Hindawi Publishing Corporation
International Journal of Digital Multimedia Broadcasting
Volume 2008, Article ID 586067, 16 pages
doi:10.1155/2008/586067

Research Article

Real-Time Communications in Large-Scale Wireless Networks

Liang Song and Dimitrios Hatzinakos

Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, 10 King's College Road, Toronto, ON, Canada M5S 3G4

Correspondence should be addressed to Liang Song, lsong@ieee.org

Received 25 March 2008; Revised 13 July 2008; Accepted 17 September 2008

Recommended by Massimiliano Laddomada

There is an emerging need for realizing real-time quality of service (QoS) over multihop wireless communications in large-scale wireless networks. The applications can include wireless mesh infrastructure for broadband Internet access supporting multimedia services, visual sensor networks for surveillance, and disaster-relief networks. However, a number of challenges still exist as revealed by recent works, where the dataflow QoS performance such as throughput and end-to-end delay can degrade fast with the number of wireless hops. We propose to use large-scale cognitive networking methods to resolve the wireless multihop challenges. By the cognitive-networking concept, data packets travel along opportunistically available paths in the network with opportunistically available spectrum in every hop. Reliable end-to-end communications can be achieved for real-time services, where we show that (1) dataflow throughput can be independent of any number of wireless hops, (2) end-to-end delay and delay variance increase linearly with the number of wireless hops, and (3) delay variance decreases to zero with higher network density. These results are supported by analysis, simulations, and experiments.

Copyright © 2008 L. Song and D. Hatzinakos. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

Large-scale wireless networks, for example, mobile ad hoc networks, wireless mesh networks, as well as wireless sensor networks, have been receiving significant attention in the past few years. Recently, there has been growing interest in realizing real-time quality of service (QoS) over multihop wireless transmissions in large-scale wireless networks. For example, the driving commercial applications include wireless mesh infrastructure for broadband Internet access [1], supporting Voice over Internet Protocols (VoIP) or Multimedia over IP, and wireless sensor network with real-time event detection/reporting, especially in video surveillance, that is, visual sensor networks [2].

Compared to the classical many-to-one (starlike) networks, such as cellular phone networks and WiFi- (IEEE 802.11) [3] enabled wireless local area networks (WLANs), a wireless mesh-based infrastructure for broadband access has the advantage of much lower capital and operational costs and can achieve higher service coverage. As one example, municipal WiFi mesh networks are rolling out in a number of major cities, providing wireless web surfing and email services to general public. Considering the popularity of Internet voice and video, further developments for supporting VoIP within the wireless mesh infrastructure are highly demanded. On the other hand, visual sensor networks appear to be one of the killer applications for wireless sensor networks, since a lot of major cities have installed public-administrated surveillance cameras for various purposes. As of now, those cameras are normally attached by cables. By adopting wireless connections, there will be significant cost reduction in the installation, which can enable much denser deployment. Moreover, related applications also include the disaster-relief wireless networks, with ad hoc formation, where the real-time information delivery can be of life and death importance.

However, many challenges still exist in realizing real-time QoS over multihop wireless networks. It has been well known that dataflow throughput, end-to-end delay, and delay jitter can degrade fast when the number of wireless hops increases [4, 5]. The limitation can be primarily introduced by dynamic network-resource availabilities including both spectrum bandwidth and wireless nodes/radios. Specifically, random spectrum availability is based on wireless fading

and interferences prevailing in unlicensed bands (e.g., ISM bands); and random radio availability is due to dynamic traffic load (congestions) and other factors such as radio failure. In traditional wireless networking, the media access control (MAC) layer sets up "logic wired links" over wireless media, which assume predetermined spectrum availability. On top of it, the network layer runs ad hoc routing protocols which assume predetermined radio availability and network topology. It is difficult to have effective QoS models in the traditional network protocol stack [6], when both spectrum bandwidth and radio availabilities cannot be predetermined in large-scale wireless networks.

In order to resolve the wireless multihop limitation, there have been proposals to install multiple radios in one single wireless node [4] and design network routing protocols based on wireless link status, as well as other combined metrics [5]. Although evidence shows that improvements could be obtained, these approaches do not address real-time QoS. Moreover, existing approaches can generally suffer from the lack of scalability. Typically, the complexity of existing protocols can grow fast with the number of wireless nodes neighboring each other (higher network density), since there are a large number of links to be managed and/or tracked on every node. Other research work that investigates real-time QoS in wireless networks could be based on inappropriate assumptions for large-scale wireless networks, such as the assumption on fixed link bandwidth [7].

In this paper, we apply the methods of large-scale cognitive networking [8] to resolve the wireless multihop limitation. The concept of cognitive networking opportunistically utilizes network resources including both spectrum bandwidth and wireless node/radio availability to realize reliable communications in large-scale wireless networks. The adopted network architecture is the Embedded Wireless Interconnect (EWI) [8–11], where wireless linkage is redefined introducing *abstract wireless links*. By definition, abstract wireless links can be arbitrary functional abstractions of mutual cooperations among proximity wireless nodes, where the categories can include broadcast, unicast, multicast/anycast, and data aggregation, among others. Therefore, *wireless link modules* are designed as the units for realizing different abstract wireless links at individual wireless nodes. The architecture of EWI is then developed upon two layers, where the bottom wireless link layer supplies a library of wireless link modules; and the upper system layer can organize the modules for achieving effective application programming interface (API).

Conforming to the cognitive networking concept, both the operating spectrum and the participating nodes of an abstract wireless link is opportunistically decided based on their instantaneous availabilities. By the unicast module design in the cognitive networking, data packets can take opportunistically available paths in the network from source to destination, with opportunistically available spectrum in every hop. In comparison, dynamic changes in spectrum and radio availability can create bottlenecks on the routing path of traditional wireless networking, where such bottlenecks are all resolved by the proposed methods. Therefore, reliable end-to-end communications can be achieved, by the opportunistic utilization of network resources. The presented unicast module design is also implemented by cognitive-radio prototyping [12]. In the performance analysis, real-time QoS metrics including throughput, expected end-to-end delay, and delay variance, are studied. The following propositions are obtained.

(i) All the investigated real-time QoS metrics improve with larger network scale, and the complexity on individual nodes remains constant independent of any network scale. It is also shown that the performance dynamics, as indicated by end-to-end delay variance, can diminish to zero with higher network density.

(ii) Reliable communications can be achieved over a large number of wireless hops. In principle, the throughput can be independent of the source-to-destination distance; while the expected end-to-end delay, and the delay variance, increase linearly with the source-to-destination distance.

(iii) Considering the resource constraints, such as energy consumption or network capacity, we identify that radio transmitting power can be an effective control knob for traffic prioritization in dealing with the tradeoff between achievable QoS performance and resource consumptions.

Therefore, the two major contributions of the paper are the following: (1) the first design & implementation of real-time communications for large-scale wireless networks; (2) the QoS performance analysis of the designed real-time unicast module. In what follows, related works are surveyed in Section 2; the unicast wireless link module design and implementation are described in Section 3; the performance analysis of real-time QoS metrics is presented in Section 4; the simulation and experiment results are reported in Section 5 and 6, respectively; the conclusion is given in Section 7.

2. RELATED WORK

Related works in cognitive networking and the EWI network architecture have previously appeared in [1, 8–12]. As a pilot architectural reference model, EWI was first introduced in the application-specific studies in wireless sensor networks [10, 11]. In [9], recent research works in the cross-layer design of wireless sensor networks are surveyed, where it is suggested that EWI can be an unified design architecture. The concept of large-scale cognitive wireless networks has been presented in [8], where EWI was adopted as the network architecture. In [1, 12], the EWI is further utilized to construct a ubiquitous wireless mesh infrastructure for broadband Internet access.

The unicast design can also be related to existing research works in opportunistic routing [13–17]. In principle, opportunistic routing schemes deal with the codesign of routing and MAC protocols. However, previous research works usually deal with single-hop performance metrics, while

the contribution of this paper deals with end-to-end QoS for real-time communications. Moreover, contributions are made in large-scale cognitive networking, with a report of experiment results.

In [13], the protocol selects the next hop relay node by a slotted ACK (acknowledge) mechanism. Having successfully received a data packet, the node calculates a priority level, which is inversely proportional to the distance between the node and the packet destination. The node with the highest priority is then selected distributively as the next hop relay node. In [14, 15], a technique named GeRaF was proposed, where the node forwarding region is divided into smaller sections with different priorities, as decided by the forwarding distances. The performance analysis of GeRaF included the single-hop energy consumption and latency tradeoff, as well as the multihop count number. The study in [16] investigated the architectural aspects of opportunistic routing, where the performance of the component design is analyzed based on single-hop metrics and light-traffic scenario. In [17], the authors discussed the opportunistic relay node selection in a two-hop transmission scenario. The relay-node priority criterion is based on the gain of source-to-relay channels and relay-to-destination channels.

3. MODULE DESIGN AND IMPLEMENTATION

A unicast wireless link is an abstraction of the proximity wireless nodes cooperation for unicasting. Since unicasting usually involves the information delivery from source to destination over multiple hops, the unicast module design decouples the end-to-end network behavior into proximity wireless nodes cooperation. In the following, some elements of the module design and implementation are presented. The module state-diagram is then shown.

3.1. Network address

Network address is related to the context, subject to a "cost of delivery" criteria. Let d and s denote the network addresses of the destination node and the source node, respectively. Given the destination d of a data packet at one wireless node n, a local parameter $c_{n,d}$ is assumed obtainable, which indicates the approximate or average cost of delivering the packet to the destination from the node n independent of dynamic changes. In large-scale wireless networks, the cost of delivery $c_{n,d}$ is usually a function of the distance from one local node n to the destination d.

In location-centric networks (e.g., [10]), where wireless nodes are aware of their own locations, for example, by global position systems (GPS) or triangulation estimations, the network address of one node n is solely decided by the node n location coordinates $\mathbf{L}_n$. Given the destination coordinates $\mathbf{L}_d$ of a packet, $c_{n,d} = |\mathbf{L}_n - \mathbf{L}_d|$ can be readily calculated in geometry, defined as the specific distance.

In data-centric networks (e.g., [18]), the network address can be decided by the application specific data. The cost of delivery $c_{n,d}$ is the application data gradient, which can be assumed as a monotonically increasing function of the distance between n and d.

In a data-collecting or fusion network, for example, wireless sensor networks, the sink (data collector) can broadcast a number of identity advertisement packets, which is thereon flooded in the network, by broadcasting. Every node can count the average smallest number of hops from the sink, on receiving the advertisement packets. The count number can be used as $c_{n,d}$ for one node n. A similar approach appeared in [19] yet in a more general scenario, where the logic address, that is, a vector of the estimated distance or hop number to a group of anchors, is maintained for every wireless node. New nodes joining in the network can estimate their own logic address by acquiring the address of neighboring nodes.

Although other types of address can be applicable, we use the location address in this paper for analysis, simulations, and experiments.

3.2. Radio implementation

The terminology of cognitive radio was first suggested in [20] as an ideal-omnipotent radio which can take all the parameters into consideration for user-centric communications. Cognitive radio was later reviewed in [21, 22] as the radio with dynamic spectrum access. In [8], two propositions are further suggested for the cognitive radio for large-scale wireless networks.

(i) The radio can opportunistically sense the spectrum resource, so that the selected spectrum usage will not be interfering with other on-going wireless communications.

(ii) The radio can opportunistically poll one or more other proximity radios onto the selected spectrum, so as to realize certain types of local cooperations.

The above two propositions can extend the concept of pure cognitive radio to cognitive networks, which implement both dynamic spectrum access and dynamic radio access. In the network architecture EWI, a wireless node can initiate an abstract wireless link, that is, certain types of local cooperations among a number of proximity wireless nodes. Hence, both the set of wireless nodes and the operating spectrum are opportunistically decided for an abstract wireless link. Further, the initiated link will not be interfering with other wireless communications, in accordance with the first cognitive proposition.

In the test-bed development of the paper, we prototyped an experimental tone-based cognitive radio [8, 12]. Specifically, the radio can access a group of predetermined data channels, where every data channel is also associated with two distinctive frequency tones, that is, one sensing tone and one polling tone. The sensing and polling tones are at distinctive frequency different from the data channel. The radio hardware is therefore composed of two transceivers, which are the tone transceiver and the data transceiver, respectively. On initiating an abstract wireless link, the radio senses for an available channel, with the vacant data channel and the vacant sensing/polling tones. It then broadcasts the polling tone associated to the selected channel, to poll its

surrounding nodes. On detecting the rising edge of the polling tone, the surrounding nodes can decide to join in the initiated abstract wireless link based on their autonomous availability. On joining in an abstract wireless link, the radio of the surrounding nodes also broadcasts the sensing tone. As such, both sensing and polling tones protect wireless link modules from spectrum interferences. We note that the use of frequency (busy) tones has also appeared in numerous works of MAC protocols [23, 24]. However, here we have the first implementation over multiple channels, and such implementation is in the context of large-scale cognitive wireless networks.

The unicast wireless link module is programmed on the cognitive radio prototype (shown in Figure 1). The experiment hardware is a stack of two radio boards (TELOSB [25]). The radio board is composed of one IEEE 802.15.4 [26] compatible transceiver, Chipcon CC2420, one TI MSP430 Microcontroller, and the program interface. The top radio board is utilized as the tone transceiver, while the bottom radio board is utilized as the data transceiver. The two are wired up by the digital interface shown in Figure 1. Three independent data channels in the 2.4 GHz band are allotted to the radio platform, while the associated tones are also in the band of 2.4 GHz. The unicast module is programmed in the firmware of TI MSP430.

3.3. Packet relaying process

Under an initiated wireless link, the packet relaying process picks one available relay node and sends the data information from the source or a previous relay node to the new relay node.

Four types of packets are used in the packet relaying process, which are RTS, CTS, DATA, and ACK, respectively. RTS, CTS, and ACK are control packets, while DATA is the data information unit. Let node n_p denote the relay node at the hop p, the next hop relay node n_{p+1} is found by the following procedure, which incurs local wireless nodes cooperation in proximity, that is, a unicast wireless link.

First, the node n_p broadcasts a RTS, including the module type (unicast), the destination address d, and the self-address n_p. Upon receiving the RTS at one node n, if the condition $c_{n,d} < c_{n_p,d}$ is satisfied, the node n initializes a timer, with the timeout period as $T_d(n_p, n, d)$. $T_d(n_p, n, d)$ is inversely proportional to $c_{n_p,d} - c_{n,d}$, which is determined by the specific physical radio technology, for example,

$$T_d(n_p, n, d) = \frac{C_0}{c_{n_p,d} - c_{n,d}} + \text{SIFS}, \tag{1}$$

where C_0 is a constant, and $n \neq d$. SIFS is the smallest (minimal) inter frame space (delay constant), which is composed of the module processing time and the transceiver RX/TX switch time. In our current test-bed implementation, $T_d(n_p, n, d)$ is also slotted according to the minimal carrier-sensing time. Note that the condition $c_{n,d} < c_{n_p,d}$ is enforced, so that the cost of delivery to the destination is always descending which prevents any loop.

The node n then backs off and monitors the data channel for the period $T_d(n_p, n, d)$. If the data channel is free during that period, the node n replies with a CTS, declaring itself as the next hop relay node n_{p+1}. Otherwise, the node n quits the unicast module. As such, the node with the smallest cost $c_{n,d}$ should be elected as the next hop relay node n_{p+1}. After having received the CTS, the node n_p transmits the DATA packet, and after receiving the DATA packet, the node n_{p+1} replies an ACK, completing the round of relaying. In the case that more than one node send out their CTS simultaneously, the transmissions will collide on some specific bits, for example, the node address, where the CTS packets are different. Since this collision can be detected by certain mechanism, such as cyclic redundancy check (CRC), at n_p, the node n_p can use an appended procedure to differentiate one node as n_{p+1}, either before or after the DATA transmission.

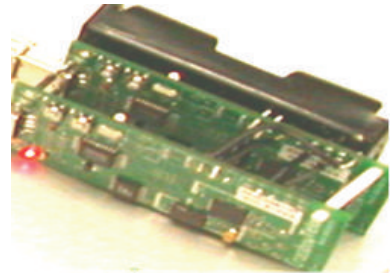


Figure 1: Cognitive radio prototype.

The described mechanism is illustrated in Figure 2. In the last hop, the destination d is always assigned with the minimal delay, that is, $T_d(n_p, d, d) = \text{SIFS}$.

3.4. Module state diagram

The state diagram of a single unicast module in the wireless link layer is shown in Figure 3, which can give a summary of the implementation. The definitions of the states and the associated transferring branches can be self-explainable, in accordance with previous descriptions.

The wireless link layer stays in the IDLE state, when no module is involved. It can initiate a unicast module, for example, on receiving the command from the system layer, so as to send certain amount of information (i.e., contained in one DATA packet) to a specified destination. In the IDLE state, the wireless link layer also monitors the set of polling tones, which are associated to the predetermined group of data channels. On detecting the rising edge of one polling tone, it sends one module request to the system layer. Upon approval, the wireless link layer listens to the RTS packet on the data channel, which is associated with the detected polling tone. Otherwise, it goes back to the IDLE state. Furthermore, after having received one DATA packet, the unicast module automatically relays it, if the current node is not the destination.

Table 1 shows an example set of primitive functions on the interface between the system layer and the wireless link layer, as related to the unicast module. The parameter *Priority* in Table 1 can be used for specifying the traffic class, which will be explored further in Section 4.5.3.

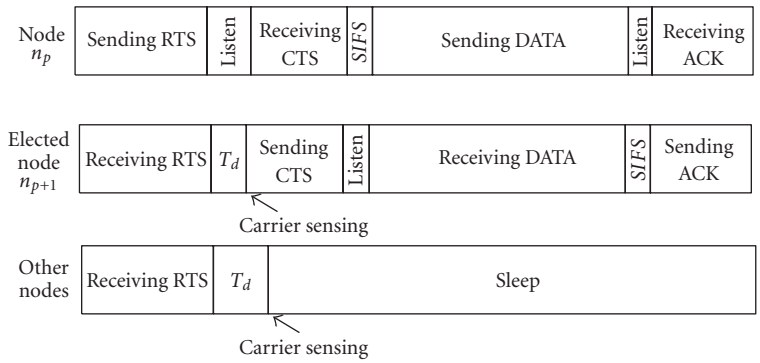


Figure 2: Packet relaying mechanism.

3.5. Discussion

By opportunistically dealing with network resources including both spectrum bandwidth and radio availability, data packets take opportunistically available paths with opportunistically available spectrum in every hop. Since the unicast module operates opportunistically in proximity wireless nodes, the end-to-end performance of real-time QoS, can be statistically assured, adding up the sequential proximity operations. This discussion can provide an intuitive explanation of the performance analysis results in Section 4.

Specifically, random spectrum availability is handled by the cognitive radio implementation, where a source or relay node tries to find an available data channel for initiating the unicast wireless link. Important aspects of random node/radio availability are as follows: the node deployment, mobility, and traffic congestion. In general, it indicates that the node n_p, that is, the pth hop relay, would be uncertain about the availability of the next hop relay node, before the RTS probing. The RTS/CTS exchange in the module design opportunistically finds the next hop relay, in a group of available candidates.

In traditional computer networks, dynamic traffic load is limited by network congestion control. The network layer drops overflowed packets by queueing management, which is then detected by the transport layer. Such methods have been known to incur problems in wireless networks [27]. By the real-time unicast module, the paradigm of the classical "network queueing" could be transferred to "queueing in network." In an individual relay node, a DATA packet is forwarded automatically without lengthy buffering in one local queue. Sequential DATA packets nominally take opportunistically decided paths from the source to the destination, which are buffered in the network, instead of any predetermined nodes.

Therefore, it can be intuitive to conceive that the QoS performance metrics should improve with larger network scale, since the network density contributes to the diversity (or more radio resource) that can be opportunistically exploited. This agrees with other theoretical results, for example, in [28], where it is stated that summed network transport capacity should increase with network scale, on the order of $\Theta(\sqrt{N})$, where N is the number of nodes.

Moreover, in the proposed method, any node does not need to consider or know every other possible node as a candidate for the next hop. In the formation of abstract wireless linkage, the participating nodes are decided by their autonomous availability. The unicast module design also assures that there would only be one RTS response in most cases by the carrier-sensing mechanism. Therefore, the module complexity at individual nodes remains constant, independent of the network scale.

4. PERFORMANCE ANALYSIS OF REAL-TIME QoS

The performance analysis mathematically quantifies the major propositions in the paper. The results can also provide useful network planning guidance for deployments.

4.1. Objectives and approaches

In supporting real-time QoS, the metrics of throughput, expected end-to-end delay, and delay variance, are studied for the unicast wireless link module. Two types of resource consumptions are considered, which are network capacity and network energy consumption. Therefore, we centrally explore the following three questions. (1) How do the QoS metrics change with network scale? (2) How do the QoS metrics change with the source-to-destination distance? (3) How to control the resource allocation, so as to achieve the QoS requirements?

Given the source s and the destination d of a unicast dataflow, let $l = |\mathbf{L}_s - \mathbf{L}_d|$ denote the source-to-destination distance. Analytical results are obtained under the asymptotic condition of long-range unicasting, $l \to \infty$, where the justification can be that long-range unicasting presents the worst-case scenario of the network performance under practical interests. On the other hand, concise analytical formulation of the QoS metrics can be obtained under this long-range asymptotic condition, which can provide analytical insights to previous questions.

4.2. Analytical models

In the performance analysis, models about node distribution, wireless channel, power consumption, bandwidth availability, and the IDLE probability are adopted.

Table 1: Exemplary primitive functions related to the unicast module.

Function	Description	Direction
SendUnicast(*DA, DATA, Priority*)	Command of sending the unicast DATA to the destination address DA;	From system to wireless link
StatusUnicast(*&Status*)	Return the command status of sendUnicast (· · ·), for example, Success or Fail;	From wireless link to system
IndicateUnicast(*SA, DATA, Priority*)	Indicate the received unicast DATA from the source address SA;	From wireless link to system
ModuleRequest()	Request for listening to the RTS, that is, defined in the unicast module;	From wireless link to system
ModuleResponse(*&Status*)	Return the request status of moduleRequest(), that is, Permitted or Denied;	From system to wireless link
GetState(*&State*)	Get the current state.	From system to wireless link

4.2.1. Node distribution model

The node distribution is modeled as a two-dimensional (2D) Poisson process [29], with the node density λ. That is, given an area $\mathbf{A}$ of the size $|\mathbf{A}|$ in the field, the number of nodes in the area follows Poisson distribution with the parameter $\lambda|\mathbf{A}|$. The Poisson modeling can be typical for random node distribution and/or random mobility.

4.2.2. Wireless channel model

Given an arbitrary transmitting node n, and a receiving node m, the successful transmission probability, in general, is a function of the transmitting power P_t, the radio data rate R, and the distance $\zeta = |\mathbf{L}_m - \mathbf{L}_n|$. We denote this function as $f(P_t, R, \zeta)$.

For example, if small-scale Rayleigh fading [30] is assumed, the channel model can be given by

$$f(P_t, R, \zeta) = \text{Prob}\left(\frac{P_t \cdot G \cdot \xi}{N_0 \cdot R \cdot \zeta^{\alpha}} \geq B\right) = e^{-(N_0 \cdot R \cdot \zeta^{\alpha} \cdot B)/(P_t \cdot G)}. \tag{2}$$

In (2), G is a propagation-loss constant; N_0 is the receiver-noise power spectrum density; α is the path loss component in wireless channel [30]; B is a threshold constant representing the receiver sensitivity; and ξ is a unit-mean exponential random variable. This model will also be utilized to obtain the numerical results in Section 5.

4.2.3. Power consumption model

Referring to the state diagram in Figure 3, the node power consumption at the IDLE state is denoted by P_I, that is, the low-power monitoring of the polling tones; the power consumption at the states of "Send DATA", "Send RTS", "Send CTS", and "Send ACK", is modeled by

$$P_S + (1 + \beta) \cdot P_t, \tag{3}$$

where P_S is the transmitter circuits power consumption; β is a constant decided by the RF power amplifier efficiency [31]; and P_t is the transmitting power as defined perviously. The node power consumption in other states or time is modeled by a constant P_R, which denotes the receiving (or idle listening) power consumption.

4.2.4. Bandwidth availability and the IDLE probability

When initiating a unicast module, we assume that the node can always find an available channel by the cognitive radio without significant delay. The IDLE probability is the probability that a given node is in the IDLE state (shown in Figure 3), and can be engaged in a unicast wireless link being initiated by other nodes around it. We assume that there is a fixed (or lower bounded) IDLE probability "ρ", in the analysis.

With limited network capacity and arbitrary traffic load, these two assumptions need to be ensured by an appropriate call admission control (CAC) mechanism. On the other hand, given analogous experiences of the popular video streaming on peer-to-peer overlay networks, it could also be appropriate to assume no CAC mechanism, while the traffic volume demanding can always fall below the planned (or unplanned) network capacity. In large-scale wireless networks, the capacity could be of relatively cheap resources, given abundant wireless nodes and the use of unlicensed band, while the traffic load is not necessarily related to the number of wireless nodes. These considerations give the optimization formulations under different types of resource consumptions in Section 4.5.

4.2.5. Other parameters and notations

The packet lengths (bit) of RTS, CTS, DATA, and ACK are denoted by L_R, L_C, L_D, and L_A, respectively. T_S denotes the time delay in the "channel sensing" state in Figure 3, and T_B denotes the time delay in the "Backoff T_d" state. We also use the notations $E(\cdot)$ and $\sigma^2(\cdot)$ to denote the mean and the variance of a random variable.

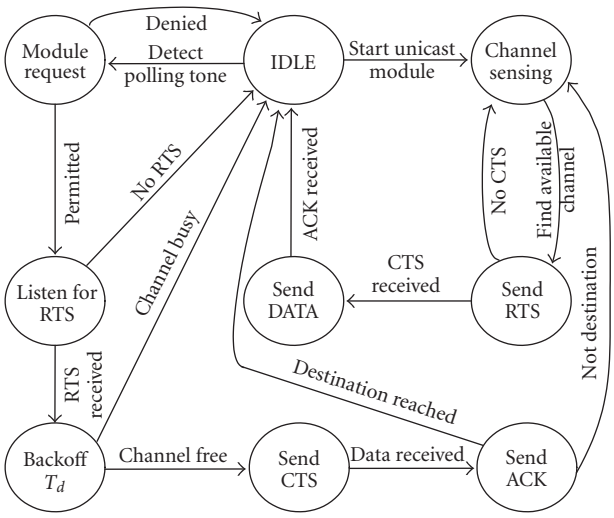


Figure 3: Wireless link layer state diagram.

4.3. Iterative performance analysis

"Iterative performance" can be similar to "single-hop performance." However, with the notation of "iterative performance," we also consider the scenario where the next hop relay node is not found in the transferring branch from "Send RTS" to "channel sensing" of Figure 3. The reason resides in the random node availability, for example, no potential relay candidate is in the IDLE states. Due to this difference, the terminology "iteration" instead of "hop" is used in the following analysis, and "iterative delay" indicates the time delay of an iteration, whereas "iterative energy consumption" indicates the network energy consumption in one iteration.

Particularly, consider that there are totally I iterations from the source s to the destination d. Let $\{n_i \mid 1 \le i \le I\}$ denote the revelent set of nodes, which initiates the unicast wireless link. Obviously, there is $n_1 = s$, and n_I sends the DATA packet directly to the destination d. The iteration number I is therefore a random variable.

4.3.1. Iterative delay

Let τ_i denote the iterative delay at an iteration i. (Other possible models are also discussed in Appendix C, which contribute to the same end-to-end performance analysis results.) If the DATA transmission is completed in the iteration, τ_i can be calculated by the formula $(L_R + L_C + L_D + L_A)/R + T_S + T_B$. Otherwise, if the DATA transmission is uncompleted since no relay is found, τ_i can be given by $L_R/R + T_S + T_B$. Obviously, for the last iteration ($i = I$), the DATA transmission is completed.

When $i < I$, the probability of uncompleted DATA transmission depends on the relay candidates availability on the forwarding plane. Due to the Poisson node distribution model with the density λ, and the node IDLE probability ρ, the distribution of relay candidates conforms to an inhomogeneous Poisson process with the density function $\lambda\cdot\rho\cdot f(P_t, R, \zeta)$ on the forwarding plane, where ζ indicates the distance to the node n_i. Therefore, the "uncompleted" probability is decided by the "equal-zero" probability of a Poisson random variable with the mean $\lambda\cdot\rho\cdot\Omega$, which is given by $e^{-\lambda\rho\Omega}$. Ω is always positive, and is given by the integration of the channel model $f(\cdot,\cdot,\cdot)$ on the forwarding plane:

$$\Omega = \int_0^{\infty}\int_{-\infty}^{\infty} f\left(P_t, R, \sqrt{x^2+y^2}\right) dx\, dy. \tag{4}$$

According to the above analysis, $\{\tau_i \mid 1 \le i < I\}$ can be modeled as a set of random variables with identical distributions, where the expectation $E(\tau_i)$ is

$$E(\tau_i) = \frac{L_R}{R} + T_S + T_B + \frac{L_C + L_D + L_A}{R}\cdot\left(1 - e^{-\lambda\rho\Omega}\right). \tag{5}$$

4.3.2. Iterative energy consumption

Let ϵ_i denote the iterative energy consumption at an iteration i. (Other possible models are also discussed in Appendix C, which contribute to the same end-to-end performance analysis results.) If the DATA transmission is not completed in the iteration, the energy consumption of the node n_i is $(T_S+T_B)\cdot P_R+L_R/R\cdot[P_S+(1+\beta)P_t]$, according to the described models. The energy consumption of one relay candidate is given by $T_S\cdot P_I+L_R/R\cdot P_R$. The number of the relay candidates around n_i, which receive the RTS, is a Poisson random variable, with the mean value $\lambda\cdot\rho\cdot\Omega$. Here Ω can be still given by (4), but is integrated on the backward plane, that is, $\int_{-\infty}^{0}\int_{-\infty}^{\infty} f(P_t, R, \sqrt{x^2+y^2})dx\, dy$, instead of the forwarding plane. Therefore, the expected iterative energy consumption under uncompleted DATA transmission is the summation: $(T_S + T_B)\cdot P_R + L_R/R\cdot[P_S + (1+\beta)P_t] + \lambda\cdot\rho\cdot\Omega\cdot(T_S\cdot P_I + L_R/R\cdot P_R)$.

If the DATA transmission is completed in the iteration, the energy consumption of the node n_i is $(T_S + T_B)\cdot P_R + (L_R + L_D)/R\cdot[P_S + (1+\beta)P_t] + (L_C + L_A)/R\cdot P_R$. The energy consumption of one relay candidate, which is not elected as the next node n_{i+1}, is still $T_S\cdot P_I+L_R/R\cdot P_R$. The number of the relay candidates is a Poisson random variable with the mean value $2\lambda\cdot\rho\cdot\Omega$, including both the forward and the backward planes. On the node n_{i+1}, the additional energy consumption is $L_D/R\cdot P_R + (L_C + L_A)/R\cdot[P_S + (1+\beta)P_t]$. Therefore, the expected iterative network energy consumption under completed DATA transmission is $(T_S + T_B + (L_C + L_D + L_A)/R)\cdot P_R + (L_R + L_C + L_D + L_A)/R\cdot[P_S + (1+\beta)P_t] + 2\lambda\cdot\rho\cdot\Omega\cdot(T_S\cdot P_I + L_R/R\cdot P_R)$.

Since the DATA transmission is always completed in the last iteration, ϵ_I is directly obtained from the previous discussion. When $i < I$, similar to the previous analysis of τ_i, the probability of uncompleted DATA transmission is decided by $e^{-\lambda\rho\Omega}$. Therefore, $\{\epsilon_i \mid 1 \le i < I\}$ can be modeled

as a set of random variables with identical distributions, where the expectation $E(\epsilon_i)$ is

$$\begin{aligned} E(\epsilon_i) = {} & (T_S + T_B)\cdot P_R + \frac{L_R}{R}\cdot[P_S + (1+\beta)P_t] \\ & + \lambda\cdot\rho\cdot\Omega\cdot\left(T_S\cdot P_I + \frac{L_R}{R}\cdot P_R\right) + (1 - e^{-\lambda\rho\Omega}) \\ & \cdot\left\{\frac{L_C + L_D + L_A}{R}\cdot[P_R + P_S + (1+\beta)P_t]\right. \\ & \left. + \lambda\cdot\rho\cdot\Omega\cdot\left(T_S\cdot P_I + \frac{L_R}{R}\cdot P_R\right)\right\}. \end{aligned} \tag{6}$$

4.3.3. Number of iterations

Given all the parameters, the total iteration number I is a random variable decided by the source-to-destination distance l. It is shown in Appendix A that the mean value $E(I)$ and the variance $\sigma^2(I)$ are

$$E(I) = \frac{l}{E(\Lambda)} + o(\sqrt{l}), \tag{7}$$

$$\sigma^2(I) = \frac{l\cdot\sigma^2(\Lambda)}{E(\Lambda)^3} + o(l), \tag{8}$$

respectively.

In the above, the random variable Λ generally indicates the forwarding distance of one single iteration. Λ is defined with the pdf (probability density function) $p_\Lambda(x)$:

$$p_\Lambda(x) = \begin{cases} \lambda\rho\cdot e^{-\lambda\rho\int_x^\infty g(u)\,du}\cdot g(x), & x > 0, \\ e^{-\lambda\rho\Omega}\cdot\delta(0), & x = 0, \\ 0, & x < 0, \end{cases} \tag{9}$$

where

$$g(u) = \int_{-\infty}^{\infty} f\left(P_t, R, \sqrt{u^2 + v^2}\right) dv. \tag{10}$$

Proved in Appendix B, we point out that the expectation $E(\Lambda)$ is an increasing function of the node density λ, but is upper bounded, as decided by the channel model $f(\cdot,\cdot,\cdot)$. Furthermore, the variance $\sigma^2(\Lambda)$ decreases to zero with higher λ. Given (7) and (8), both $E(I)$ and $\sigma^2(I)$ can decrease with higher node density λ. These relations can also be intuitively understood, which will be exploited for interpreting the end-to-end performance analysis results next.

4.4. End-to-end performance analysis

Analytical results are presented for the maximal throughput Φ, the expected end-to-end delay Υ, the end-to-end delay variance Θ, and the expected network energy consumption per DATA packet Ξ, of a dataflow from the source s to the destination d, under the long-range asymptotic condition $l \to \infty$. The discussion shows how these metrics change with the source-to-destination distance l, as well as the network density λ.

4.4.1. Throughput

The maximal throughput Φ is decided by how many DATA packets can be sent out from the source s to the destination d in a period of time, subject to the condition that this time period is much larger than the packet traveling time. According to previous analysis in Section 4.3.1, the probability of the completed DATA transmission in an iteration is $1-e^{-\lambda\rho\Omega}$. Therefore, the expected number of iterations initiated by s for one DATA packet is $1/(1-e^{-\lambda\rho\Omega})$, and the associated time expectation is $(L_R/R+T_S+T_B)/(1-e^{-\lambda\rho\Omega})+(L_C+L_D+L_A)/R$.

Given the DATA packet length L_D, the maximal throughput is thus obtained as

$$\Phi = \frac{L_D\cdot R}{(L_R + (T_S + T_B)\cdot R)/(1 - e^{-\lambda\rho\Omega}) + L_C + L_D + L_A}. \tag{11}$$

It is shown that the maximal throughput Φ can be independent of the source-to-destination distance l. It also increases monotonically with the node density λ, approaching the limit decided by the radio data rate R and the overhead ratio.

4.4.2. Expected end-to-end delay

The end-to-end delay is given by $\sum_{i=1}^{I}\tau_i$. The expected end-to-end delay Υ is

$$\begin{aligned} \Upsilon &=^{(a)} E\left(\sum_{i=1}^{I}\tau_i\right) \\ &=^{(b)} [E(I) - 1]\cdot E(\tau_1) + \tau_I \\ &=^{(c)} l\cdot\frac{L_R + (T_S + T_B)\cdot R + (L_C + L_D + L_A)\cdot(1 - e^{-\lambda\rho\Omega})}{R\cdot E(\Lambda)} \\ &\quad + o(\sqrt{l}), \end{aligned} \tag{12}$$

where (a) is given by the definition; (b) is due to the fact that $\{\tau_i \mid 1 \le i < I\}$ are of identical probability distribution [32]; and (c) is obtained directly from (5) and (7).

Therefore, the expected end-to-end delay Υ increases linearly with the source-to-destination distance l. It also decreases with the node density λ, approaching a bounded limit.

4.4.3. Delay variance

Let I_c denote the number of the iterations with completed DATA transmission, that is, the number of wireless hops, and let I_u denote the number of the iterations with uncompleted DATA transmission. I_c and I_u are independent random

variables, and $I = I_c + I_u$. The delay variance Θ is calculated by the following:

$$\begin{aligned}
\Theta &=^{(a)} \sigma^2\left(\sum_{i=1}^{I}\tau_i\right)\\
&=^{(b)} \sigma^2(I_c)\cdot\left(\frac{L_R+L_C+L_D+L_A}{R}+T_S+T_B\right)^2\\
&\quad+\sigma^2(I_u)\cdot\left(\frac{L_R}{R}+T_S+T_B\right)^2\\
&=^{(c)} \sigma^2(I_c)\cdot\left[\left(\frac{L_C+L_D+L_A}{R}\right)^2+2\cdot\left(T_S+T_B+\frac{L_R}{R}\right)\right.\\
&\qquad\left.\cdot\frac{L_C+L_D+L_A}{R}\right]+\sigma^2(I)\cdot\left(\frac{L_R}{R}+T_S+T_B\right)^2\\
&=^{(d)} l\cdot\frac{\sigma^2(\Lambda)}{E(\Lambda)^3}\cdot\left\{\left(1-e^{-\lambda\rho\Omega}\right)^2\right.\\
&\qquad\cdot\left[\left(\frac{L_C+L_D+L_A}{R}\right)^2+2\cdot\left(T_S+T_B+\frac{L_R}{R}\right)\right.\\
&\qquad\left.\left.\cdot\frac{L_C+L_D+L_A}{R}\right]+\left(\frac{L_R}{R}+T_S+T_B\right)^2\right\}\\
&\quad-l\cdot e^{-\lambda\rho\Omega}\cdot\frac{1-e^{-\lambda\rho\Omega}}{E(\Lambda)}\cdot\left[\left(\frac{L_C+L_D+L_A}{R}\right)^2+2\cdot\left(T_S+T_B+\frac{L_R}{R}\right)\right.\\
&\qquad\left.\cdot\frac{L_C+L_D+L_A}{R}\right]+o(l),
\end{aligned} \tag{13}$$

where (a) is given by the definition; (b) is due to the analysis in Section 4.3.1, that is, the iterative delay under completed and uncompleted transmissions respectively; (c) is due to the fact that $\sigma^2(I) = \sigma^2(I_c) + \sigma^2(I_u)$; (d) is obtained from the result of $\sigma^2(I)$ in (8), as well as the $\sigma^2(I_c)$ derived in (A.4) of Appendix A.

Therefore, the end-to-end delay variance Θ increases linearly with the source-to-destination distance l. Θ also diminishes to zero with the network density λ, since the component $\sigma^2(\Lambda)$ also decreases to zero with higher λ.

4.4.4. Network energy consumption

The expected network energy consumption for one DATA packet Ξ is

$$\begin{aligned}
\Xi &=^{(a)} E\left(\sum_{i=1}^{I}\epsilon_i\right)\\
&=^{(b)} [E(I)-1]\cdot E(\epsilon_1)+\epsilon_I\\
&=^{(c)} l\cdot\frac{E(\epsilon_1)}{E(\Lambda)}+o(\sqrt{l}),
\end{aligned} \tag{14}$$

where (a) and (b) are given by the definition and the fact that $\{\epsilon_i \mid 1 \le i < I\}$ are of identical probability distributions [32]; (c) is obtained directly from (7), and the component ϵ_1 can be replaced by the result in (6). Therefore, Ξ also increases linearly with the source-to-destination distance l.

4.5. Optimizations under QoS constraints

Consider that the dataflow has certain QoS requirements on throughput, expected end-to-end delay, and delay variance, for example, $\Phi \ge \Phi_{\text{th}}$, $\Upsilon \le \Upsilon_{\text{th}}$, and $\Theta \le \Theta_{\text{th}}$. Here, Φ_{th}, Υ_{th}, and Θ_{th} are the threshold constants representing those QoS constraints, respectively. The optimizations aim to minimize the resource consumption, subject to these QoS constraints.

Two types of resource consumptions, including network energy consumption and network capacity, are considered here, which lead to different optimization formulations. As pointed out in Section 4.2.4, these two formulations can be representing different networking scenarios. In capacity limited networks, every node can have cabled power supply. The number of nodes is therefore limited by engineering costs, and the possible use of long-range radio also limits the available bandwidth. On the other hand, in energy limited networks, the nodes can be powered by energy scavenging, for example, solar cells. Therefore, the number of wireless nodes can be large, which can also use short-range radio with vast unlicensed bandwidth. These contribute to large network capacity. Other works with similar considerations have been discussed in [33].

4.5.1. Capacity limited networks

The objective of the optimization is to minimize the transmitting power P_t, so as to limit the occupied geographic area of an unicast link. In accordance with the results in [28], this approach can save the use of wireless network capacity. Therefore, the optimization formulation is

$$\begin{aligned}
&\min P_t\\
&\text{subject to: } \Phi \ge \Phi_{\text{th}};\quad \Upsilon \le \Upsilon_{\text{th}};\quad \Theta \le \Theta_{\text{th}}.
\end{aligned} \tag{15}$$

4.5.2. Energy limited networks

The objective of the optimization is to minimize the expected network energy consumption per DATA packet Ξ:

$$\begin{aligned}
&\min \Xi\\
&\text{subject to: } \Phi \ge \Phi_{\text{th}};\quad \Upsilon \le \Upsilon_{\text{th}};\quad \Theta \le \Theta_{\text{th}}.
\end{aligned} \tag{16}$$

4.5.3. Solutions and traffic priority

In order to solve the optimizations in (15) and (16), the transmitting power P_t can be an ideal control knob (optimization variable), for the resource allocation. Both (15) and (16) can be found as convex optimization problems over P_t, where efficient solutions are guaranteed [34].

Specifically, as shown in (11), Φ monotonically increases with P_t, since Ω (defined in (4)) monotonically increases with P_t, provided by any realistic channel model $f(\cdot,\cdot,\cdot)$. As shown in (12), Υ monotonically decreases with P_t, since $E(\Lambda)$ monotonically increases with P_t (proved in Appendix B).

Θ also monotonically decreases with P_t in (13), since $\sigma^2(\Lambda)$ monotonically decreases with P_t (proved in Appendix B). Moreover, as shown in (14), Ξ can be verified as a convex function of P_t, where both $E(\epsilon_1)$ and $E(\Lambda)$ monotonically increase with P_t. Due to the described monotonicity, the convexity of the optimizations in (15) and (16) can then be obtained, over the control knob P_t.

Therefore, we suggest that the transmitting power P_t can provide an effective control knob for the tradeoff between the resource consumptions and the real-time QoS requirements. In fact, different levels of P_t can be configured for different classes of traffics, for example, voice, video (streaming), or data, which then decide the unicasting "priority".

Further investigation may also include the radio data rate R, so as to generate the problem of joint power and rate control. Unfortunately, the convexity over R cannot be obtained for a general channel propagation model $f(\cdot,\cdot,\cdot)$. Future studies may address the problem on some specific models.

4.6. Summary

In summary, the analytical results indicate that all the investigated real-time QoS metrics can improve with larger network scale. In particular, the maximal throughput Φ increases with the network density λ to a predetermined bound, while the expected end-to-end delay Υ decreases with λ, approaching a predetermined bound, as decided by other related parameters. On the other hand, the delay variance Θ falls to zero with higher network density λ, which indicates that the network performance can be made arbitrarily stable, simply by dropping more nodes in the network. For supporting long-range communications over a relatively large number of wireless hops, the analytical results show that the maximal throughput Φ can be independent of the source-to-destination distance, while both the expected delay Υ and the delay variance Θ increase linearly with the source-to-destination distance. Furthermore, it is identified that the radio transmitting power P_t can be an ideal control knob for the tradeoff between the real-time QoS requirements and the resource consumptions, for example, energy consumption or network capacity.

We consider that all these propositions could be intuitively conceivable for large-scale wireless networks. Since the analytical results are obtained under the long-range asymptotic condition $l \to \infty$, they may present the worst-case scenario analysis for practical considerations. In the next section, the simulation results are also provided, showing how these metrics converge.

5. SIMULATION RESULTS

Some fixed parameters in the simulations are listed in Table 2. In particular, the radio transmission parameters are conforming to the developed test-bed, while others, for example, power consumption and channel model parameters, are typical for indoor short-range radios. For example, under these parameters, the radio range is about 10–15 meters, when the transmitting power P_t is 1 mW.

Table 2: Parameters in the simulations.

Parameter	Value	Unit
ρ	0.2	N/A
R	250 K	bits/sec
L_R	120	bit
L_C, L_A	96	bit
L_D	1024	bit
T_S	250	μsec
T_B	750	μsec
P_S, P_R	1	mW
P_I	100	μW
β	3	N/A
G	−40	dB
N_0	−150	dBm/Hz
B	10	dB
α	4	N/A

Given the source and the destination nodes, separated by the distance l, random Poisson points are generated, for every iteration. The generation of the Poisson points conforms to the density $\lambda\cdot\rho$ on the 2D plane, which represent the available wireless nodes. 1000 DATA packets are then "sent out" consecutively from the source to the destination, according to the described real-time unicast module. The end-to-end delay instance of every DATA packet is then recorded, from which the simulated end-to-end delay expectation and the simulated delay variance are estimated by standard statistical methods. The simulated throughput is also obtained, by dividing the data bulk size $1000\cdot L_D$ with the time consumption of sending out the 1000 DATA packets at the source. In the following figures, we also compare the simulation results to the analytical results side-by-side, where the lower-order terms, for example, $o(\sqrt{l})$ or $o(l)$, are neglected in plotting the analytical results.

Figures 4, 5, and 6 show how the QoS metrics change with the source-to-destination distance l, where the node density λ is fixed at $0.1/\text{m}^2$. Then Figures 7, 8, and 9 show how the QoS metrics change with the network density λ, where the source-to-destination l is fixed at 160 m. The number of wireless hops in the simulations can be up to 20. All the obtained simulation results match well with the analytical results, as well as the main propositions given in Section 4.6. Especially, Figures 4, 5, and 6 show that close-matching also exists in short-range unicasting, although the analytical results were obtained under long-range asymptotic approximation.

Moreover, Figure 10 shows how the expected network energy consumption per DATA packet changes with the transmitting power P_t, which is observed as a convex function. It is also observed in Figure 10 that the expected network energy consumption increases linearly with l, and it decreases with λ in the simulation. These simulation results conform to previous performance analysis.

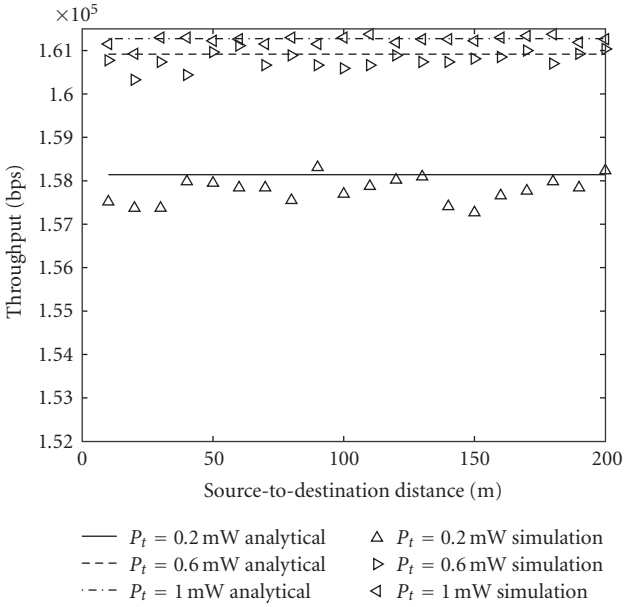


Figure 4: Throughput versus source-to-destination distance l ($\lambda = 0.1/\text{m}^2$).

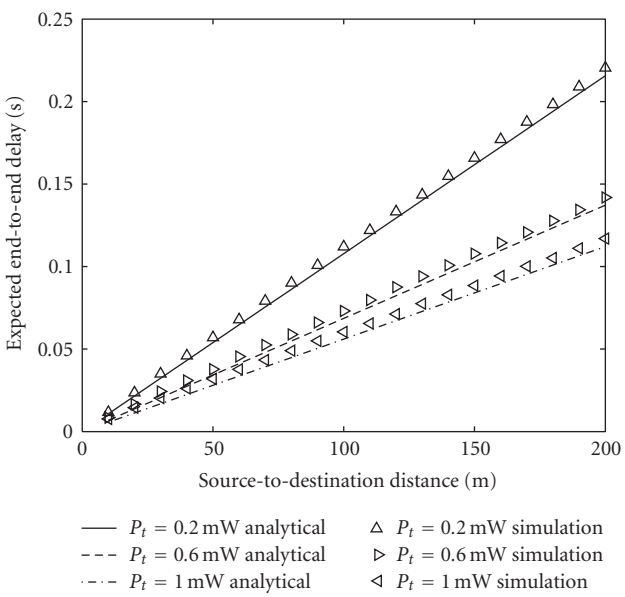


Figure 5: Expected delay versus source-to-destination distance l ($\lambda = 0.1/\text{m}^2$).

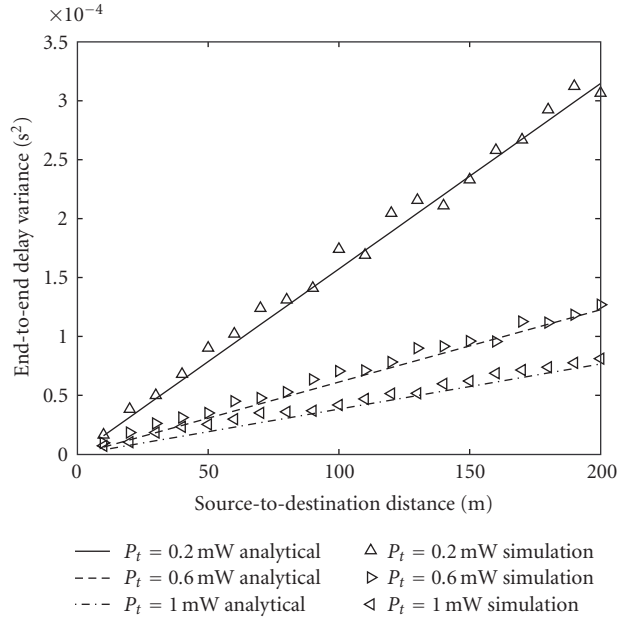


Figure 6: Delay variance versus source-to-destination distance l ($\lambda = 0.1/\text{m}^2$).

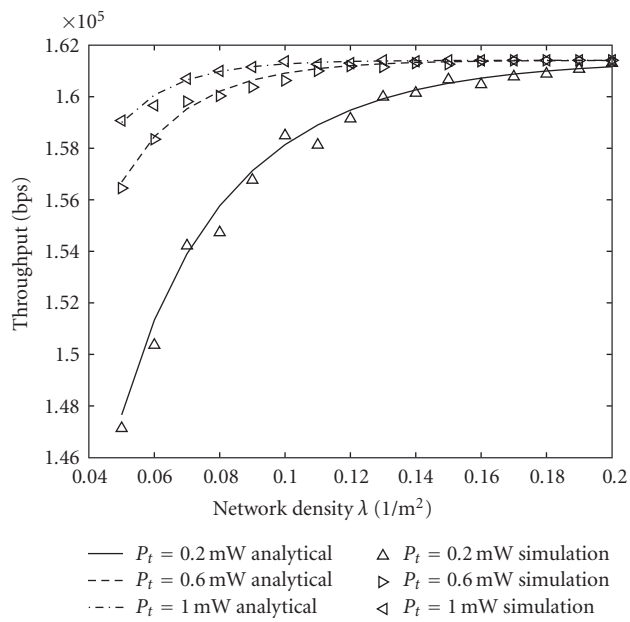


Figure 7: Throughput versus network density λ ($l = 160$ m).

6. EXPERIMENT RESULTS

In the test-bed (Figure 11), the source and the destination nodes are connected to a laptop computer for the purpose of collecting experiment records. The wireless nodes are based on the cognitive radio prototype (Figure 1), with a stack of two radio boards. The cognitive radio can opportunistically access 3 data channels in 2.4 GHz band, as described in Section 3.2, where each channel has an air-interface data rate 250 kbps as defined in IEEE 802.15.4 [26].

The real-time unicast module is implemented in the cognitive radio prototype, with preconfigured address. The packet length parameters, that is, L_R, L_C, L_D, L_A, are about the same as those listed in Table 2. The node transmitting power is however fixed at -15 dBm. The experiment results on throughput, end-to-end delay expectation, and delay variance, are obtained by the following methods. (1) Throughput is calculated at the source, where the laptop computer forwards 1000 DATA packets to the source node.

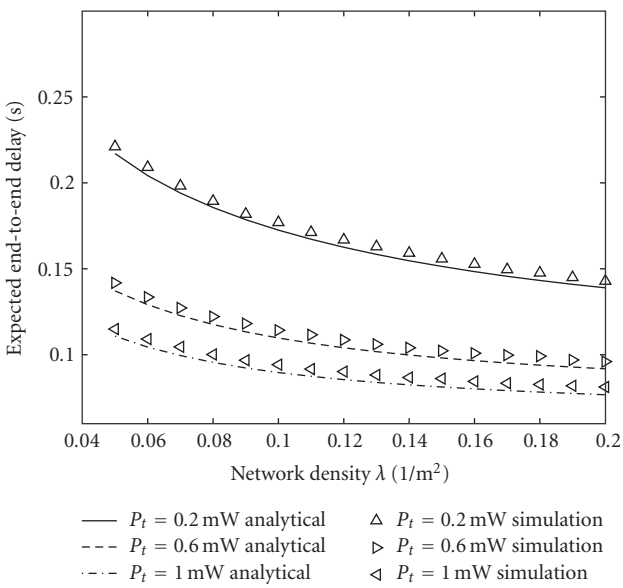


Figure 8: Expected delay versus network density λ ($l = 160$ m).

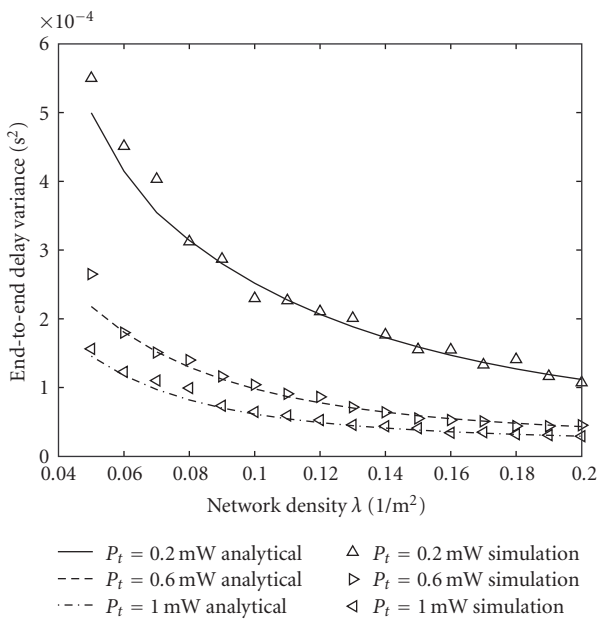


Figure 9: Delay variance versus network density λ ($l = 160$ m).

Since the source node accepts a new DATA packet from the computer only after the previous one has been sent out, the computer can count the total time expenditure in sending the 1000 DATA packets. The measured throughput is obtained by dividing the bulk size with the time expenditure. (2) On receiving a DATA packet, the destination also forwards the DATA packet to the laptop computer, where the end-to-end delay instance of every DATA packet is calculated at the computer by the difference between the packet departure and arrival time. The measured end-to-end delay expectation and

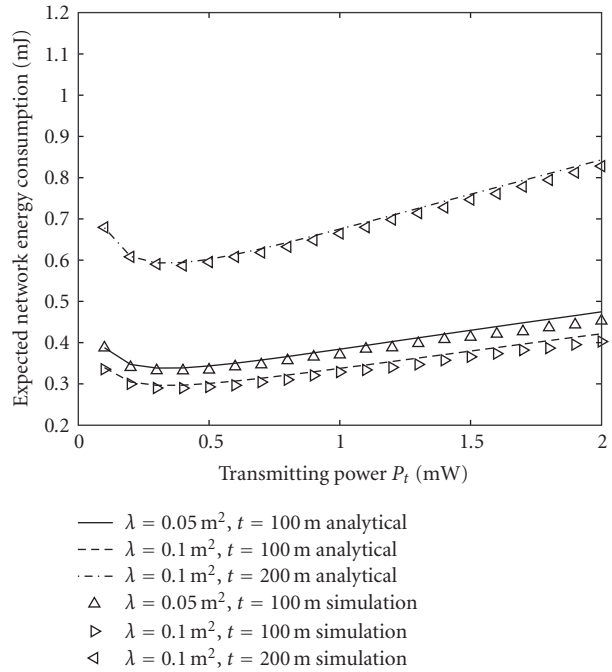


Figure 10: Network energy consumption versus transmitting power P_t.

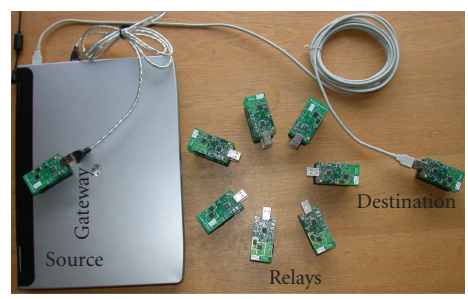


Figure 11: A snapshot of the test-bed.

the delay variance are then obtained by standard statistical methods. Since the operating system running at the laptop computer (Windows XP Professional Edition) is by no means real-time, we tried to minimize the computer running threads in the experiments.

The experiments were conducted in a standard office environment, where external interferences in the utilized 2.4 GHz band can be intensive, for example, from adjacent WLAN hot spots. Two sets of experiments are illustrated in Figure 12. In Set-I, the nodes are placed in a straight line, such that the four configurations correspond approximately to one-hop to four-hop communications. For example, in the second configuration of Set-I, the node placement is such that the chance of direct transmission from s to d is small. In Set-II, more relay nodes are added to every hop being illustrated in Figure 12. The purpose of the two-set experiments is to compare how the designed unicast module can improve the test-bed QoS performance. In fact,

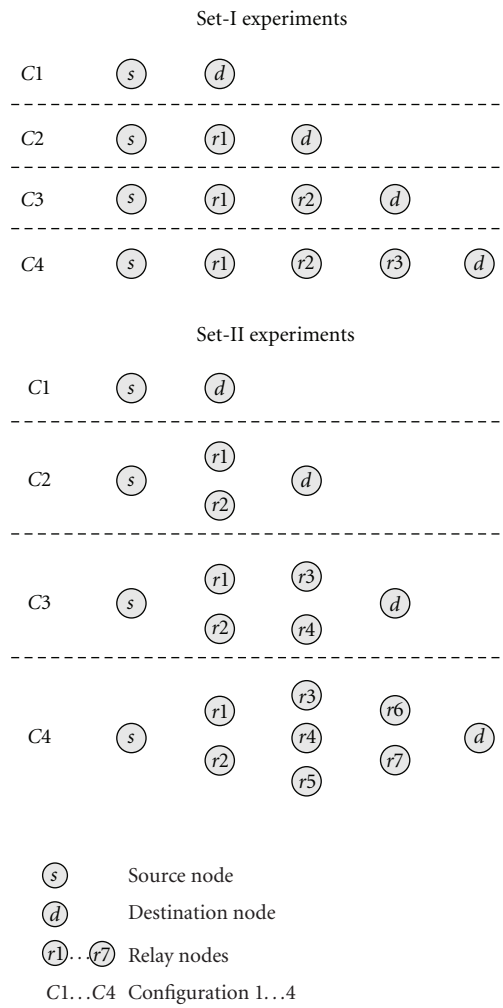


Figure 12: Illustration of the two sets of experiments.

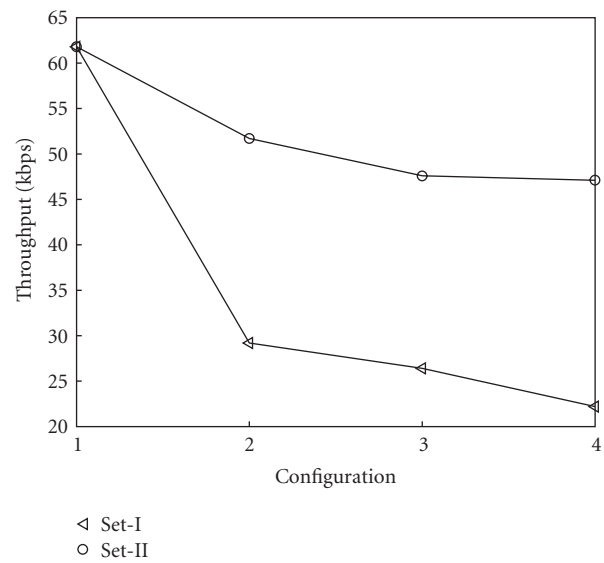


Figure 13: Measurement results on the throughput.

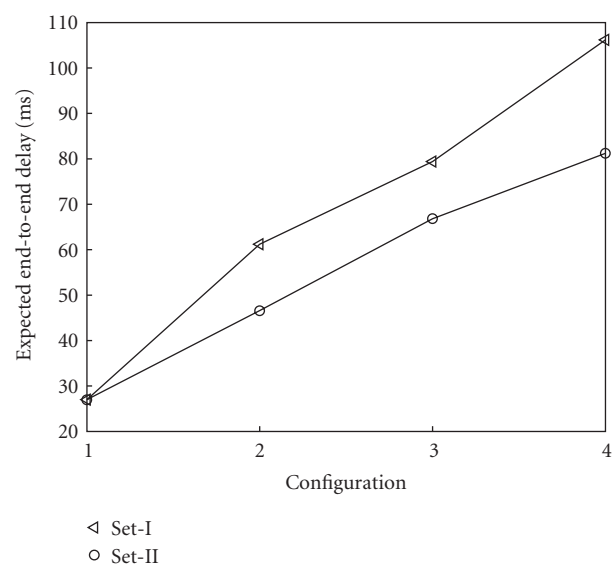


Figure 14: Measurement results on the expected end-to-end packet delay.

Set-I experiments can be comparable to traditional wireless networking where a routing path is predetermined with best available spectrum in every hop. However, in Set-II experiments, data is forwarded from s to d opportunistically without any predetermined path.

Figures 13, 14, and 15 show the measurement results of the two-set experiments on throughput, expected end-to-end delay, and delay variance, respectively. Evidently, the results of Set-II are strictly better than those of Set-I, beyond the first configuration. In Figure 13, the throughput of Set-I is reduced by more than 50% from the first to the second configuration, that is, from one-hop to two-hop, due to the impact of half-duplex radios. A comparable reduction of throughput is not observed in Set-II. In general, the throughput of Set-II changes smoothly with different configurations (number of wireless hops), where the limited reduction can be due to the limited number of nodes. In Figures 14 and 15, it is further observed that the expected end-to-end delay and the delay variance increase linearly with the number of hops (configuration index) in Set-II experiments; whereas they tend to increase much faster in Set-I experiments.

7. CONCLUSIONS

In the paper, we have implemented a real-time unicast module in the context of large-scale cognitive wireless networking. By opportunistically using both spectrum bandwidth and wireless nodes/radio availability, the most significant result is that the network performance can improve with larger network scale. Specifically, it is shown that the performance dynamics, that is, indicated by dataflow end-to-end delay variance, can fall to zero with higher network density. Furthermore, additional results show that dataflow throughput can be independent of the number of wireless hops or the source-to-destination distance, whereas end-to-end delay expectation and delay variance increase

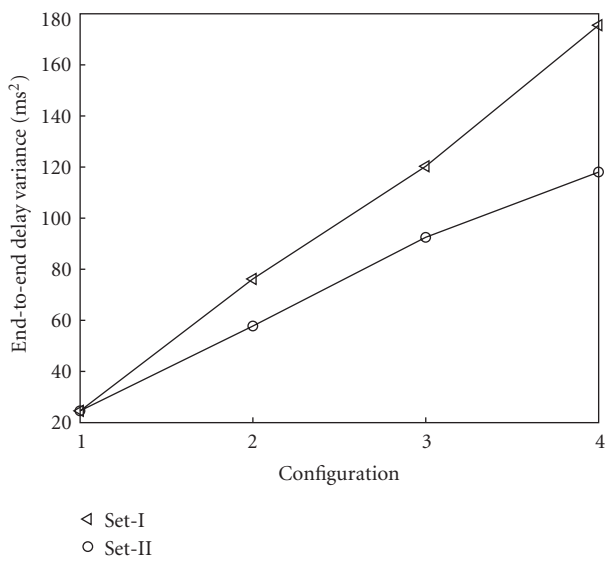


Figure 15: Measurement results on the packet delay variance.

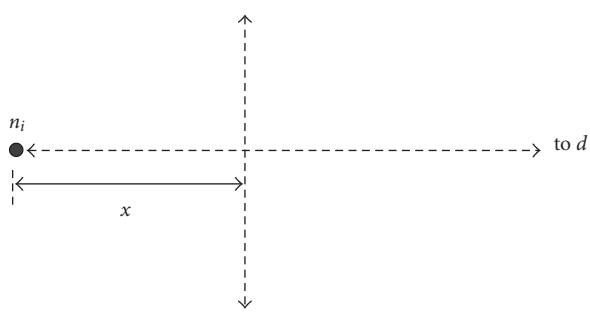


Figure 16: Calculation of $p_\Lambda(x)$.

linearly. The obtained analytical, simulation, and experiment results support these propositions. It is also identified that node transmitting power can be an effective control knob, for the tradeoff between the real-time QoS performance metrics and resource consumptions. We suggest that the presented research could contribute to the development of diverse commercial and scientific applications with large-scale wireless networking.

APPENDICES

A. CALCULATIONS OF THE ITERATION NUMBER I

At one iteration i, let X_i denote the distance advance to the destination d, that is,

$$X_i = \begin{cases} \left|\mathbf{L}_{n_i} - \mathbf{L}_d\right| - \left|\mathbf{L}_{n_{i+1}} - \mathbf{L}_d\right|, & 1 \le i < I, \\ \left|\mathbf{L}_{n_I} - \mathbf{L}_d\right|, & i = I. \end{cases} \tag{A.1}$$

Furthermore, let $p_{X_i}(x)$ denote the pdf of the positive random variable X_i. Given $1 \le i < I$, $p_{X_i}(x)$ is decided by the distance between the node n_i and the destination d, $|\mathbf{L}_{n_i} - \mathbf{L}_d|$. Under the condition that $|\mathbf{L}_{n_i} - \mathbf{L}_d| \to \infty$, the probability of $X_i < x$, that is, the cumulative distribution function (cdf), is given by the probability that there is no available relay candidates to the right of the vertical line shown in Figure 16. Particularly, the relay candidates distribution is an inhomogeneous Poisson process with the density function $\lambda \cdot \rho \cdot f(P_t, R, \zeta)$, where ζ is the distance to n_i. Therefore, for $x \ge 0$, the cdf of X_i is obtained:

$$\text{Prob}(X_i \le x) = e^{-\lambda\rho \cdot \int_x^\infty \int_{-\infty}^\infty f(P_t, R, \sqrt{u^2+v^2})\, du\, dv}. \tag{A.2}$$

The pdf of X_i, $p_{X_i}(x)$, can be calculated by taking the differentiation of the cdf function in (A.2), which equals to the $p_\Lambda(x)$ defined in (9).

Therefore, $p_{X_i}(x) \to p_\Lambda(x)$, that is, X_i converges to Λ in distribution, when the distance $|\mathbf{L}_{n_i} - \mathbf{L}_d| \to \infty$. In obtaining the statistics of the iteration number I, we divide the source-to-destination distance l into two segments, with the lengths $l - \sqrt[3]{l}$ and $\sqrt[3]{l}$, respectively. Let I_1 denote the first passage iteration of first segment, which is $I_1 = \min\{i \mid |\mathbf{L}_{n_i} - \mathbf{L}_d| < \sqrt[3]{l}\}$, and let I_2 denote the rest number of iterations in the second segment. Given l, I_1 and I_2 are independent variables, and $I = I_1 + I_2$. Therefore, we obtain $E(I) = E(I_1) + E(I_2)$, and $\sigma^2(I) = \sigma^2(I_1) + \sigma^2(I_2)$.

According to the above definitions, under the asymptotic condition $l \to \infty$, the set of random variables $\{X_i \mid 1 \le i \le I_1\}$ converge to the random variable Λ in distribution. Therefore, I_1 is virtually the first passage time of a positive-drift stopped random walk beyond the level $l - \sqrt[3]{l}$, where the random walk steps are i.i.d. (The proposition is mathematically rigorous under the uniform integrability [35] of $\{|X_i|^r \mid 1 \le i \le I_1, r \ge 1\}$, which can be verified for realistic wireless propagation models $f(\cdot,\cdot,\cdot)$.) according to the distribution $p_\Lambda(x)$. Due to the results of stopped random walks [36, Theorem 9.1, page 94], there are

$$E(I_1) = \frac{l}{E(\Lambda)} + o(\sqrt{l}), \qquad \sigma^2(I_1) = \frac{l \cdot \sigma^2(\Lambda)}{E(\Lambda)^3} + o(l) \tag{A.3}$$

Now, if we assume that $E(I) = C_1 \cdot l^{c_1} + o(l^{c_1})$, and $\sigma^2(I) = C_2 \cdot l^{c_2} + o(l^{c_2})$, where C_1, C_2, c_1, c_2 are constants, there will be $E(I_2) = C_1 \cdot l^{c_1/3} + o(l^{c_1/3})$, and $\sigma^2(I_2) = C_2 \cdot l^{c_2/3} + o(l^{c_2/3})$, since the difference between I and I_2 is only by the distances l and $\sqrt[3]{l}$ as defined. Due to the statistics of I_1 in (A.3), it is easy to obtain $C_1 = 1/E(\Lambda)$, $C_2 = \sigma^2(\Lambda)/E(\Lambda)^3$, and $c_1 = c_2 = 1$. Therefore, the results on the iteration number I, as given in (7) and (8), are obtained directly.

We further consider I_c, which denotes the number of iterations with completed DATA transmissions, that is, the number of wireless hops. Corresponding to the random variable Λ in the previous analysis, the asymptotic step length Λ_c under the prior condition of completed DATA transmission is decided by the pdf, $p_{\Lambda_c}(x) = p_\Lambda(x)/(1 - e^{-\lambda\rho\Omega})$, since the probability of an iteration with uncompleted DATA transmission is $e^{-\lambda\rho\Omega}$. Therefore, simply by replacing Λ with Λ_c in the previous analysis, the statistics of I_c can be obtained.

In particular, $E(I_c) = l/E(\Lambda_c) + o(\sqrt{l}) = l\cdot(1-e^{-\lambda\rho\Omega})/E(\Lambda) + o(\sqrt{l})$, and

$$\begin{aligned}\sigma^2(I_c) &= l\cdot\frac{\sigma^2(\Lambda_c)}{E(\Lambda_c)^3} + o(l)\\ &= l\cdot\left(\frac{E(\Lambda_c^2)}{E(\Lambda_c)^3} - \frac{1}{E(\Lambda_c)}\right) + o(l)\\ &= l\cdot\left(\frac{E(\Lambda^2)}{E(\Lambda)^3}\cdot(1-e^{\lambda\rho\Omega})^2 - \frac{1-e^{-\lambda\rho\Omega}}{E(\Lambda)}\right) + o(l)\\ &= l\cdot\left[\frac{\sigma^2(\Lambda)}{E(\Lambda)^3}\cdot(1-e^{-\lambda\rho\Omega})^2 - \frac{e^{-\lambda\rho\Omega}(1-e^{-\lambda\rho\Omega})}{E(\Lambda)}\right] + o(l).\end{aligned} \tag{A.4}$$

B. EXPRESSIONS OF $E(\Lambda)$ AND $\sigma^2(\Lambda)$

Due to the pdf $p_\Lambda(x)$ in (9), we obtain

$$\begin{aligned}E(\Lambda) &= \lim_{C\to\infty}\int_0^C x\cdot de^{-\lambda\rho\int_x^\infty g(u)du}\\ &= \lim_{C\to\infty} x\cdot e^{-\lambda\rho\int_x^\infty g(u)du}\Big|_{x=0}^{C} - \int_0^C e^{-\lambda\rho\int_x^\infty g(u)du}dx\\ &= \lim_{C\to\infty} C - \int_0^C e^{-\lambda\rho\int_x^\infty g(u)du}dx,\end{aligned} \tag{B.1}$$

and similarly

$$\begin{aligned}E(\Lambda^2) &= \lim_{C\to\infty}\int_0^C x^2\cdot de^{-\lambda\rho\int_x^\infty g(u)du}\\ &= \lim_{C\to\infty} C^2 - 2\int_0^C x\cdot e^{-\lambda\rho\int_x^\infty g(u)du}dx.\end{aligned} \tag{B.2}$$

Therefore,

$$\begin{aligned}\sigma^2(\Lambda) &= E(\Lambda^2) - E(\Lambda)^2\\ &= \lim_{C\to\infty} 2C\cdot\int_0^C e^{-\lambda\rho\int_x^\infty g(u)du}dx\\ &\quad - \left(\int_0^C e^{-\lambda\rho\int_x^\infty g(u)du}dx\right)^2\\ &\quad - 2\int_0^C x\cdot e^{-\lambda\rho\int_x^\infty g(u)du}dx.\end{aligned} \tag{B.3}$$

From (B.1) and (B.3), it is shown that $E(\Lambda)$ increases with λ, approaching a limit decided by the characteristics of $g(u)$, and $\sigma^2(\Lambda)$ decreases with λ, approaching zero. Since $g(u)$, as defined in (10), is a monotonically increasing function of P_t, it is also shown that $E(\Lambda)$ increases monotonically with P_t, and $\sigma^2(\Lambda)$ decreases monotonically with P_t.

C. MODELS OF THE ITERATIVE DELAY AND ENERGY CONSUMPTION

We point out that more accurate models of the iterative delay τ_i $(1 \le i < I)$ and the iterative energy consumption ϵ_i $(1 \le i < I)$ can include the distance $|\mathbf{L}_{n_i} - \mathbf{L}_d|$. Under such models, τ_i and ϵ_i will converge to the results in Sections 4.3.1 and 4.3.2, respectively, when $|\mathbf{L}_{n_i} - \mathbf{L}_d| \to \infty$. With similar discussion in Appendix A, this will contribute to the same end-to-end analytical results, under the asymptotic condition $l \to \infty$.

ACKNOWLEDGMENTS

The authors would like to acknowledge the support from Natural Sciences and Engineering Research Council of Canada (NSERC) and Ontario Research Commercialization Program (ORCP).

REFERENCES

[1] L. Song, "Wireless mesh infrastructure supporting broadband internet with multimedia services," in *Proceedings of the 4th IEEE International Conference on Circuits and Systems for Communications (ICCSC '08)*, pp. 792–796, Shanghai, China, May 2008.

[2] D. Kundur, C.-Y. Lin, and C.-S. Lu, "Visual sensor networks," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, Article ID 21515, 3 pages, 2007.

[3] ANSI/IEEE Standard for Local and Metropolitan Area Networks Part 11 (802.11), "Wireless LAN medium access control (MAC) and physical layer (PHY) specifications," revised 2003.

[4] R. Draves, J. Padhye, and B. Zill, "Routing in multi-radio, multi-hop wireless mesh networks," in *Proceedings of the 10th Annual International Conference on Mobile Computing and Networking (MOBICOM '04)*, pp. 114–128, Philadelphia, Pa, USA, September 2004.

[5] Y. Zhu, H. Liu, M. Wu, D. Li, and S. Mathur, "Implementation experience of a prototype for video streaming over wireless mesh networks," in *Proceedings of the 4th Annual IEEE Consumer Communications and Networking Conference (CCNC '07)*, pp. 23–28, Las Vegas, Nev, USA, January 2007.

[6] A. Kumar, D. Manjunath, and J. Kuri, *Communication Networking: An Analytical Approach*, Morgan Kaufmann, San Francisco, Calif, USA, 2004.

[7] P. Djukic and S. Valaee, "Link scheduling for minimum delay in spatial re-use TDMA," in *Proceedings of the 26th IEEE International Conference on Computer Communications (INFOCOM '07)*, pp. 28–36, Anchorage, Alaska, USA, May 2007.

[8] L. Song, "Cognitive networks: standardizing the large scale wireless systems," in *Proceedings of the 5th IEEE Consumer Communications and Networking Conference (CCNC '08)*, pp. 988–992, Las Vegas, Nev, USA, January 2008.

[9] L. Song and D. Hatzinakos, "Embedded wireless interconnect for sensor networks: concept and example," in *Proceedings of the 4th Annual IEEE Consumer Communications and Networking Conference (CCNC '07)*, pp. 850–854, Las Vegas, Nev, USA, January 2007.

[10] L. Song and D. Hatzinakos, "A cross-layer architecture of wireless sensor networks for target tracking," *IEEE/ACM Transactions on Networking*, vol. 15, no. 1, pp. 145–158, 2007.

[11] L. Song and D. Hatzinakos, "Architecture of wireless sensor networks with mobile sinks: sparsely deployed sensors," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 4, pp. 1826–1836, 2007.

[12] L. Song, D. Hatzinakos, and X. Wang, "Wireless mesh infrastructure for ubiquitous voice and video," in *Proceedings of the 5th IEEE Consumer Communications and Networking Conference (CCNC '08)*, pp. 1267–1268, Las Vegas, Nev, USA, January 2008.

[13] S. Biswas and R. Morris, "Opportunistic routing in multi hop wireless networks," in *Proceedings of the 2nd ACM Workshop on Hot Topics in Networks (HotNets '03)*, Cambridge, Mass, USA, November 2003.

[14] M. Zorzi and R. R. Rao, "Geographic random forwarding (GeRaF) for ad hoc and sensor networks: energy and latency performance," *IEEE Transactions on Mobile Computing*, vol. 2, no. 4, pp. 349–365, 2003.

[15] M. Zorzi and R. R. Rao, "Geographic random forwarding (GeRaF) for ad hoc and sensor networks: multihop performance," *IEEE Transactions on Mobile Computing*, vol. 2, no. 4, pp. 337–348, 2003.

[16] R. C. Shah, S. Wietholter, A. Wölisz, and J. M. Rabaey, "Modeling and analysis of opportunistic routing in low traffic scenarios," in *Proceedings of the 3rd International Symposium on Modeling and Optimization in Mobile, Ad hoc, and Wireless Networks (WIOPT '05)*, pp. 294–304, Trentino, Italy, April 2005.

[17] A. Bletsas, A. Khisti, D. P. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 3, pp. 659–672, 2006.

[18] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, and F. Silva, "Directed diffusion for wireless sensor networking," *IEEE/ACM Transactions on Networking*, vol. 11, no. 1, pp. 2–16, 2003.

[19] Q. Cao and T. Abdelzaher, "Scalable logical coordinates framework for routing in wireless sensor networks," *ACM Transactions on Sensor Networks*, vol. 2, no. 4, pp. 557–593, 2006.

[20] J. Mitola III and G. Q. Maguire Jr., "Cognitive radio: making software radios more personal," *IEEE Personal Communications*, vol. 6, no. 4, pp. 13–18, 1999.

[21] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201–220, 2005.

[22] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: a survey," *Computer Networks*, vol. 50, no. 13, pp. 2127–2159, 2006.

[23] Z. J. Haas and J. Deng, "Dual busy tone multiple access (DBTMA)-a multiple access control scheme for ad hoc networks," *IEEE Transactions on Communications*, vol. 50, no. 6, pp. 975–985, 2002.

[24] F. A. Tobagi and L. Kleinrock, "Packet switching in radio channels: part II. The hidden terminal problem in carrier sense multiple-access and the busy-tone solution," *IEEE Transactions on Communications*, vol. 23, no. 12, pp. 1417–1433, 1975.

[25] J. Polastre, R. Szewczyk, and D. Culler, "Telos: enabling ultra-low power wireless research," in *Proceedings of the 4th International Symposium on Information Processing in Sensor Networks (IPSN '05)*, pp. 364–369, Los Angeles, Calif, USA, April 2005.

[26] IEEE Standard for Local and Metropolitan Area Networks Part 15.4 (802.15.4), "Wireless medium access control (MAC) and physical (PHY) layer specifications for low-rate wireless personal area networks (LR-WPANs)," 2003.

[27] R. Abdelmoumen, M. Malli, and C. Barakat, "Analysis of TCP latency over wireless links supporting FEC/ARQ-SR for error recovery," in *Proceedings of IEEE International Conference on Communications (ICC '04)*, vol. 7, pp. 3994–3998, Paris, France, June 2004.

[28] P. Gupta and P. R. Kumar, "Towards an information theory of large networks: an achievable rate region," *IEEE Transactions on Information Theory*, vol. 49, no. 8, pp. 1877–1894, 2003.

[29] J. F. C. Kingman, *Poisson Processes*, Clarendon Press, Oxford, UK, 1993.

[30] T. S. Rappaport, *Wireless Communications: Principles and Practice*, Prentice Hall, Upper Saddle River, NJ, USA, 1996.

[31] B. Otis, Y. H. Chee, and J. Rabaey, "A 400 μW Rx, 1.6 mW Tx super-regenerative transceiver for wireless sensor networks," in *Proceedings of the 42nd IEEE International Solid-State Circuits Conference (ISSCC '05)*, vol. 1, pp. 396–606, San Francisco, Calif, USA, February 2005.

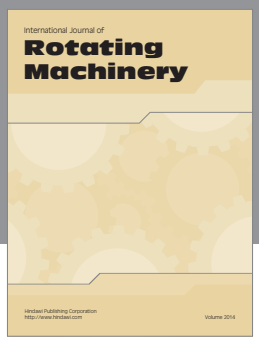
[32] S. M. Ross, *Stochastic Processes*, John Wiley & Sons, New York, NY, USA, 1983.

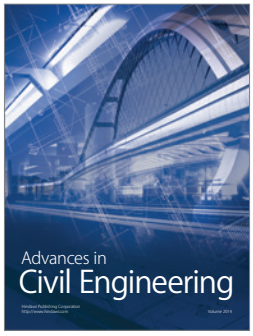
[33] S. Verdu, "Spectral efficiency in the wideband regime," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1319–1343, 2002.

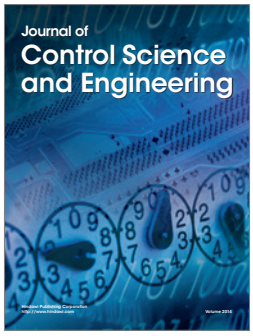
[34] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.

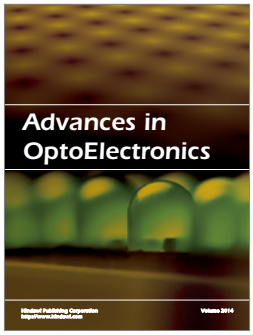
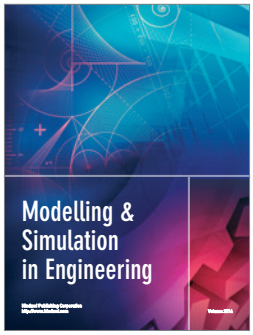
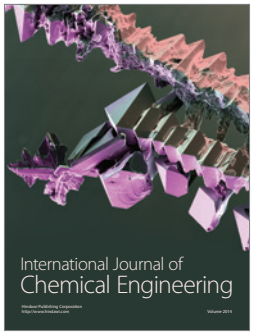
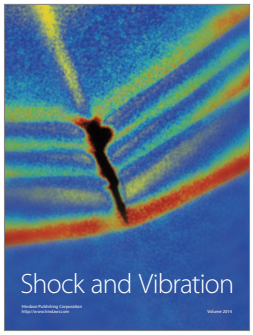
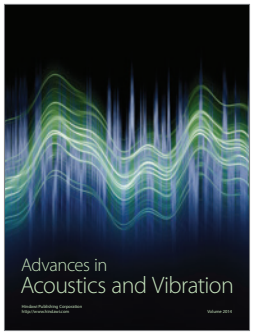
[35] P. Billingsley, *Convergence of Probability Measures*, John Wiley & Sons, New York, NY, USA, 1968.

[36] A. Gut, *Stopped Random Walks: Limit Theorems and Applications*, Springer, New York, NY, USA, 1988.







Hindawi

Submit your manuscripts at
http://www.hindawi.com