CCAs and number of failed CCAs in the interval $(0, t]$, respectively, we have:

$$\alpha_i = \lim_{t \rightarrow \infty} \frac{N_i^{(\text{f})}(t)}{N_i^{(\text{cca})}(t)} \text{ a.s.} = \lim_{t \rightarrow \infty} \frac{\frac{N_i^{(\text{f})}(t)}{T_i^{(\text{ne})}(t)}}{\frac{N_i^{(\text{cca})}(t)}{T_i^{(\text{ne})}(t)}}$$

Applying Renewal-Reward Theorem (RRT) [24], we get

$$\alpha_i = \frac{\frac{N_i^{(\text{f})}}{W_i}}{\frac{N_i^{(\text{cca})}}{W_i}} = \frac{N_i^{(\text{f})}}{N_i^{(\text{cca})}} \quad (1)$$

where $N_i^{(\text{f})}$ is the mean number of failed CCAs and $N_i^{(\text{cca})}$ is the mean number of total CCAs in a renewal cycle. W_i is the mean renewal cycle length.

To obtain $N_i^{(\text{cca})}$, we can write

$$N_i^{(\text{cca})} = \eta_i + (1 - \eta_i)c_i + (1 - \eta_i)(1 - c_i)(\beta_i T_i^{(\text{eff})} + N_i^{(\text{cca})})$$

Here, the first two terms correspond to the events where node i 's CCA succeeds, or node i is involved in a simultaneous channel sensing such that the cycle ends after one CCA. When neither of these two events happen, node i performs CCAs at the rate of β_i for the entire dilated transmission period $T_i^{(\text{eff})}$ i.e., a total of $\beta_i T_i^{(\text{eff})}$ CCAs all of which fail, and the cycle continues after this period. Upon rearranging terms, we have

$$N_i^{(\text{cca})} = \frac{\eta_i + (1 - \eta_i)c_i + (1 - \eta_i)(1 - c_i)\beta_i T_i^{(\text{eff})}}{1 - (1 - \eta_i)(1 - c_i)} \quad (2)$$

Using similar arguments, the mean number of failed CCA attempts by node i in a renewal cycle, i.e., $N_i^{(\text{f})}$, can be written using (S7) as

$$N_i^{(\text{f})} = (1 - \eta_i)(1 - c_i)(\beta_i T_i^{(\text{eff})} + N_i^{(\text{f})})$$

Rearranging and using Equation (1), the CCA failure probability is given by

$$\alpha_i = \frac{(1 - \eta_i)(1 - c_i)\beta_i T_i^{(\text{eff})}}{\eta_i + (1 - \eta_i)c_i + (1 - \eta_i)(1 - c_i)\beta_i T_i^{(\text{eff})}} \quad (3)$$

Remark: Note that for no hidden node network, $T_i^{(\text{eff})} = T_{\text{tx}}$.

We shall now show how the dilated channel activity period perceived by node i , $T_i^{(\text{eff})}$, can be expressed in terms of the CCA attempt rates β_i and $\bar{\tau}_j^{(i)}$, $j \in \Omega_i$. This, combined with (3), enables us to express the CCA failure probability, α_i , only in terms of the CCA attempt rates, β_i and $\bar{\tau}_j^{(i)}$, $j \in \Omega_i$.

4.1.2. The dilated activity period, $T_i^{(\text{eff})}$

As mentioned earlier, the length of the dilated activity period as perceived by a node depends upon the set of nodes in its CS range. We propose two different models for calculating $T_i^{(\text{eff})}$.

1. **$M/D/\infty$ model** : Since each node has a different set of neighbours, we can make the following approximation to simplify our analysis:

(A1) The mean length of the dilated activity period as perceived by node i is equal to the length of the mean busy period of an $M/D/\infty$ queue where the deterministic service time is equal to a single transmission period T_{tx} , and the arrival process is approximated as a Poisson process having rate equal to the aggregate transmission initiation rate of *all* the nodes in the CS range of node i .

Note that the above assumption is equivalent to saying that all nodes in Ω_i are hidden from each other, and hence this approximation results in a larger mean for the dilated activity period.

Note that the transmission initiation rate of any node $j \in \Omega_i$, as perceived by node i , is $\bar{\tau}_j^{(i)}$. For any node i , let ζ_i be the aggregate transmission initiation rate for nodes in Ω_i . Then

$$\zeta_i = \sum_{j \in \Omega_i} \bar{\tau}_j^{(i)} \quad (4)$$

With this model, the expression for $T_i^{(\text{eff})}$ can be written recursively as

$$\begin{aligned} T_i^{(\text{eff})} &= T_{\text{tx}} \exp\{-\zeta_i T_{\text{tx}}\} + \int_0^{T_{\text{tx}}} (u + T_i^{(\text{eff})}) \zeta_i \exp\{-\zeta_i u\} du \\ &= \frac{1}{\zeta_i} (\exp\{\zeta_i T_{\text{tx}}\} - 1) \end{aligned} \quad (5)$$

where, in the first equality, the first term corresponds to the event that some node $j \in \Omega_i$ started transmission, and no other node in Ω_i initiated transmission during T_{tx} so that the busy period ends after T_{tx} ; the second term corresponds to the event that some node $j \in \Omega_i$ started transmission, and another node $k \in \Omega_i$ starts at time $u < T_{\text{tx}}$, thus extending the busy period by $T_i^{(\text{eff})}$.

2. **Boorstyn et al. [25] model**: For node i in Figure 8, Figure 9 shows its evolution process in conditional time. The grey shaded regions indicate times during which all the nodes in Ω_i as well as node i are contending for the medium. We see that when nodes in Ω_i start packet transmission, then we have a dilated transmission period, $T_i^{(\text{eff})}$ as perceived by Node i .

Here we will use a CTMC model as suggested by Boorstyn et al. [25] for determining $T_i^{(\text{eff})}$ during times when node i is non empty. Let ϕ_i be the state when all the nodes in Ω_i as well as node i are in backoff (thus, in Figure 9, ϕ_i is the state of the system in the grey shaded regions). Then the state transition diagram for computing $T_i^{(\text{eff})}$ for the nodes in Figure 8 is shown in Figure 10. By the *insensitivity property* explained in [25], the steady state probability of being in state ϕ_i , say π_{ϕ_i} , does not depend on the distribution of packet transmission time except by its mean. So we can take the packet service times to be exponentially distributed. Hence π_{ϕ_i} can be obtained as in [25], i.e.,

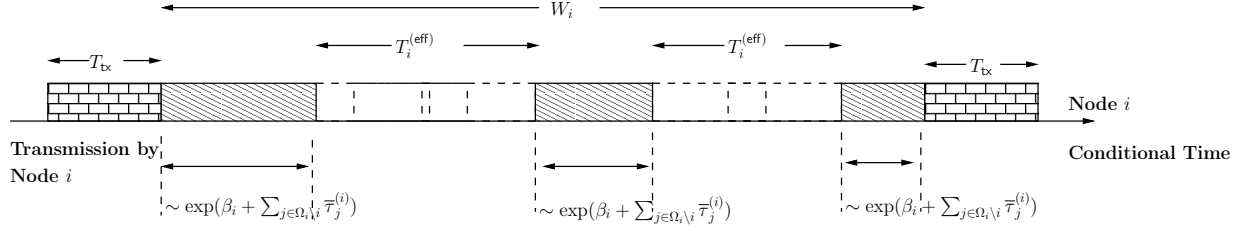


Figure 9: Evolution Process of Node i in conditional time. Grey shaded regions indicate that all nodes in Ω_i as well as node i are in back-off. We see that when nodes in Ω_i attempt packet transmission, we have a dilated transmission period, $T_i^{(\text{eff})}$ as perceived by Node i . The instants at which channel activity ends, i.e., the starting epochs of the grey shaded regions are renewal instants.

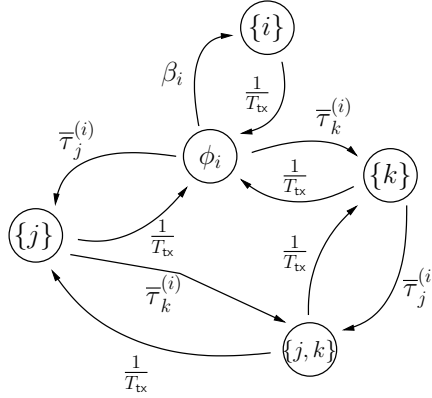


Figure 10: State transition diagram for computing $T_i^{(\text{eff})}$ for the nodes in Figure 8

$$\pi_{\phi_i} = \left[\sum_{D_i \in \mathcal{D}_i} \prod_{j \in D_i} \bar{\tau}_j^{(i)} T_{\text{tx}} \right]^{-1} \quad (6)$$

where D_i is a set of nodes in Ω_i that are actively transmitting, and \mathcal{D}_i is the collection of all such sets.

Again, note that return to ϕ_i are renewal instants since the transition out of ϕ_i are all exponentially distributed (see Figure 9). Hence, using the RRT we obtain π_{ϕ_i} as the ratio of the mean time the system is in state ϕ_i and the mean time between the ends of channel activity (transmission) during a busy period of Node i (see Figure 9), i.e.,

$$\pi_{\phi_i} = \frac{\frac{1}{\beta_i + \sum_{j \in \Omega_i} \bar{\tau}_j^{(i)}}}{\frac{1}{\beta_i + \sum_{j \in \Omega_i} \bar{\tau}_j^{(i)}} + \frac{\beta_i}{\beta_i + \sum_{j \in \Omega_i} \bar{\tau}_j^{(i)}} \times T_{\text{tx}} + \frac{\sum_{j \in \Omega_i} \bar{\tau}_j^{(i)}}{\beta_i + \sum_{j \in \Omega_i} \bar{\tau}_j^{(i)}} \times T_i^{(\text{eff})}}$$

which yields

$$\pi_{\phi_i} = \frac{1}{1 + \beta_i \times T_{\text{tx}} + \left(\sum_{j \in \Omega_i} \bar{\tau}_j^{(i)} \right) \times T_i^{(\text{eff})}} \quad (7)$$

Hence, from Equation (6) and (7), we can obtain $T_i^{(\text{eff})}$ for each node i in terms of β_i and $\bar{\tau}_j^{(i)}, j \in \Omega_i$. This,

together with 3, yields the CCA failure probability, α_i , in terms of β_i and $\bar{\tau}_j^{(i)}$, $j \in \Omega_i$. Observe that the dilated activity period computed using the Boorstyn et al. model would be less conservative than the $M/D/\infty$ model. Also note that identification of \mathcal{D}_i is an NP-complete combinatorial problem [25]; however, efficient algorithms exist that handle networks of arbitrary topology and moderate size (50-100 nodes).

Next, we proceed to derive an expression for the packet transmission failure probability at a node i .

4.1.3. Packet failure probability, γ_i

A transmitted packet can fail to be decoded by its intended receiver due to a collision, or due to noise. We do not take packet capture into account here (see **(S3)**). Define $M_i(t)$ as the total number of transmissions from node i in the interval $(0, t]$, $M_i^{(c)}(t)$ as the total number of collisions encountered by packets from node i in the interval $(0, t]$ and l_i as the probability of *data* packet error (known by **(S5)**) on the link between node i and its receiver $r(i)$ due to noise. Recall that ACKs are not corrupted by PHY noise (see **(S6)**). The probability of packet failure, γ_i , is then defined as:

$$\gamma_i = \left(\lim_{t \rightarrow \infty} \frac{M_i^{(c)}(t)}{M_i(t)} \right) + \left(1 - \lim_{t \rightarrow \infty} \frac{M_i^{(c)}(t)}{M_i(t)} \right) l_i \text{ a.s.}$$

Dividing the numerator and denominator of each term inside the limits by $T_i^{(ne)}(t)$ and applying RRT, we get

$$\gamma_i = \left(\frac{\mathbf{M}_i^{(c)}}{\mathbf{W}_i} \right) + \left(1 - \frac{\mathbf{M}_i^{(c)}}{\mathbf{W}_i} \right) l_i$$

where $\mathbf{M}_i^{(c)}$ is the mean number of collided packets, and \mathbf{M}_i is the mean number of transmitted packets from node i in a renewal cycle. Recall that W_i is the mean length of a renewal cycle at node i (see Figure 4 and the discussion thereafter). Note that there is a single transmission by node i in each renewal cycle so that \mathbf{M}_i becomes unity. We can rewrite

$$\gamma_i = p_i + (1 - p_i)l_i \tag{8}$$

where p_i is the probability of packet collision and equals $\mathbf{M}_i^{(c)}$. Since l_i is given for each node i , we need to compute probability of packet collision, p_i , to compute packet failure probability, γ_i . We proceed to do this next.

Computation of packet collision probability, p_i

We denote the set of nodes that can cause interference in successful reception at $r(i)$ by $I_{r(i)}$, which is composed of two sets:

$$\begin{aligned} C_{r(i)}^{(1)} &= \{j \in \mathcal{N} : j \in \Omega_i \text{ and } j \in I_{r(i)}\} \\ C_{r(i)}^{(2)} &= \{j \in \mathcal{N} : j \notin \Omega_i \text{ and } j \in I_{r(i)}\} \end{aligned}$$

such that

$$C_{r(i)}^{(1)} \cap C_{r(i)}^{(2)} = \emptyset; C_{r(i)}^{(1)} \cup C_{r(i)}^{(2)} = I_{r(i)}$$

Note that the receiver $r(i)$ is in $C_{r(i)}^{(1)}$ (see Figure 11 for a depiction of these sets).

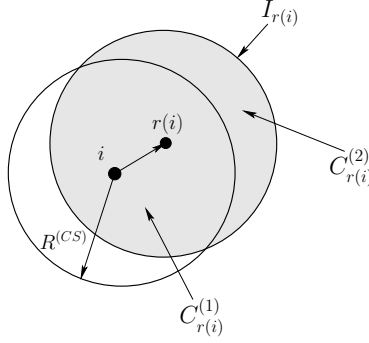


Figure 11: Interference region around the receiver. $R^{(CS)}$ is the CS range which is assumed to be equal for all nodes by (S9).

Recall that the collision probability, p_i , is nothing but $\mathbf{M}_i^{(c)}$, the mean number of collisions of node i 's packets in a renewal cycle, which can be written recursively as

$$\mathbf{M}_i^{(c)} = (1 - \eta_i)(1 - c_i)\mathbf{M}_i^{(c)} + \overline{\mathbf{M}}_i^{(c)}$$

where $\overline{\mathbf{M}}_i^{(c)}$ is the mean number of collisions of node i 's packets in a renewal cycle when node i attempts a packet transmission or does a simultaneous channel sensing. On rearranging, the packet collision probability at node i is given by

$$p_i = \frac{\overline{\mathbf{M}}_i^{(c)}}{1 - (1 - \eta_i)(1 - c_i)} \quad (9)$$

As we shall see in a moment, to compute $\overline{\mathbf{M}}_i^{(c)}$, we first need to find the probabilities that nodes in certain given sets are not transmitting. We use a product approximation to do this due to the unavailability of joint distribution of processes.

Let \bar{h}_i be the unconditional fraction of time node i is not transmitting and is given by

$$\bar{h}_i = \lim_{t \rightarrow \infty} \frac{I_i(t)}{t} \quad \text{a.s} = (1 - q_i) + q_i b_i$$

where we recall that the first term is the fraction of time that node i 's queue is empty, and the second term is the fraction of time that node i 's queue is non-empty, and the node is in backoff.

With this, the probability that the nodes in a given set \mathcal{S} are not transmitting is approximated by $\prod_{j \in \mathcal{S}} \bar{h}_j$.

The expression for $\overline{\mathbf{M}}_i^{(c)}$ can now be found as the sum of several components; each of these components represents the probability of a particular collision event as explained below, and these events are mutually exclusive. The rationale behind this computation is straightforward; each of these mutually exclusive collision events will result in a *single* packet collision for node i in a renewal cycle. Thus, the mean number of collisions in a renewal cycle when node i attempts a packet transmission, or does a simultaneous channel sensing is simply the sum of the probabilities of these mutually exclusive collision events.

1. The first term accounts for the fact that Node i started transmitting *in the presence of at least one transmission by the hidden nodes set*, i.e., $C_{r(i)}^{(2)}$.

$$R_i^{(1)} = \eta_i \left(1 - \prod_{j \in C_{r(i)}^{(2)}} \bar{h}_j \right)$$

2. The second term accounts for the scenario when Node i started its transmission as a *simultaneous channel sensing with a node in Ω_i in the presence of at least one transmission by the hidden nodes set*, i.e., $C_{r(i)}^{(2)}$.

$$R_i^{(2)} = (1 - \eta_i) c_i \left(1 - \prod_{j \in C_{r(i)}^{(2)}} \bar{h}_j \right)$$

3. The third term considers the case where Node i starts transmitting in the absence of any ongoing transmission by a hidden node but it encounters a *simultaneous transmission by a node in $C_{r(i)}^{(1)}$ anywhere in the corresponding vulnerable window of $12 T_s$ and/or a transmission by a node in $C_{r(i)}^{(2)}$, i.e., a hidden node anywhere in its activity period T_{tx} .*

$$R_i^{(3)} = \eta_i \left(\prod_{j \in C_{r(i)}^{(2)}} \bar{h}_j \right) \left(1 - \exp \left\{ -12 T_s \left(\sum_{j \in C_{r(i)}^{(1)}} \bar{\tau}_j^{(i)} \right) \right\} \cdot \exp \left\{ -T_{\text{tx}} \left(\sum_{j \in C_{r(i)}^{(2)}} \bar{\tau}_j \right) \right\} \right)$$

where $\bar{\tau}_j$ is the rate of successful CCA attempts of node j over its non-transmitting period, $I_j(t)$ and is equal to (using RRT)

$$\bar{\tau}_j = \lim_{t \rightarrow \infty} \frac{N_j^{(\text{cca})}(t) - N_j^{(\text{f})}(t)}{I_j(t)} \quad \text{a.s} = \frac{\beta_j b_j q_j (1 - \alpha_j)}{1 - q_j + q_j b_j}$$

Observe the intuition behind the expression for $\bar{\tau}_j$; the numerator can be interpreted as the unconditional rate of successful CCA attempts, while the denominator is the fraction of time that the node is not transmitting.

4. The fourth term says that Node i started its transmission as a *simultaneous channel sensing with a node in $C_{r(i)}^{(1)}$ in the absence of any hidden node*. This event surely ends up in a collision at the receiver $r(i)$.

$$R_i^{(4)} = \left(\frac{\sum_{j \in C_{r(i)}^{(1)}} \bar{\tau}_j^{(i)}}{\beta_i + \sum_{j \in \Omega_i} \bar{\tau}_j^{(i)}} \right) c_i \left(\prod_{j \in C_{r(i)}^{(2)}} \bar{h}_j \right)$$

5. The final term says that Node i started its transmission as a *simultaneous channel sensing with a node in $\Omega_i \setminus C_{r(i)}^{(1)}$ in the absence of any hidden node but it encounters a simultaneous transmission by a node in $C_{r(i)}^{(1)}$ and/or by a node in $C_{r(i)}^{(2)}$, i.e., a hidden node.*

$$R_i^{(5)} = \left(\frac{\sum_{j \in \Omega_i \setminus C_{r(i)}^{(1)}} \bar{\tau}_j^{(i)}}{\beta_i + \sum_{j \in \Omega_i} \bar{\tau}_j^{(i)}} \right) c_i \left(\prod_{j \in C_{r(i)}^{(2)}} \bar{h}_j \right) \left(1 - \exp \left\{ -12 T_s \left(\sum_{j \in C_{r(i)}^{(1)}} \bar{\tau}_j^{(i)} \right) \right\} \exp \left\{ -T_{\text{tx}} \left(\sum_{j \in C_{r(i)}^{(2)}} \bar{\tau}_j \right) \right\} \right)$$

Observe that in writing down the expressions for $R_i^{(3)}$, $R_i^{(4)}$ and $R_i^{(5)}$, we have used the assumption of independent Poisson attempt processes at the nodes.

The overall packet collision probability is given by

$$p_i = \frac{R_i^{(1)} + R_i^{(2)} + R_i^{(3)} + R_i^{(4)} + R_i^{(5)}}{\eta_i + (1 - \eta_i) c_i} \quad (10)$$

The packet failure probability, γ_i , can now be calculated using Equations (8) and (10).

Remark: Note that for no hidden node network $C_{r(i)}^{(2)} = \phi$ and $\prod_{j \in C_{r(i)}^{(2)}} \bar{h}_j = 1$.

Once we have the estimates of CCA failure probability, α_i , and transmission failure probability, γ_i at each node, we can use these to compute several useful measures at each node, most notably among them, the discard probability at each node. This is the next step in our derivation.

4.1.4. Discard probability (δ_i), aggregate arrival rate (ν_i), and goodput (θ_i) for a node i

The probability of discarding a packet due to consecutive CCA failures, or successive failed retransmission attempts, δ_i , is given by

$$\delta_i = \alpha_i^{n_c} (1 + r_i + r_i^2 + \dots + r_i^{n_t - 1}) + r_i^{n_t} \quad (11)$$

where $r_i = \gamma_i(1 - \alpha_i^{n_c})$ is the probability that the HOL packet at node i was transmitted, and it encountered a transmission failure. n_c and n_t denote respectively, the values of the parameters *macMaxCSMABackoffs* and $1 + aMaxFrameRetries$ (see Section 2).

The arrival process at each node consists of packets which are generated at the same node (if it is a sensor) and the packets to be forwarded. The aggregate arrival rate at a node i can be written as:

$$\nu_i = \lambda_i + \sum_{k \in \mathcal{P}_i} \theta_k \quad (12)$$

where λ_i is the packet generation rate at sensor node i , \mathcal{P}_i is the set of all its in-neighbours, and θ_i is the rate of packet transmission by the node which are *successfully received* at $r(i)$ (known as goodput). An enqueued packet at a node is successfully received by the intended receiver in a manner shown in Figure 12.

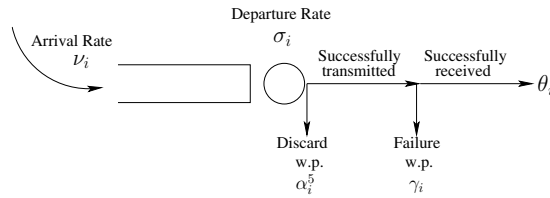


Figure 12: The goodput of a node is defined as the rate of successfully received packets by its receiver. σ_i is the mean service rate at node i .

Assuming that the queueing system is stable and has a steady-state solution, the goodput θ_i is given by

$$\theta_i = \nu_i(1 - \delta_i) \quad (13)$$

Note that if the queue at node i is saturated, then the goodput θ_i is equal to $\sigma_i(1 - \delta_i)$, where σ_i is the mean service rate of HOL packets at node i , which we compute next.

4.1.5. Average service rate (σ_i)

Each packet that reaches the HOL position in the queue at a node can have multiple backoffs and transmissions before it is successfully transmitted or discarded. We define \bar{Z}_i as the average time spent in backoff, and \bar{Y}_i as the

average transmission time until the packet is successfully transmitted or discarded at the MAC layer. Then the average service rate, σ_i , at node i can be calculated as:

$$\frac{1}{\sigma_i} = \bar{Z}_i + \bar{Y}_i \quad (14)$$

Using the values of *macminBE* and *macmaxBE*, the mean backoff durations after each successive CCA failure can be computed; for example, the mean backoff durations calculated using the default values from the standard are shown in Table 3. Let b_k , $k = 0, \dots, n_c - 1$ denote the mean backoff duration (in symbol times) after the k^{th} successive CCA failure. We can then define

$$\bar{B}_i = b_0 + \alpha_i b_1 + \dots + \alpha_i^{n_c-1} b_{n_c-1}$$

where \bar{B}_i refers to the mean backoff duration until the packet is transmitted or discarded due to successive CCA failures, i.e., the mean time spent in backoff during a *single* transmission attempt. Then the quantities \bar{Z}_i and \bar{Y}_i can be calculated as:

$$\begin{aligned} \bar{Z}_i &= \bar{B}_i(1 + r_i + r_i^2 + \dots + r_i^{n_t-1}) \\ \bar{Y}_i &= (1 - \alpha_i^{n_c})T_{\text{tx}}(1 + r_i + r_i^2 + \dots + r_i^{n_t-1}) \end{aligned}$$

where, we recall that $r_i = \gamma_i(1 - \alpha_i^{n_c})$ can be interpreted as the probability that the HOL packet at node i was transmitted (i.e., it did not encounter n_c successive CCA failures), and it encountered a transmission failure (either due to collision or due to link error). Thus, the second factor in the expression for \bar{Z}_i can be interpreted as the mean number of transmission attempts until the HOL packet is successfully transmitted or discarded. Thus, while \bar{B}_i is the mean time spent in backoff during a *single* transmission attempt, \bar{Z}_i is the mean time spent in backoff over *all* transmission attempts until the packet is finally *successfully* transmitted or discarded. Similarly, observe that $(1 - \alpha_i^{n_c})T_{\text{tx}}$ is the mean transmission time in a *single* transmission attempt, while \bar{Y}_i is the average transmission time over *all* transmission attempts until the packet is *successfully* transmitted or discarded.

Now we have obtained several quantities of interest at each node in terms of the CCA failure probabilities, $\{\alpha_i\}_{i \in \mathcal{N}}$, and the transmission failure probabilities, $\{\gamma_i\}_{i \in \mathcal{N}}$, which themselves were obtained in terms of the CCA attempt rates, $\{\beta_i\}$, $\{\bar{\tau}_j^{(i)}\}$, and the fractions $\{q_i\}$, and $\{b_i\}$. However, these CCA attempt rates, and the fractions are not known in advance. Thus, to complete the fixed point loop, we must express these in terms of $\{\alpha_i\}$ and $\{\gamma_i\}$. We start with the queue non-empty probability, q_i .

Backoff Exponent	Mean Backoff Duration(T_s)
3	70 + 8
4	150 + 8
5	310 + 8

Table 3: Mean backoff durations in symbol times, T_s

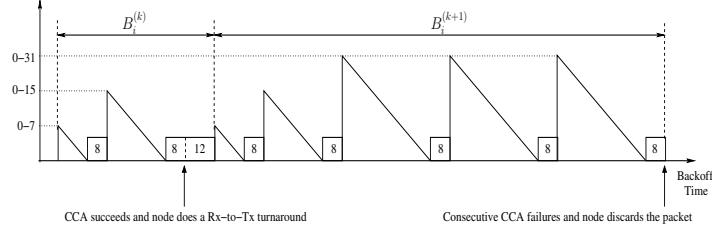


Figure 13: Evolution of backoff periods in conditional (backoff) time. The y -axis is in units of a backoff slot ($20 T_s$). The range indicates that the random backoff duration is uniformly distributed within it. The CCA duration, and Rx-to-Tx turnaround duration are 8 and 12 symbol times respectively, as indicated in the boxes.

4.1.6. The node non-empty probability, q_i

To find the expression for q_i , assuming that all the arriving packets reach the HOL position (i.e., no tail drops) and applying Little's Theorem, we get

$$q_i = \frac{\nu_i}{\sigma_i} \quad (15)$$

Further, for a saturated node, the quantity q_i is equal to 1.

4.1.7. Obtaining b_i and β_i

In order to find b_i , the fraction of time a node is in backoff provided it is non-empty, we embed a renewal process in conditional time where the renewal epochs are those instants at which the node enters the random backoff period after a packet transmission or packet discard. We use the RRT to find the expression for b_i as

$$b_i = \frac{\overline{B}_i}{\overline{B}_i + (1 - \alpha_i^{n_c})T_{tx}} \quad (16)$$

The expression for b_i is also intuitively appealing, since the numerator is the mean time spent in backoff, and the denominator is the total mean time (including transmission) until the packet is transmitted or discarded due to n_c successive CCA failures.

The quantity β_i , the rate of CCA attempts in backoff times, is based on the backoff completion times irrespective of the transmission attempt of the packet since a packet retransmission is considered the same as a new packet transmission. Thus, for the calculation of β_i , it suffices to observe the process at node i only in the backoff times. Figure 13 shows the residual backoff process where after completion of the backoff duration, the node performs a CCA. Using the result of Kumar et al. [6] with their collision probability replaced by our CCA failure probability, α_i , the expression for β_i can be written as

$$\beta_i = \frac{1 + \alpha_i + \alpha_i^2 + \alpha_i^3 + \dots + \alpha_i^{n_c - 1}}{\overline{B}_i} \quad (17)$$

Observe that the numerator is the mean number of CCA attempts, and the denominator is the mean backoff duration until the packet is transmitted or discarded due to successive CCA failures.

Finally, we are left to compute $\overline{\tau}_j^{(i)}$, the CCA attempt rate of node j as perceived by node i .

4.1.8. Perceived CCA attempt rate, $\bar{\tau}_j^{(i)}$

We can write

$$\begin{aligned}\bar{\tau}_j^{(i)} &= \lim_{t \rightarrow \infty} \frac{N_j^{(\text{cca})}(t) - N_j^{(\text{ccaf}_{-i})}(t)}{I_j(t)} \text{ a.s} \\ &= \lim_{t \rightarrow \infty} \frac{\frac{N_j^{(\text{cca})}(t)}{T_j^{(\text{bo})}(t)} \times \frac{T_j^{(\text{bo})}(t)}{T_j^{(\text{ne})}(t)} \times \frac{T_j^{(\text{ne})}(t)}{t} \times \left(1 - \frac{N_j^{(\text{ccaf}_{-i})}(t)}{N_j^{(\text{cca})}(t)}\right)}{\frac{I_j(t)}{t}}\end{aligned}$$

Recalling the definitions of β_j , b_j and q_j given at the start of Section 4.1, we have

$$\bar{\tau}_j^{(i)} = \frac{\beta_j \times b_j \times q_j \times (1 - \alpha_j^{(-i)})}{1 - q_j + q_j \times b_j} \quad (18)$$

where $\alpha_j^{(-i)}$ is the probability of CCA failure of node j only due to the nodes in $\Omega_j \setminus \Omega_i$, and is derived below.

4.1.9. Computation of $\alpha_j^{(-i)}$

Using the notation introduced earlier, we can write

$$\begin{aligned}\alpha_j^{(-i)} &= \lim_{t \rightarrow \infty} \frac{N_j^{(\text{ccaf}_{-i})}(t)}{N_j^{(\text{cca})}(t)} \text{ a.s} \\ &= \lim_{t \rightarrow \infty} \frac{\frac{N_j^{(\text{ccaf}_{-i})}(t)}{T_j^{(\text{ne})}(t)}}{\frac{N_j^{(\text{cca})}(t)}{T_j^{(\text{ne})}(t)}}\end{aligned}$$

Applying the Renewal-Reward Theorem (RRT), we have

$$\alpha_j^{(-i)} = \frac{\frac{\mathbf{N}_j^{(\text{ccaf}_{-i})}}{W_j}}{\frac{\mathbf{N}_j^{(\text{cca})}}{W_j}} = \frac{\mathbf{N}_j^{(\text{ccaf}_{-i})}}{\mathbf{N}_j^{(\text{cca})}} \quad (19)$$

where we recall that W_j is the mean time between the ends of transmission during a busy period of node j (i.e., the mean renewal cycle length). The mean number of failed CCA attempts by node j in a renewal cycle due to transmissions by nodes in $\Omega_j \setminus \Omega_i$, say $\mathbf{N}_j^{(\text{ccaf}_{-i})}$, can be expressed recursively, using **(S7)** as

$$\mathbf{N}_j^{(\text{ccaf}_{-i})} = \left(\frac{\sum_{k \in \Omega_j \cap \Omega_i} \bar{\tau}_k^{(j)}}{\beta_j + \sum_{l \in \Omega_j} \bar{\tau}_l^{(j)}} \right) (1 - c_j) \left(\mathbf{N}_j^{(\text{ccaf}_{-i})} \right) + \left(\frac{\sum_{k \in \Omega_j \setminus \Omega_i} \bar{\tau}_k^{(j)}}{\beta_j + \sum_{l \in \Omega_j} \bar{\tau}_l^{(j)}} \right) (1 - c_j) \left(\beta_j T_{\text{tx}} + \mathbf{N}_j^{(\text{ccaf}_{-i})} \right) \quad (20)$$

Clearly, in case of a simultaneous channel sensing event involving node j , the cycle ends with zero reward, since no CCA failure happens for node j . Therefore, the terms in Equation (20) only consider the case where node j is not involved in simultaneous channel sensing. The first term in (20) corresponds to transmission attempts by nodes in $\Omega_j \cap \Omega_i$; this results in zero reward (since no CCA failure due to nodes in $\Omega_j \setminus \Omega_i$ happens), and the cycle continues thereafter. The second term accounts for transmission attempts by nodes in $\Omega_j \setminus \Omega_i$; this results in a total reward of

$\beta_j T_{\text{tx}}$ (since those many CCA attempts of node j fail on an average due to transmissions from nodes in $\Omega_j \setminus \Omega_i$) and the cycle continues thereafter. Upon rearranging terms, we have

$$N_j^{(\text{ccaf}_{-i})} = \frac{\left(\frac{\sum_{k \in \Omega_j \setminus \Omega_i} \bar{\tau}_k^{(j)}}{\beta_j + \sum_{l \in \Omega_j} \bar{\tau}_l^{(j)}} \right) (1 - c_j) (\beta_j T_{\text{tx}})}{1 - (1 - \eta_j)(1 - c_j)}$$

Hence from 19, the CCA failure probability of node j only due to the nodes in $\Omega_j \setminus \Omega_i$ is given by

$$\alpha_j^{(-i)} = \frac{\left(\frac{\sum_{k \in \Omega_j \setminus \Omega_i} \bar{\tau}_k^{(j)}}{\beta_j + \sum_{l \in \Omega_j} \bar{\tau}_l^{(j)}} \right) (1 - c_j) (\beta_j T_{\text{tx}})}{\eta_j + (1 - \eta_j)c_j + (1 - \eta_j)(1 - c_j)\beta_j T_j^{(\text{eff})}} \quad (21)$$

Remark: For networks with no hidden nodes, $\alpha_j^{(-i)} = 0$.

5. Iterative Solution and Calculation of Performance Measures

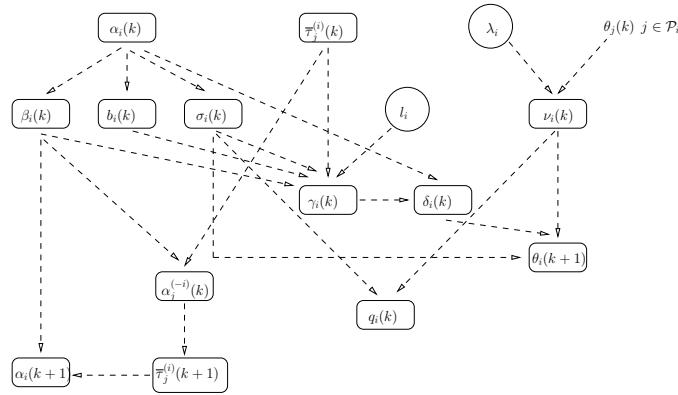


Figure 14: The global iteration scheme. The square boxes indicate the derived quantities and the round boxes indicate the known quantities.

5.1. Global Iteration Scheme

The global iteration scheme to solve the fixed point equations derived above is shown in Figure 14. We start with a vector $\{0, 10, \lambda_i, l_i\}$ corresponding to $\{\alpha_i, \bar{\tau}_j^{(i)}, \nu_i, l_i\}$ for each node i and node $j \in \Omega_i$, and repeat the procedure until the quantities converge.

5.2. Existence of a fixed point

A careful look at the derivation of the fixed point equations in Section 4.1 reveals that the fixed point variables can be taken to be $\{(\alpha_i, \gamma_i)\}_{i=1}^N$, where α_i is the CCA failure probability at node i , and γ_i is the transmission failure probability at node i . These variables are related via the set of equations (3), (6), (7), (8), (10), (11)-(18), and (21). We observe that all the functions involved in the fixed point equations are continuous (since compositions, sums, and

products of continuous functions are continuous, and also minimum of continuous functions is continuous). Hence, the fixed point equations can be represented as a *continuous* map from $[0, 1]^{2N}$ to $[0, 1]^{2N}$. Hence, the existence of a solution to the fixed point equations follows from Brouwer's fixed point theorem.

We do not have proofs of the uniqueness of the fixed point, and convergence of the iterative procedure to the fixed point. For a detailed discussion of the complexity associated with these proofs for general multi-hop topologies, see Section IV.D in [18]. To the best of our knowledge, the only work to have attempted a formal proof of the uniqueness of the fixed point in the context of multi-hop networks is [22], but that too for a rather simplified version of CSMA policies. However, *in our numerical experiments, the fixed point iterations always converged to a solution.*

5.3. Packet Delivery Probability for each Source Node

The packet delivery probability, $p_i^{(del)}$, for each source node i is defined as the fraction of generated packets at source node i that reach the base station without any time bound. Let the set of nodes constituting the path from a source node i to the BS be L_i . Assuming that the drop events are independent from node to node, the expression for $p_i^{(del)}$ is

$$p_i^{(del)} = \prod_{j \in L_i} (1 - \delta_j) \quad (22)$$

where δ_j is the packet discard probability of node j .

5.4. End-to-End Delay Calculation

We use the approximation techniques in Whitts Queueing Network Analyzer (QNA, [26]) to compute the packet delay between a source node and the sink node. The arrival processes at the source nodes are modeled as independent Poisson processes. We recall that our network is a tree, rooted at the sink. Thus, each interior node receives packets from one or more upstream nodes, and feeds exactly one downstream node. Each node is modeled by a GI/GI/1 queue. The fixed point equations obtained as explained earlier are iterated until a convergence tolerance is met. The resulting values of the attempt rates, collision rates, CCA failure rates, etc., are used as explained later in this section to obtain the first two moments of the time spent by an HOL packet at a queue, i.e., the service time at the queue. Given the first two moments of the interarrival time and of the service time, QNA *uses a standard GI/GI/1 mean delay approximation for the mean delay at a queue.* QNA also provides a method (described later in this section) for approximately modeling the superposition of the output processes of two or more GI/GI/1 queues by a renewal process, and thus provides approximations for the first two moments of the interarrival times into downstream queues. In this manner, starting with the model of the input processes at the source nodes, the end-to-end mean delay can be approximated.

To obtain the first two moments of the service time at a node, we use an approximate Moment Generating Function (MGF) of service time. We can write an expression for service time in a recursive manner by allowing infinite CCAs and retransmissions. To account for the discarded packets, we allow only a fraction of transmitted packets to join the next-hop neighbour's queue based on the discard probability at that node. We denote the service time at node i by S_i ,

and let B_i denote the length of random backoff duration, which is assumed to have an exponential distribution with rate β_i for node i . Then, S_i can be written as

$$S_i = \begin{cases} B_i + \tilde{S}_i & \text{w.p. } \alpha_i \\ B_i + T_{\text{tx}} & \text{w.p. } (1 - \alpha_i)(1 - \gamma_i) \\ B_i + T_{\text{tx}} + \tilde{S}_i & \text{w.p. } (1 - \alpha_i)\gamma_i \end{cases}$$

where \tilde{S}_i is a random variable with the same distribution as S_i . The MGF of B_i , denoted by $M_{B_i}(z)$, is equal to $\frac{\beta_i}{z + \beta_i}$. Therefore, we can express the MGF $M_{S_i}(z)$ of service time S_i as

$$M_{S_i}(z) = \frac{\beta_i(1 - \alpha_i)(1 - \gamma_i)e^{-zT_{\text{tx}}}}{z + \beta_i(1 - \alpha_i)(1 - \gamma_i e^{-zT_{\text{tx}}})}.$$

The first two moments of the service time, $\mathbb{E}(S_i)$ and $\mathbb{E}(S_i^2)$ can be calculated by differentiating the MGF $M_{S_i}(z)$:

$$\begin{aligned} \mathbb{E}(S_i) &= -\left. \frac{d(M_{S_i}(z))}{dz} \right|_{z=0} \\ \mathbb{E}(S_i^2) &= \left. \frac{d^2(M_{S_i}(z))}{dz^2} \right|_{z=0} \end{aligned}$$

The QNA procedure commences from the leaf nodes and converges to the base-station in a sequential manner. We first calculate the squared coefficient of variance of service time, denoted by $c_{S_i}^2$, at each node, i.e.,

$$c_{S_i}^2 = \frac{\mathbb{E}(S_i^2)}{(\mathbb{E}(S_i))^2} - 1$$

To obtain the first two moments of the interarrival times into the downstream nodes, we proceed as follows. Ignoring the packet discards at the MAC layer, the net arrival rate Λ_i at a node i is given by (this is equivalent to obtaining the first moment of the interarrival time):

$$\Lambda_i = \begin{cases} \lambda_i + \sum_{k \in \mathcal{P}_i} \theta_k & \text{for source-cum-relay nodes} \\ \sum_{k \in \mathcal{P}_i} \theta_k & \text{for relay nodes} \end{cases}$$

For any node i , let ρ_i be defined as

$$\rho_i = \begin{cases} \lambda_i \mathbb{E}(S_i) & \text{for leaf nodes} \\ \Lambda_i \mathbb{E}(S_i) & \text{for internal nodes} \end{cases}$$

We now calculate the squared coefficient of variance for the departure process ($c_{D_i}^2$) for a node which uses the squared coefficient of variance for the interarrival times at that node ($c_{A_i}^2$). Note that for the leaf nodes, $c_{A_i}^2 = 1$ due to our assumption of Poisson arrival process. Here we include the probability of a packet discard which is equal to δ_i .

$$c_{D_i}^2 = (1 - \delta_i)(1 + \rho_i^2(c_{S_i}^2 - 1)) + (1 - \rho_i^2)(c_{A_i}^2 - 1)$$

Further, $c_{A_i}^2$ can now be calculated as (this is equivalent to computing the second moment of the interarrival time at a downstream node)

$$c_{A_i}^2 = \frac{1}{\Lambda_i} \left(\lambda_i + \sum_{j \in \mathcal{P}_i} \Lambda_j c_{D_j}^2 \right)$$

Finally, the mean sojourn time at a node i is given by

$$\bar{\Delta}_i = \frac{\rho_i \mathbb{E}(S_i)(c_{A_i}^2 + c_{S_i}^2)}{2(1 - \rho_i)} + \mathbb{E}(S_i) \quad (23)$$

The end-to-end mean packet delay for a source node j , provided that the set of nodes along the path from this node to the BS is L_j , is given by

$$\Delta_j = \sum_{i \in L_j} \bar{\Delta}_i \quad (24)$$

6. Discussion on Validity of the Fixed Point Approach

We recall that we developed the analysis under the premise that the system of queues is stable, and, hence, all the steady state quantities involved in the fixed point equations exist. The next question we ask is whether, having performed the above analysis, we can use the results to conclude that the system of queues is indeed stable, which would provide a consistency in the overall approach.

We proceed by modeling the CSMA/CA multihop network analyzed in Section 3 as a Discrete Time Markov Chain (DTMC), and therefrom, deriving a sufficient condition for the stability of the network.

6.1. A DTMC Model

The system evolves *synchronously* over slotted time (this is an idealization only for the analysis in this section), with the duration of each slot being 1 symbol time⁵. We assume that the *start instants of all backoffs are aligned with the slot boundaries*. Further, the external arrival process to each source node is assumed to be Poisson; hence the number of arrivals in successive slots are independent (independent increment property). The evolution of the queueing system can then be modeled as a DTMC embedded at these slot boundaries, with the state of each queue comprising of the following components:

1. Queue length at each node
2. Residual backoff at each node
3. Residual CCA time at each node
4. Backoff stage (i.e., the number of CCA attempts for the current packet) at each node
5. Transmission state (whether the node is transmitting a packet or not) at each node

⁵Note that in a CSMA/CA network, all durations (e.g., backoff, transmission, CCA) are multiples of the symbol time

6. Retransmission stage (the number of retransmissions so far for the current packet) at each node
7. Residual packet length when a node is transmitting (in symbol times)

Note that except the queue lengths, all the other components of the state space are *finite*, while the state space of the queue lengths is countable. Hence, the state space is countable. We denote the state at time step n by $(\mathbf{X}(n), \mathbf{Y}(n))$, where $\mathbf{X}(n) = (X_1(n), \dots, X_N(n))$ denotes the queue length process, and $\mathbf{Y}(n)$ denotes the rest of the components of the state taken together. Thus, $\mathbf{X}(n) \in \mathbb{Z}_+^n$, and $\mathbf{Y}(n) \in K_y$, where K_y is a finite set of finite valued vectors.

We adopt the following convention from [27] (see Figure 4 in Section 3.1 in [27]) for counting the queue lengths at each time step: all arrivals that occur during a slot are counted immediately before the end boundary of the slot, all packets that leave the queues during a slot are counted at the end boundary of the slot, and the queue lengths are computed immediately after the end boundary of the slot, so as to account for all the arrivals and departures during the slot.

With this setup, the queueing system evolves as a DTMC over the state space $\mathbb{Z}_+^n \times K_y$. The transition probabilities of the DTMC are governed by the distribution of the arrival process, and the backoff distribution. Note that a similar DTMC model was proposed in [6] for IEEE 802.11 CSMA/CA under saturation assumption.

Now note that the operation of the system starts with all the queues empty, i.e., in the state $(\mathbf{X}(0) = \mathbf{0}, \mathbf{Y}(0) = \mathbf{0})$. We are, therefore, only interested in the communicating class of the DTMC containing the all-zero state, $(\mathbf{X}(0) = \mathbf{0}, \mathbf{Y}(0) = \mathbf{0})$. Let us denote this class by \mathcal{C}_0 .

Proposition 1. *The class \mathcal{C}_0 is closed, and aperiodic.*

Proof. See the Appendix. □

It follows from Proposition 1 that the DTMC evolving from the all-zero state is confined to the class \mathcal{C}_0 . We denote by $(\mathbf{X}(n), \mathbf{Y}(n))$, the DTMC evolving from the all-zero state.

6.2. A Sufficient Condition for Network Stability

Since we have a DTMC, by “stability”, we mean that the DTMC $(\mathbf{X}(n), \mathbf{Y}(n))$ is *positive recurrent*. Then, since $(\mathbf{X}(n), \mathbf{Y}(n))$ is also aperiodic, it can be verified that the steady state rates introduced in Section 3 *exist*. Therefore, we shall derive conditions for positive recurrence of the DTMC.

Note that the queue non-empty probability of queue i , $1 \leq i \leq N$, in the steady state is⁶

$$q_i = \lim_{n \rightarrow \infty} \Pr[X_i(n) > 0 | \mathbf{X}(0) = \mathbf{0}, \mathbf{Y}(0) = \mathbf{0}]$$

We have the following sufficient condition on q_i , $1 \leq i \leq N$, for positive recurrence of the DTMC $(\mathbf{X}(n), \mathbf{Y}(n))$.

Theorem 1. *If $\sum_{i=1}^N q_i < 1$, then the DTMC $(\mathbf{X}(n), \mathbf{Y}(n))$ (evolving from the all-zero state) is positive recurrent.*

Proof. See the Appendix. □

⁶under the condition of aperiodicity, this limit always exists

Discussion: We began this section by asking the question whether the results from our analytical model can be used to verify the initial assumption of stability. Since the analysis involves many approximations, evidently no definite answer can be given to this question. However, we have found from extensive simulations (see Section 7 for details of the simulation procedure) that the solution to the fixed point analysis models $\sum_{i=1}^N q_i$ accurately (*within an error of up to 10%*) for external arrival rates of up to about 6 packets/sec (see Figure 15 for example). This observation along

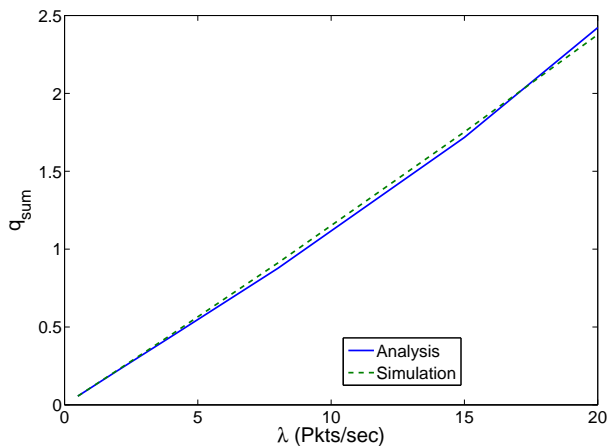


Figure 15: Plot of measure $q_{sum} \triangleq \sum_{i=1}^N q_i$ for the topology in Figure 16

with Theorem 1 can support the claim that if $\sum_{i=1}^N q_i < 0.9$ in the solution to the fixed point analysis, then we can safely assume the network to be stable. We will see later that for the practical operating range of the networks (i.e., the range of arrival rates at which the networks can provide a reasonable QoS), this sum is substantially less than 1 (See Observation 1 in Section 7.1, and Observation 3 in Section 7.2).

7. Numerical and Simulation Results

For the verification of our analytical model, we use QualNet (v4.5) simulator [28] with the default parameter values and a fixed packet size of 131 bytes⁷. However, the QualNet implementation is devoid of ACKs. Therefore, we compare results only for the *ACK-less* scenarios (although the analysis permits the modeling of ACKs). We use the following simulation models in Qualnet to declare that a packet is in error:

1. **Collision and Link Error Model:** If a receiver is receiving a packet from a node, and there is another transmission in the carrier sense range of the receiver, then the receiver concludes collision of all the packets that are intended for it. In addition, if a receiver $r(i)$ of node i receives a packet successfully from node i since there is no other packet transmission in the receiver's carrier sense range, then the received packet is concluded to be in error with probability l_i , the probability of data packet error on the link between node i and the receiver $r(i)$ due to noise.

⁷This includes 70 bytes of data, 8 bytes of UDP header, 20 bytes IP header, 27 bytes MAC header, and 6 bytes Phy header.

2. Capture Model: If a receiver is receiving a packet from a node, and there is another transmission in the carrier sense range of the receiver, then the receiver computes the PER (from the SINR) of the packet that is meant for it, and rejects the packet if the computed PER is greater than a random number generated between 0 and 1.

To increase the accuracy of simulation for each arrival rate, we generate the packets for 1500 seconds, and average the results over 25 repetitions with different random number seeds. 1500 seconds was chosen so that even at lower arrival rates ($\leq 2\text{pkt/sec}$), sufficient number of packets would be generated from each source, thereby allowing the simulation to reach steady state.

7.1. Simulation of an Example Network Topology

The network shown in Figure 16 has hidden nodes (the dependency graph is also shown). Note that all the nodes are sources (some of which also serve as relays) with identical packet generation rates that are simultaneously increased. Figures 17, 18, 19, and 20 are plots of measures α_i , γ_i , Δ_i and $p_i^{(\text{del})}$ for the nodes in Figure 16.

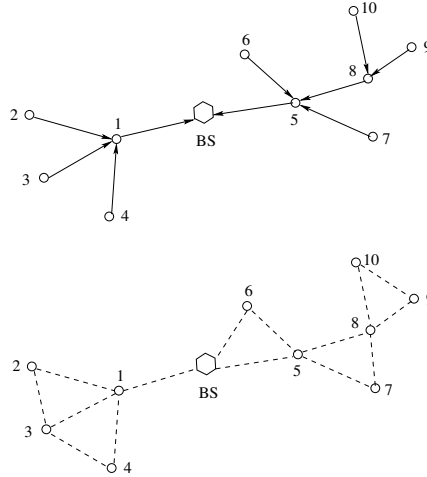


Figure 16: A 10 nodes TREE topology. Also shown is the dependency graph where the dotted lines connecting two nodes indicate that the nodes are in CS range of each other.

Observations:

1. We observe from the plots that the errors in the measures increase at higher arrival rates, i.e., the accuracy of the analysis decreases at higher arrival rates. For example, whereas for Nodes 1 to 4, and Nodes 6 to 10, the errors in the packet failure probability, γ_i , were between 1.5% and 20%, for Node 5, the error was up to 22.2% with arrival rate of 20 packets per second. We will see, however, that the practical operating range of these networks is closer to 1 measurement per second from each node, for which value the approximation is excellent.
2. The differences in the simulation measures under the Capture Model and the Collision + Link Error Model were well within 10% for all nodes at lower arrival rates, and within 20% even at an arrival rate of 20 packets per second. Hence, we conclude that the Collision + PER Model does not yield performance significantly different from the Capture Model. For other examples illustrating this fact, see [29].

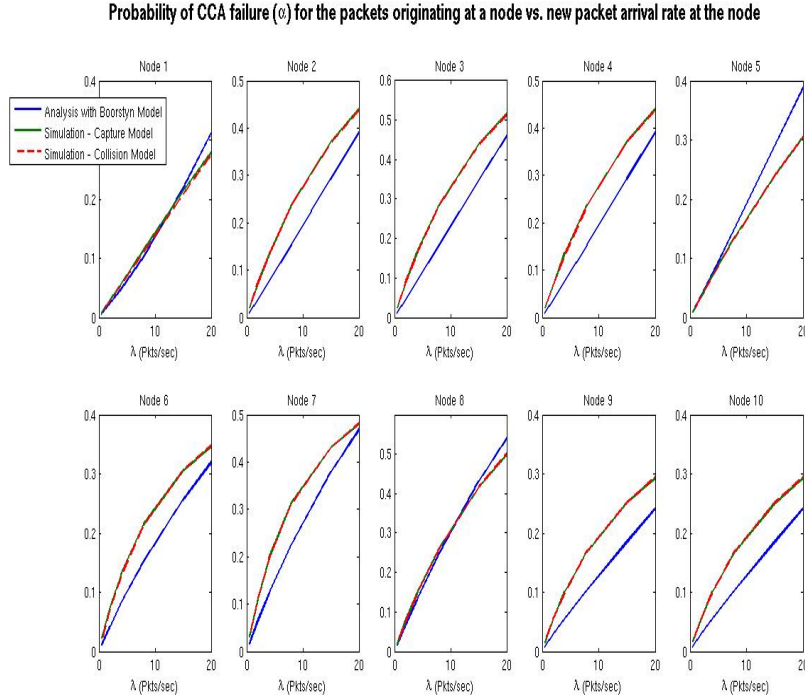


Figure 17: Plots of CCA failure rate, α , for the nodes in Figure 16; blue solid line indicates analytical results, green solid line indicates simulation with Capture Model, and Red dotted line indicates simulation with Collision + PER Model

7.2. Extensive Simulation Results

We now present a summary of an extensive simulation study, where we only use Collision + Link Error as the model. We have performed simulations with a variety of scenarios where one or more of the following features were varied: the topology, number of nodes in the network, average number of nodes in the CS range of a node, and the probability of data packet error on the link due to noise. For compact representation of the various cases, we use the following notation:

- (i) *tree-nN-CSm-PERl*: These are the cases in which the network is a tree with N nodes (excluding the base station), all of which are sources. On an average, m nodes are in the CS range of a node, i.e., $m = \frac{\sum_{i=1}^N \text{number of nodes in the CS range of node } i}{N}$. Moreover, the PER takes the same value, l , on all the links.
- (ii) *treeR-nN-CSm-PERl*: These are the cases in which the network is a tree with N nodes (excluding the base station) such that 10 nodes among them are source nodes, and the rest of the nodes are relay nodes (the “R” in “treeR” signifies that there are relays as well as sources). On an average, m nodes are in the carrier sense range of a node, and the PER takes the same value, l , on all the links.

These network scenarios, in particular, were generated randomly as follows. A 150×150 m^2 area was partitioned into square cells of side 10 meters. 10 sources were placed uniformly at random over the lattice created

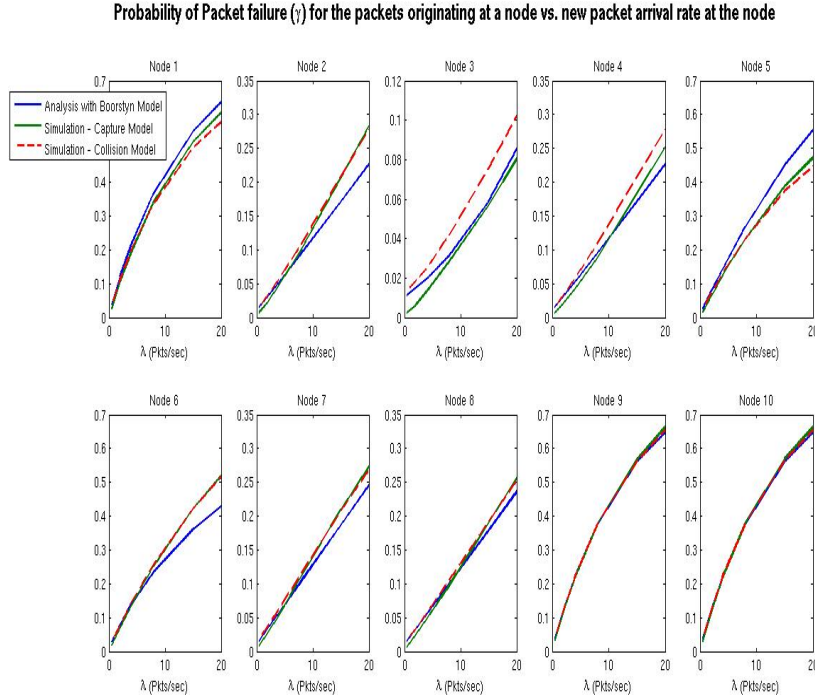


Figure 18: Plots of the packet failure probability, γ , for the nodes in Figure 16; blue solid line indicates analytical results, green solid line indicates simulation with Capture Model, and Red dotted line indicates simulation with Collision + PER Model

by the corner points of these cells. Then 30 potential relay locations were chosen uniformly at random over the area. The base station was located at (0,0). The power of the nodes were chosen so as to achieve a maximum communication range of 30 meters for a PER of 1%. On the network graph thus obtained, the SPTiRP algorithm proposed in [30] was run with a hop constraint of 6 to obtain a tree topology connecting the sources to the base station, and using a few relays. Qualnet simulation was run on the tree topology thus obtained.

Remark: Note that there are some seemingly identical entries in the Topology column of Table 4 (e.g., Sl. No. 2, 3). These scenarios are randomly generated network scenarios, as explained above. They happen to have the same number of nodes, and the same average number of CS range neighbors per node. However, the scenarios differ from one another in the position of the source nodes, the relays, as well as the routes from the sources to the sink, i.e., the associated tree topology.

- (iii) *star- nN -CS m -PER l :* These are the cases in which the network is a star with N nodes (excluding the base station), all of which are sources, positioned symmetrically on a circle centred at the base station. The m nodes nearest to a node are in the carrier sense range of it, and the PER on all the links has the same value l .
- (iv) *line- nN -CS m -PER l :* These are the cases in which the network is a line with N nodes (excluding the base station) all of which are sources. m nodes on either side of a node are in the carrier sense range of it. The PER on all the links is the same, and is equal to l .

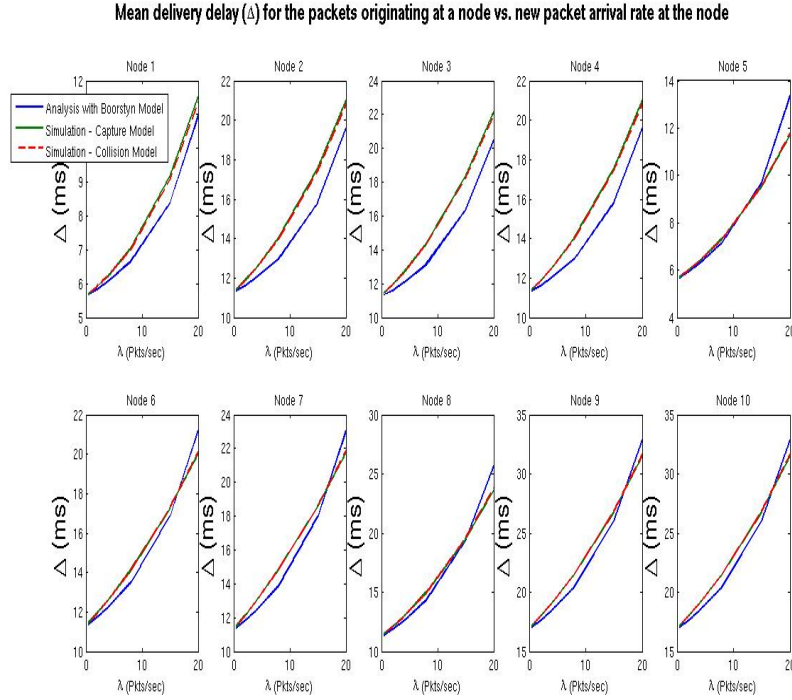


Figure 19: Plots of the mean delivery delay, Δ , for the nodes in Figure 16; blue solid line indicates analytical results, green solid line indicates simulation with Capture Model, and Red dotted line indicates simulation with Collision + PER Model

For example, the case in Figure 16 is tree-n10-CS3-PER0.01.

We compute the fractional errors in performance measures as $\frac{\text{Simulation} - \text{Analysis}}{\text{Simulation}}$. The Boorstyn et al. [25] model based analysis (see Section 4.1.2) is employed for computing $T_i^{(\text{eff})}$.

Table 4 tabulates the accuracy of the analysis compared to simulation, and also the computation times of both simulation and analysis. We use the following notation to indicate the range of errors: (i) \checkmark indicates that the error is within $\pm 10\%$; (ii) $+$ indicates overestimate by analysis with error from 10% to 25%; (iii) $++$ indicates overestimate by analysis with error more than 25%; (iv) $-$ indicates underestimate by analysis with error from 10% to 25%; and (v) $--$ indicates underestimate by analysis with error more than 25%. The entry in every row is a, b , where $a, b \in \{\checkmark, +, ++, -, --\}$ with a denoting the error summary for a range of λ where, the discard probability, $\delta_i \leq 0.01$, and b denoting the error summary for $\lambda > 2$ pkts/sec.

Observations and Discussion:

1. As can be observed from Table 4, for relatively small arrival rates at which the discard probability, $\delta_i \leq 0.01$, the errors in $\bar{P}^{(\text{del})}$ and $\bar{\Delta}$ are less than 10% in all the scenarios tested, whereas at higher arrival rates, the error in $\bar{P}^{(\text{del})}$ sometimes exceeds 25% (scenarios 7,8, and 16), which is in agreement with our earlier observation from Figure 20 as well. However, the error in $\bar{\Delta}$ even at higher arrival rates was within 10% for all but one scenario (scenario 22), where the analysis overestimated the delay within an error of 25%.

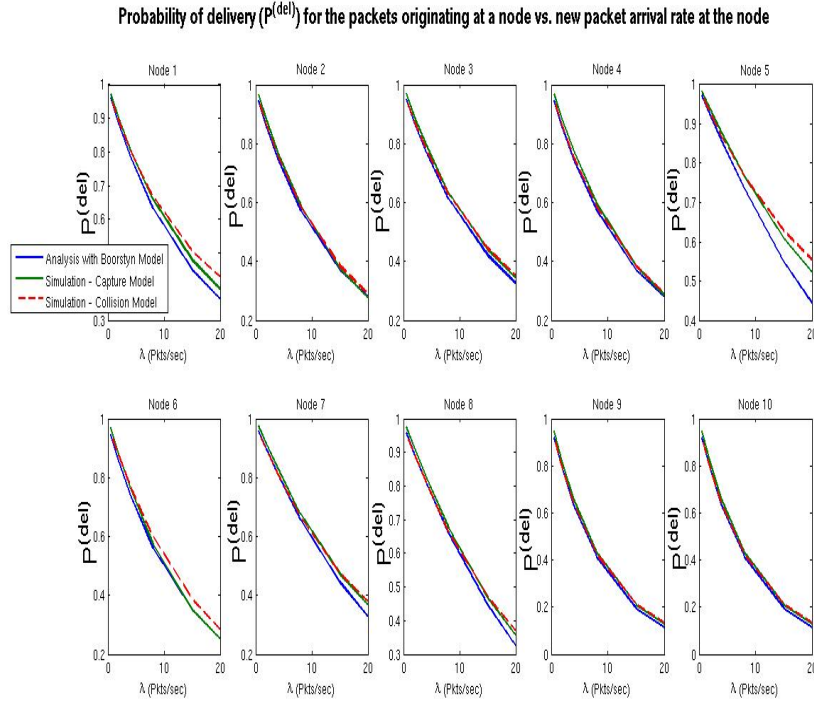


Figure 20: Plots of the average probability of delivery, $P^{(\text{del})}$, for the nodes in Figure 16; blue solid line indicates analytical results, green solid line indicates simulation with Capture Model, and Red dotted line indicates simulation with Collision + PER Model

One possible explanation for the degradation in the accuracy of the analysis at higher arrival rates can come from the discussion in Section 6. Note that at higher arrival rates, the queue non-empty probability of the nodes are typically higher, and hence the condition $\sum_{i=1}^N q_i < 1$ (see Theorem 1) is more likely to be violated, in which case, the system may not be stable, and the validity of the fixed point analysis is questionable. For example, note from Figure 15 that for the example topology in Figure 16, $\sum_{i=1}^N q_i$ approaches 1 at around $\lambda = 10$ pkts/sec, and it can be seen from Figures 17, 19, 18, and 20 that the errors in the measures are more pronounced for $\lambda \geq 10$ pkts/sec.

2. We also observe from Table 4 that the average time to analytically compute the performance measures for each arrival rate (using the Boorstyn et al. [25] model for computing $T_i^{(\text{eff})}$) was of the order of seconds, whereas the average simulation time ran into several hours.

Further, it was observed that for hidden-nodes networks with 20 nodes, and an average of 3 nodes in the CS range of a node, the analysis with Boorstyn et al. [25] model for computing $T_i^{(\text{eff})}$ takes around 13 secs per arrival rate, and if $T_i^{(\text{eff})}$ is computed using $M/D/\infty$, then the analysis takes around 9 secs per arrival rate. Hence we can trade-off between the accuracy of analysis and computation time since analysis with $M/D/\infty$ model usually incurs more error ($> 25\%$) at higher arrival rates, but less than 10% for $\lambda < 0.5$ pkts/sec.

3. Finally, it was observed that the packet discard probability at a node increases to an impractical value much

Sl. No	Topology	$\overline{P}^{(\text{del})}$	$\overline{\Delta}$	Simulation Time (Min, Mean, Max) in seconds per arrival rate until 10pkts/sec	Analysis Time in seconds per arrival rate until 10pkts/sec
1	treeR-n23-CS4-PER0.01	✓, -	✓, ✓	3120, 9720, 17580	16
2	treeR-n20-CS3-PER0.01	✓, -	✓, ✓	2220, 7740, 15060	12
3	treeR-n20-CS3-PER0.01	✓, -	✓, ✓	2580, 8640, 16800	14
4	treeR-n19-CS3-PER0.01	✓, -	✓, ✓	2460, 7680, 16020	10
5	treeR-n19-CS2-PER0.01	✓, ✓	✓, ✓	1800, 7440, 14340	10
6	treeR-n20-CS3-PER0.01	✓, -	✓, ✓	2880, 9480, 17040	16
7	treeR-n20-CS2-PER0.01	✓, --	✓, ✓	2280, 8340, 15180	12
8	treeR-n22-CS4-PER0.01	✓, --	✓, ✓	3420, 11820, 19680	36
9	treeR-n19-CS2-PER0.01	✓, -	✓, ✓	2040, 7440, 14340	14
10	treeR-n19-CS3-PER0.01	✓, -	✓, ✓	2460, 8760, 17040	17
11	treeR-n20-CS3-PER0.01	✓, -	✓, ✓	2640, 9180, 17460	17
12	treeR-n19-CS3-PER0.01	✓, -	✓, ✓	2340, 7920, 16740	15
13	treeR-n19-CS3-PER0.01	✓, -	✓, ✓	2400, 8280, 16980	12
14	treeR-n20-CS3-PER0.01	✓, -	✓, ✓	2340, 8340, 16440	11
15	treeR-n19-CS3-PER0.01	✓, -	✓, ✓	2280, 7860, 16260	11
16	treeR-n19-CS2-PER0.01	✓, --	✓, ✓	1740, 7260, 14220	9
17	treeR-n20-CS3-PER0.01	✓, -	✓, ✓	2520, 8520, 16020	11
18	treeR-n19-CS2-PER0.01	✓, -	✓, ✓	2040, 7560, 14760	10
19	treeR-n19-CS2-PER0.01	✓, -	✓, ✓	1920, 7140, 14460	9
20	line-n10-CS2-PER0.01	✓, ✓	✓, ✓	780, 6240, 12060	5
21	line-n10-CS3-PER0.01	✓, ✓	✓, ✓	1080, 7860, 13980	8
22	line-n10-CS4-PER0.01	✓, ✓	✓, +	1260, 8700, 16080	12
23	star-n20-CS9-PER0.01	✓, ✓	✓, ✓	900, 6480, 12660	36
24	star-n20-CS11-PER0.01	✓, ✓	✓, ✓	1140, 7620, 14880	48
25	tree-n13-CS2-PER0.01	✓, ✓	✓, ✓	960, 7320, 13260	5

Table 4: Error in measures $P_i^{(\text{del})}$ and Δ_i where $\overline{P}^{(\text{del})}$ and $\overline{\Delta}$ are the errors averaged over all nodes for every input arrival rate. Note that run timing of simulation and analysis are in *seconds*.

before the average delay at the node becomes substantial, e.g., when λ is around 1 measurements/sec from each node, tree-n10-CS3-PER0.01-Node-1 discards 2-5 packets for every 1000 packets whereas the average node delay is 5-6 msec. Hence, the overall packet delivery probabilities for the sources act as performance bottlenecks for these networks.

8. Network Design

In this section, we shall consider a problem of QoS constrained network design for a given positive traffic arrival rate to demonstrate the usefulness of the analytical model in an iterative network design process.

Consider the following setting: *a set of sensor nodes and a base station (BS) are deployed over an area; each sensor node is a data source, and can also act as a relay. The transmit power level of each node can be adjusted over a range.* Consider the complete graph over these nodes. Each directed link in this graph will have a certain packet error rate (PER) that *depends on the transmit power level of the sender node on that link.*

We consider the following network topology design problem over the given set of nodes: for a given traffic arrival

rate $\lambda > 0$ at the sensor nodes, *minimize the maximum transmit power level used by the sensor nodes, such that the resulting network has the following properties.*

1. Each sensor node has a path to the BS, with the PER on each link in each path being upper bounded by a predefined target threshold p .
2. For the given λ , the end-to-end packet delivery probability (i.e., the probability that a packet is not discarded) on any path is at least p_{del} .
3. For the given λ , the mean delay (computed over the successfully delivered packets) on any path is upper bounded by a predefined target d_{max} .

Note that for a design (network) to satisfy the QoS objectives for a given arrival rate, it is *necessary* that the network satisfies the QoS objectives under zero/light traffic load, i.e., in the limit as $\lambda \rightarrow 0$, in which limit a packet that enters the network departs from the network before another packet arrives, i.e., each packet traverses the network *alone*, prompting us to call this limit the *lone packet* traffic model (for a more formal proof⁸ of this fact, see [32]). We, therefore, adopt the following two step approach:

1. We first focus on the QoS constrained network design problem under the lone-packet traffic model in Section 8.1; we formulate the problem as a network design problem on graphs, and propose an algorithm to solve the problem optimally.
2. Then, in Section 8.2, we combine the lone-packet based design algorithm with the analytical tool developed in Section 3 to address the more general (and more complex) problem of QoS constrained network design for a given positive traffic arrival rate.

Note that we can ensure a given target PER on a link by ensuring that the *average* received power across the link (averaged over shadowing and fading) meets a received power target. Such a target average received power would be obtained by deriving a margin above the minimum value of received power, from statistics of shadowing and fading. If we take this worst case approach over the joint probability distribution of shadowing and fading, then for a given target PER, the required transmit power on a link is a non-decreasing function of the link length. Hence, the problem of minimizing the maximum transmit power is *equivalent to minimizing the maximum link length in the resulting network*.

8.1. Network Design under the Lone-Packet Model

Given the target PER p on any link, the mean delay on a link under the “lone packet” model can be computed using an elementary analysis (see [33]), taking into account the backoff behavior of 802.15.4 CSMA/CA, and using

⁸A formal proof is necessary for this seemingly obvious statement since, in CSMA/CA networks, in general, the performance is not monotone with the arrival rates (see, e.g., [31]); hence, the statement about the lone-packet traffic model needs to be made with care.

the backoff parameters given in the standard [2]. Then, to meet the mean delay requirement of d_{\max} on a path with h hops, we require that

$$h \leq \left\lfloor \frac{d_{\max}}{\bar{d}_{\text{single-hop}}} \right\rfloor \triangleq h_{\max}^{\text{delay}} \quad (25)$$

where, $\bar{d}_{\text{single-hop}}$ is the mean link delay computed as explained earlier.

Again, given the target link PER of p , and the number of retries r (obtained from the standard) before a packet is discarded on a link, the packet discard probability on a link can be obtained as $q = p^{r+1}$. Hence, to ensure a packet delivery probability of at least p_{del} on a path with h hops, we require that

$$h \leq \frac{\ln p_{\text{del}}}{\ln(1-q)} \triangleq h_{\max}^{\text{delivery}} \quad (26)$$

Hence, to ensure the QoS constraints under the ‘‘lone packet’’ model, we require that the *hop count on each path is upper bounded* by $h_{\max} = \min\{h_{\max}^{\text{delay}}, h_{\max}^{\text{delivery}}\}$.

Thus, the problem of QoS constrained network design under the ‘‘lone packet’’ model can be reformulated as the following graph design problem:

Given a set of source nodes Q , and a base station (BS), indexed as node 0, consider the graph $G = (V, E)$, where $V = Q \cup \{0\}$, and the edge set E consists of all feasible edges, i.e., edges with $\text{PER} \leq p$ at maximum possible transmit power. Given a hop count constraint h_{\max} , the problem is to extract from G , a spanning tree on Q rooted at the BS, such that the hop count on any path is upper bounded by h_{\max} , and the maximum edge length in the spanning tree is minimized.

We call this, the MinMax Spanning Tree with Hop Constraint (MMST-HC) problem.

The following is an algorithm to obtain an optimal solution to the MMST-HC problem. At each iteration, we prune all edges with edge length more than or same as the maximum edge length in the current feasible solution, form a shortest path tree (SPT) with hop count as cost using only the remaining edges, and keep doing this until the resulting SPT violates the hop constraint, at which point we stop and declare the last feasible solution as the final solution.

8.1.1. SPTiEP: Shortest Path Tree based iterative Edge Pruning Algorithm

- (i) **Initialize:** Set $k \leftarrow 0$, $G^{(0)} \leftarrow G$
- (ii) **Checking feasibility:** In iteration k , find a shortest path tree (SPT) $T^{(k)}$ on the graph $G^{(k)}$. Check if all the paths from the sources to the BS in $T^{(k)}$ satisfy the hop constraint.
 - If the hop constraint is not met for some of the sources in $T^{(k)}$, **STOP**.
 - * If $k = 0$, declare the problem *infeasible*.
 - * If $k > 0$, output $T^{(k-1)}$ as the final solution.
 - If the hop constraint is met for all the sources in $T^{(k)}$, proceed to the next step.

(iii) **Edge pruning:** Let $\bar{w}^{(k)}$ be the maximum edge length in $T^{(k)}$. Remove from $G^{(k)}$, all edges of length $\geq \bar{w}^{(k)}$ to obtain the graph $G^{(k+1)}$.

(iv) **Iterate:** Set $k \leftarrow k + 1$. Go to Step 2.

Since the time complexity of finding an SPT in a graph with N nodes is $O(N \log N)$, it can be easily verified that the time complexity of the SPTiEP algorithm is $O(N^3 \log N)$ for a graph with N nodes. Thus, *the SPTiEP algorithm is polynomial time.*

8.1.2. Proof of correctness of SPTiEP

We define

$\mathcal{F}^{(k)} = \{T \subset G^{(k)} : T \text{ satisfies the hop constraint}\}$: set of all feasible solutions contained in the graph $G^{(k)}$

w_{opt} : The minmax edge length, i.e., the maximum edge length in an optimal solution

Also recall the definitions of $T^{(k)}$ and $\bar{w}^{(k)}$ from the description of the algorithm.

Clearly, $\mathcal{F}^{(k)}$ is non-empty *if and only if* $T^{(k)} \in \mathcal{F}^{(k)}$.

To prove the correctness of SPTiEP algorithm, it is enough to show the following:

Proposition 2. *Given that $\mathcal{F}^{(k)} \neq \emptyset$,*

$$\mathcal{F}^{(k+1)} = \emptyset \Rightarrow w_{opt} = \bar{w}^{(k)}$$

Proof. See the Appendix. □

8.2. Network Design for a Given Positive Arrival Rate

Now we come back to the more general problem of QoS constrained network design for a given positive arrival rate. This problem is rendered much more difficult compared to its lone-packet version due to the complex stochastic interaction between contending nodes, which, unlike the lone-packet model, makes it hard to map the QoS constraints to simple explicit constraints on certain graph properties. Therefore, unlike the lone-packet version where we posed the problem as a pure graph design problem, in the positive traffic case, we shall use the analytical model explicitly in an iterative design process to evaluate for QoS, the designs obtained systematically in every iteration. While traditional network simulation tools can also be used, in principle, to evaluate a given network for QoS, *the time required for such network simulation is significantly more than that required by the analysis* (as we saw in Table 4, and shall see again in our numerical experiments in Section 8.3), and that makes network simulation, an impractical option in an iterative design process.

Also, a naive approach to the design problem would be to consider all possible trees from the given graph G , and evaluate each of them for QoS (using either the analysis, or network simulation), and choose one whose maximum edge length is minimum among all those that meet the QoS objectives. While this exhaustive search appears to be the only way that is guaranteed to obtain a feasible solution whenever there exists one, this approach has exponential time complexity (since the number of possible trees is exponential in the number of nodes), and is therefore, not practical.

We shall present below, a polynomial time algorithm for the proposed positive traffic design problem, using the SPTiEP algorithm along with the analytical model; note that because of the stochastic nature of the interaction between contending nodes in different possible networks, in absence of an exhaustive evaluation of all possibilities, the algorithm is not theoretically guaranteed to return a feasible solution whenever there exists one; however, in our numerical experiments, the algorithm was always found to return a feasible solution, and moreover, simulations confirmed that designs proposed with the analytical tool did meet the QoS requirements (see Section 8.3).

8.2.1. *Extended SPTiEP: An algorithm for QoS constrained network design at given positive arrival rate $\lambda > 0$*

As we had mentioned before, to meet the QoS objectives at a positive arrival rate, it is *necessary* (but not sufficient) to meet the objectives under the lone-packet model. Hence, our design still needs to satisfy the hop count constraint h_{\max} derived from the QoS constraints under the lone-packet model. With this in mind, we proceed as follows: if the outcome of the SPTiEP algorithm meets the QoS constraints at the given λ , then that is an optimal solution (since it is an optimal solution for the lone-packet design problem, and it is also QoS feasible for the given $\lambda > 0$). If, however, the SPTiEP outcome does not meet the QoS constraint, we need to change the design by *adding some of the edges of greater length that were pruned in course of the SPTiEP algorithm*; note that adding edges is the only option as pruning any more edge from the SPTiEP solution will cause us to violate the hop constraint, and hence the lone-packet QoS. The detailed steps are presented below.

1. **Lone-packet design:** Run the SPTiEP algorithm on the graph G to obtain a tree T_0 .
 - If T_0 does not satisfy the hop constraint, declare the problem infeasible, as we cannot satisfy the QoS objectives even for $\lambda = 0$.
 - Else, go to the next step.
2. Set $k \leftarrow 0$. Mark all edge lengths in G as *not examined*.
3. **Checking feasibility for $\lambda > 0$:** Evaluate T_k for QoS requirements (i.e., p_{del} and d_{\max}) at the given arrival rate λ , using the analytical model.⁹
 - **Stopping criteria 1:** If QoS is met, output T_k as the final solution.
 - **Stopping criteria 2:** If QoS is not met, and all edge lengths in G have been *examined*, declare the problem *possibly* infeasible.
 - Else, go to next step.
4. Let \bar{w}_k be the maximum edge length in T_k . Identify the least edge length $> \bar{w}_k$ in the graph G that is not yet *examined*; let us denote this as $w_{least,k}$.

⁹for the purposes of this design, we completely trust the outcome of the analytical model, i.e., we assume the analytical model to be 100% accurate.

5. **Edge augmentation:** Augment T_k with all edges (in G) of length $\leq w_{least,k}$, to obtain the graph G_{k+1} . Mark all the edge lengths in G_{k+1} as *examined*.

Comment: Since our objective is to minimize the maximum edge length, in Steps 4 and 5, we add back the edges pruned during SPTiEP *in increasing order of their lengths*.

6. **Redesign:** Find an SPT T_{k+1} in G_{k+1} . Observe that if T_k satisfies the hop constraint, T_{k+1} also satisfies the hop constraint.
7. **Iterate:** Set $k \leftarrow k + 1$. Go to Step 3.

Comment: In Steps 6 and 7, we check if the resulting shortest path tree in the augmented graph satisfies the QoS constraints for the given positive arrival rate.

Note that in Step 6, there could be several possible SPTs on the graph G_{k+1} ; while the total network wide traffic load, as well as the total number of nodes is the same in all of these SPTs, the individual loads on the various nodes may vary from one SPT to another, resulting in potentially different delay and delivery performance. Since it is not possible to evaluate all the possible SPTs in polynomial time, we find and evaluate only one of them, and move on to a higher power/edge length design in case of QoS not being satisfied; this may lead to suboptimal design. Moreover, for the same reason, the algorithm may not always return a feasible solution even when there exists one.

Remarks:

1. It is interesting to note that the above heuristic is basically the reverse procedure of the SPTiEP algorithm, with the additional step of evaluating each design using the analytical model. At the end of each iteration, the maximum edge length in the solution changes at most to the immediate higher value not used in previous iterations. We stop as soon as a QoS feasible solution is obtained in some iteration.
2. An alternate approach could be to start with a shortest path tree on the entire graph, and then prune edges, checking for feasibility with positive load in each iteration. But since meeting the lone-packet QoS is necessary to meet the positive-load QoS, we have adopted the above approach of first designing a network satisfying the lone-packet QoS, and then backtracking.

8.3. Numerical Experiments

To demonstrate the strengths and limitations of the analysis in an iterative design process, we performed several test runs of the proposed algorithm on randomly generated network scenarios.

30 random networks were generated in a $50 \times 50 m^2$ area as follows: The entire area was partitioned into square cells of side 10 meters. Consider the lattice created by the corner points of the cells. 10 source nodes were placed at random over these lattice points. The minimum power level of the nodes is assumed to be sufficient so that all nodes are within CS range of one another, i.e., we have a “no hidden node” scenario. We chose the target link PER to be

Table 5: Execution time comparison of the analysis based design algorithm and qualnet simulation

Scenarios	Algorithm execution time						Qualnet simulation time					
	in sec						in sec					
	SPTiEP			E-SPTiEP			SPTiEP outcome			E-SPTiEP outcome		
	Average	Max	Min	Average	Max	Min	Average	Max	Min	Average	Max	Min
30	0.0468	0.2532	0.0300	0.3827	0.5084	0.1609	2412.3	3000.6	2041.2	1765.7	2233.8	1326.6

0.01, $p_{del} = 95\%$, and $d_{max} = 25 msec$. These result in a hop constraint of $h_{max} = 5$ for the “lone packet” model. Also, the traffic arrival rate at each source was chosen to be $\lambda = 1$ packet/s, which is actually quite adequate for many wireless sensing applications.

The design algorithm using the analytical model was run on the 30 random scenarios. *In each case, the algorithm was found to return a feasible solution.*

Also, to validate the solution provided by the algorithm, we performed Qualnet simulation on both the outcome of the SPTiEP algorithm as well as the final solution for each of the 30 cases. The observations are as follows:

Observations:

1. In all 30 cases, simulations of the final solution showed QoS performance better than the target. That is, *the solution provided by the algorithm using the analytical model did actually satisfy the QoS constraints as verified by the simulation in all cases.*
2. In 12 out of the 30 cases, simulations suggested that the SPTiEP solution (“lone packet” design) met the QoS constraints even for the given $\lambda = 1$ packet/s (and hence was optimal), but the analytical model suggested otherwise. Thus, *the analytical model was somewhat conservative in its prediction of QoS*, which eventually led to a design that used more power than the optimal.
3. However, while the analytical model was conservative in its prediction of QoS, the final design met the QoS objectives and predicted the performance to within 10%, i.e., the model was *quite accurate* in its prediction.
4. Finally, we compared the time taken to run the design algorithm using the analytical model against the time taken to run the Qualnet simulations on the designed networks in the 30 cases. To save time, *in each case, we performed Qualnet simulations only for the lone packet design, and the final design obtained using E-SPTiEP algorithm.* While the algorithm was run in MATLAB 7.11 on a Windows Vista based Dell Inspiron 1525 laptop with 3 GB RAM, and 2 GHz Intel Core 2 Duo CPU, the simulations were run in Qualnet 4.5 on a Linux based Dell server with 32 GB main memory, and 3 Ghz clock speed. The results are summarized in Table 5.

From Table 5, we see that while *the execution time of the analysis based algorithm is of the order of milliseconds*, simulation of each of the designed networks takes several minutes, which indicates that *if the QoS evaluation step in the iterative design is performed using simulation instead of the analysis, the execution time of the algorithm can go well beyond an hour.*

From the above discussion, we can summarize the following strengths and limitation of the analytical model for use in a network design process.

1. **Strengths:**

- Design based on analysis is much *faster* compared to that based on network simulation, and the predictions are *accurate* to well within 10% compared to the simulation results. Such speed and accuracy also make the analytical tool a good choice for *online/field-interactive* design process, or even for as-you-go deployment of an impromptu wireless network.
- Designs provided by the analysis based algorithm do actually work, i.e., satisfy the QoS constraints in practice, as validated by simulations of the designed networks.

2. **Limitation:** The analysis, while quite accurate compared to the simulation, is somewhat *conservative in its prediction of QoS performance*. This may sometimes lead to *suboptimal* design, or even *declaration of infeasibility*, when actually there may exist a feasible solution.

9. Conclusion

We have developed an approximate stochastic model for the performance analysis of beacon-less IEEE 802.15.4 multihop wireless networks, with arrivals. The model permits the estimation of several time-average performance measures. Our model is accurate at small arrival rates (at which packet discard probability is small) in terms of the packet delivery probability, CCA failure probability, transmission failure probability, and mean delay. We calculated the mean end-to-end delays and packet delivery probabilities for each source in the network for specific packet generation rates at the source nodes. The results suggest that, for the relatively small sized tree networks that we have studied, to operate in the low-discard low-delay region, the packet arrival rates at nodes should not be greater than a packet every few seconds (e.g. a packet inter-generation time of 5 to 10 seconds at the sources). Finally, we have formulated and solved a problem of QoS constrained network design for given positive traffic arrival rate to demonstrate the usefulness of the analytical model in an iterative network design process.

Appendix

9.1. Discussion on the Poisson approximation for the conditional CCA attempt process at a node

The actual backoff distribution at a node under IEEE 802.15.4 CSMA/CA is uniform with mean *dependent on the system history*. However, an attempt at exact modeling of the CCA attempt process (or, equivalently, the backoff process) would involve keeping track of too many states, and would lead to a high-dimensional Markov chain, such as the one outlined in Section 6. Analysis of such a Markov chain is intractable, even for much simpler systems, as was demonstrated in [6]. Hence, researchers have generally taken recourse to the so called *decoupling approximation* explained in Section 3.1. As part of this approximation, many researchers have implicitly or explicitly approximated the attempt process at a node by a Poisson process (or, equivalently, a Bernoulli process for a discrete time system); see,

for example, [5], [6], [16], [12], [34], [35], [36]. The key motivation behind this approximation is that the *memoryless property* of these distributions allows one to embed a renewal process (or a Markov renewal process) in the evolution of a node; this process can then be analyzed using standard tools, thus making the analysis tractable. For example, in the seminal work of Bianchi [5], the author assumed that when in backoff, a node attempts in each slot with a certain probability independent of all the other nodes, and the system history, i.e., the attempt process conditioned on being in backoff is Bernoulli (thus, it becomes a renewal process with Geometric life time distribution). To compute the mean of this process, he used the actual distribution of the backoff process. But later, it was shown in [6] that to compute this mean, one need not consider the detailed distribution of the backoff process. Thus, in essence, the model in [5] does not use the actual distribution of the backoff process in any essential manner. In [34], [35], the authors directly assumed that the evolution at a node occurs as a renewal process; this implicitly requires the distribution of the attempt process to be memoryless, i.e., either Bernoulli or Poisson.

It was shown in all these works that despite this approximation, the resulting analyses match with simulation results with excellent accuracy. The same has been observed in our numerical experiments in this paper as well. In [37], Bordenave et al. have attempted to provide a theoretical justification as to why such approximations work so well in practice.

9.2. Proof of Proposition 1

Proof. First observe that transition from any state to the all-zero state is possible if we consider a series of “only down” transitions, i.e., no further external arrival occurs into any of the queues, and the packets leave the queues either due to successful transmission, or drop due to excessive retries or CCA failures. *No external arrival* would mean that the number of packets in the system can only decrease in each transition, and eventually we will reach the all-zero state. Note that this is possible since the maximum number of retransmission attempts, and maximum allowed CCA failures are *finite*.

Now, suppose the class is open. Then, there exists a state $i \in \mathcal{C}_0$, and a state $j \notin \mathcal{C}_0$ such that $i \rightarrow j$, but $j \not\rightarrow i$. But from our earlier argument, we know, state j can reach the all-zero state. Since both state i and the all-zero state belong to the class \mathcal{C}_0 , the all-zero state can reach state i . It follows that $j \rightarrow i$, which is a contradiction. Hence, the class \mathcal{C}_0 is closed.

Clearly, \mathcal{C}_0 is aperiodic since starting in the all-zero state, the system can remain in the all-zero state if no external arrivals occur in a slot (since the arrival process is Poisson, this event has a positive probability). \square

9.3. Proof of Theorem 1

Proof. For any $n > 0$,

$$\begin{aligned} \Pr[\cap_{i=1}^N \{X_i(n) = 0\} | \mathbf{X}(0) = \mathbf{0}, \mathbf{Y}(0) = \mathbf{0}] &= 1 - \Pr[\cup_{i=1}^N \{X_i(n) > 0\} | \mathbf{X}(0) = \mathbf{0}, \mathbf{Y}(0) = \mathbf{0}] \\ &\geq 1 - \sum_{i=1}^N \Pr[X_i(n) > 0 | \mathbf{X}(0) = \mathbf{0}, \mathbf{Y}(0) = \mathbf{0}] \end{aligned} \quad (27)$$

where, in writing (27), we have used the union bound.

Taking limit as $n \rightarrow \infty$ on both sides,

$$\begin{aligned} \lim_{n \rightarrow \infty} \Pr[\cap_{i=1}^N \{X_i(n) = 0\} | \mathbf{X}(0) = \mathbf{0}, \mathbf{Y}(0) = \mathbf{0}] &\geq 1 - \sum_{i=1}^N q_i \\ &> 0, \text{ when } \sum_{i=1}^N q_i < 1 \end{aligned}$$

Thus, if $\sum_{i=1}^N q_i < 1$, we have that $\lim_{n \rightarrow \infty} \Pr[\cap_{i=1}^N \{X_i(n) = 0\} | \mathbf{X}(0) = \mathbf{0}, \mathbf{Y}(0) = \mathbf{0}] > 0$, which, in conjunction with Proposition 1, implies that $(\mathbf{X}(n), \mathbf{Y}(n))$ is positive recurrent. \square

9.4. Proof of Proposition 2

Proof. Since $\mathcal{F}^{(k)} \neq \emptyset$,

$$w_{opt} \leq \bar{w}^{(k)} \tag{28}$$

This is because $\bar{w}^{(k)}$ is the maximum edge length of $T^{(k)}$, a feasible solution, while w_{opt} is the maximum edge length of an optimal solution.

Again, $\mathcal{F}^{(k+1)} = \emptyset \Rightarrow \nexists T \subset G^{(k+1)}$ such that T satisfies hop constraint.

But by construction (Step 3 of SPTiEP algorithm), $G^{(k+1)}$ contains in it, all trees with maximum edge length $< \bar{w}^{(k)}$. Therefore, no tree with maximum edge length $< \bar{w}^{(k)}$ is a feasible solution to the MMST-HC problem.

Hence,

$$w_{opt} \not\leq \bar{w}^{(k)} \tag{29}$$

Combining (28) and (29), it follows that $w_{opt} = \bar{w}^{(k)}$. \square

References

- [1] J.Stankovic, Wireless sensor networks, A Report, University of Virginia.
- [2] I. S. 802.15.4, Wireless medium access control and physical layer specifications for low-rate wireless personal area networks.
- [3] V. Gungor, G.P.Hancke, Industrial wireless sensor networks: Challenges, design principles, and technical approaches, IEEE Transactions on Industrial Electronics 56 (2009) 4258–4265.
- [4] C. Bordenave, D. McDonald, A. Proutiere, Performance of random medium access control an asymptotic approach, in: Sigmetrics, ACM, 2008.
- [5] G.Bianchi, Performance analysis of the IEEE 802.11 distributed coordination function, IEEE Journal on Selected Areas in Communications 18 (2000) 535–547.
- [6] A.Kumar, E.Altman, D.Miorandi, M.Goyal, New insights from a fixed point analysis of single cell IEEE 802.11 wireless LANs, IEEE/ACM Transactions on Networking 15 (2007) 588–601.

- [7] Z. J. Haas, J. Deng, Dual busy tone multiple access (dbtma) a multiple access control scheme for ad hoc networks, *IEEE TRANSACTIONS ON COMMUNICATIONS* 50 (6) (2002) 975–985.
- [8] C. L. Fullmer, J. Garcia-Luna-Aceves, Solutions to hidden terminal problems in wireless networks, in: *ACM Sigcomm*, 1997.
- [9] T. Kim, H. Kim, J. Lee, J. S. Park, B. D. Choi, Performance analysis of IEEE 802.15.4 with non-beacon enabled CSMA/CA in non-saturated condition, *Lecture Notes in Computer Science* 4096 (2006) 884–893.
- [10] L. Qiu, Y. Zhang, F. Wang, M. K. Han, R. Mahajan, A General Model of Wireless Interference, in: *MobiCom*, 2007.
- [11] T. Kim, J. S. Park, H. J. Chong, K. J. Kim, B. D. Choi, Performance analysis of IEEE 802.15.4 non-beacon mode with the unslotted CSMA/CA, *IEEE Communications Letters* 12.
- [12] C. K. Singh, A. Kumar, P. M. Ameer, Performance evaluation of an IEEE 802.15.4 sensor network with a star topology, *Wirel. Netw.* 14 (2008) 543–568.
- [13] C. Buratti, R. Verdone, Performance analysis of IEEE 802.15.4 non beacon-enabled mode, *IEEE Transaction on Vehicular Technology* 58.
- [14] J. He, Z. Tang, H.-H. Chen, Q. Zhang, An accurate and scalable analytical model for IEEE 802.15.4 slotted CSMA/CA networks, *Wireless Communications, IEEE Transactions on* 8 (1) (2009) 440–448.
- [15] M. Martalo, S. Busanelli, G. Ferrari, Markov chain-based performance analysis of multihop IEEE 802.15.4 wireless networks, *Performance Evaluation, Elsevier* (2009) 722–741.
- [16] M. Goyal, D. Rohm, H. Hosseini, K. Trivedi, A. Divjak, Y. Bashir, A Stochastic Model for Beaconless IEEE 802.15.4 MAC Operation, *SPECTS* 41 (2009) 199–207.
- [17] B. Lauwens, B. Scheers, A. Capelle, Performance analysis of unslotted CSMA/CA in wireless networks, *Springer Telecommunication Systems*.
- [18] A. Jindal, K. Psounis, The achievable rate region of 802.11-scheduled multihop networks, *IEEE/ACM TRANSACTIONS ON NETWORKING* 17 (4).
- [19] P. D. Marco, P. Park, C. Fischione, K. H. Johansson, Analytical modelling of IEEE 802.15.4 for multi-hop networks, in: *IEEE Globecom*, 2010.
- [20] S. Sen, S. De, IEEE 802.15.4 multihop forwarding throughput analysis in presence of hidden/exposed terminals, *Wireless Personal Multimedia Communications*.
- [21] M. Shyam, A. Kumar, Performance analysis of unslotted CSMA/CA in IEEE 802.15.4 networks with star and tree topologies, Master's thesis, Indian Institute of Science, Bangalore (2010).

- [22] P. Marbach, A. Eryilmaz, A. Ozdaglar, Asynchronous csma policies in multihop wireless networks with primary interference constraints, *IEEE Transactions on Information Theory* 57 (6).
- [23] P. Park, C. Fischione, K. H. Johansson, Modeling and Stability Analysis of Hybrid Multiple Access in the IEEE 802.15.4 Protocol, *ACM Transactions on Sensor Networks* 9 (2).
- [24] V. G. Kulkarni, Modeling and analysis of stochastic systems, Chapman and Hall, Ltd, 1995.
- [25] R.Boorstyn, A.Kershenbaum, B.Maglaris, V.Sahin, Throughput analysis in multihop CSMA packet radio networks, *IEEE Transactions on Communications* 35 (1987) 267–274.
- [26] W.Whitt, The queueing network analyzer, *The Bell System Technical Journal* 62.
- [27] M. K. Panda, A. Kumar, State dependent attempt rate modeling of single cell IEEE 802.11 WLANs with homogeneous nodes and poisson packet arrivals, *Performance Evaluation* 69 (2012) 413–439.
- [28] <http://web.scalable-networks.com/content/qualnet>.
- [29] S. M. Ladwa, An Analytical Performance Model for Beaconless IEEE 802.15.4 Multi-Hop Networks, M.E thesis, Indian Institute of Science (June 2012).
- [30] A. Bhattacharya, A. Kumar, A shortest path tree based algorithm for relay placement in a wireless sensor network and its performance analysis, *Computer Networks* 71 (2014) 48–62.
- [31] S. Borst, M. Jonckheere, L. Leskela, Stability of Parallel Queueing Systems with Coupled Service Rates, *Discrete Event Dynamic Systems* 18 (2008) 447–472.
- [32] A. Bhattacharya, A. Kumar, QoS Aware and Survivable Network Design for Planned Wireless Sensor Networks, Tech. rep., available at arxiv.org/pdf/1110.4746 (2011).
- [33] A. Bhattacharya, Node Placement and Topology Design for Planned Wireless Sensor Networks, M.E thesis, Indian Institute of Science (June 2010).
- [34] M. Garetto, T. Salonidis, E. W. Knightly, Modeling per-flow throughput and capturing starvation in csma multi-hop wireless networks, *IEEE/ACM Transactions on Networking* 16 (4) (2008) 864–877.
- [35] K. Medepalli, F. A. Tobagi, Towards performance modeling of ieee 802.11 based wireless networks: A unified framework and its applications, in: *IEEE Infocom*, 2006.
- [36] G. Kuriakose, S. Harsha, A. Kumar, V. Sharma, Analytical models for capacity estimation of IEEE 802.11 wlans using DCF for internet applications, *Wireless Networks* (2009) 259–277.
- [37] C. Bordenave, D. McDonald, A. Proutiere, Random multi-access algorithms, a mean field analysis, in: *Forty-Second Annual Allerton Conference on Communication, Control, and Computing*, 2005.