

Performance Modeling and Analysis of a Class of ARQ Protocols in Multi-Hop Wireless Networks

Teerawat Issariyakul, *Student Member, IEEE*, and Ekram Hossain, *Senior Member, IEEE*

Abstract—This paper models and analyzes the performances of a class of ARQ (Automatic Repeat reQuest) protocols in a multi-hop wireless data network. The performance metric here is the number of transmissions required for successful delivery of a packet over a multi-hop path. By using a discrete-time Markov model, the distribution for the total required number of transmissions is modeled as *phase type* distribution. The effects of different network parameters—such as packet error rate in each hop, maximum number of allowable retransmissions at each hop and retransmission probability at each hop—on the required total number of transmissions are investigated. The novelty of this model is that the probability mass function (*pmf*) for the number of transmissions required for successful end-to-end delivery of a packet can be easily obtained under different hop-level error control policies. Using the *pmf*, the tradeoff between transmission energy and percentage of data delivery (i.e., reliability) in a multi-hop path can be analyzed. The analytical model is validated by simulations. While the proposed analytical framework is general enough to capture the impact of any MAC (Medium Access Control) mechanism at each hop, we specifically present typical performance results under IEEE 802.11 DCF (Distributed Coordination Function) MAC.

Index Terms—Multi-hop wireless networks, ARQ, phase type distribution, discrete-time Markov process.

I. INTRODUCTION

MULTI-HOP wireless networks are characterized by the lack of direct communication link between source and destination nodes. End-to-end data transmission between a pair of nodes in this type of networks requires intermediate nodes to forward data packets. For example, in a multi-hop cellular wireless network [1], the communications between some mobiles and the base station are carried out via relay nodes. The use of relay nodes helps increase service area and prevent network partitioning. In addition, short-range transmission improves spectral efficiency, increases spatial reuse, and leads to higher energy efficiency. A wireless backhaul network is another type of multi-hop network which consists of a collection of TAPs (Transit Access Points) [2]. These wireless TAPs forward traffic from mobiles to the Internet gateway in a multi-hop manner. For successful end-to-end transmission of a packet, the packet needs to be successfully transmitted across all the links. Therefore, if the transmission fails (due to collision and/or channel fading) in one of the nodes *en route* the source and the destination nodes, retransmissions based on an ARQ (Automatic Repeat reQuest) policy will be necessary.

Manuscript received February 16, 2004; revised February 13, 2004; accepted March 23, 2005. The associate editor coordinating the review of this paper and approving it for publication was H. Yanikomeroglu.

The authors are with the Department of Electrical and Computer Engineering of University of Manitoba, Winnipeg, MB, Canada R3T 5V6 (e-mail: {teerawat, ekram}@ee.umanitoba.ca).

Digital Object Identifier 10.1109/TWC.2006.04121.

The end-to-end performance in a wireless multi-hop network depends strongly on the hop-level protocols and parameters. The study in [3] showed that the energy efficiency and end-to-end throughput (e.g., TCP (Transmission Control Protocol) throughput) depend greatly on hop-level error probability, transmission range of each node, and maximum number of retransmissions at each node. However, only average values of the performance metrics such as energy efficiency or throughput were obtained. Performance of end-to-end congestion and flow control mechanism in TCP over IEEE 802.11 DCF (Distributed Coordination Function) was investigated in [4] through simulations. In [5], the impact of TCP on end-to-end throughput performance in a multi-hop wireless network was analyzed and the optimal transmission window size was determined to be $H/4$ for a single TCP flow over an H -hop linear chain topology. However, hop-level error control policies were not considered.

An analytical model for computing average steady state TCP throughput for a two-hop chain topology was presented in [6] assuming a collision-free and error-free wireless channel. For a two-hop chain path, [7] and [8] modeled batch transmission under *probabilistic retransmission with infinite persistence* ARQ and under *infinite retransmission* ARQ with multi-rate transmission¹, respectively. After all, a generalized analysis of the impacts of different hop-level error control policies on the end-to-end performance in a multi-hop network with *arbitrary number of hops* has not been reported in the literature.

This paper presents an analytical methodology to study the impact of radio link error and different hop-level ARQ policies on the end-to-end performance in a multi-hop wireless path. Specifically, by using *phase type* (PH) distribution, we derive the probability mass function (*pmf*) of total required number of hop-level transmissions for successful end-to-end delivery of a packet in an H -hop chain topology. The usefulness of our analysis comes from the following facts. First, a general end-to-end transmission cost function can be defined in terms of number of hop-level transmissions. Based on this cost, the optimal routing paths (e.g., minimum energy paths [9]) can be determined for reliable communication in a multi-hop wireless network. Secondly, the cumulative distribution function (*cdf*) of the required number of transmissions can be utilized to quantify the reliability-energy tradeoff, since it is the probability to deliver a packet to the destination within a limited number of transmissions (and hence limited amount of transmission energy). Thirdly, statistics for end-to-end latency can be estimated if the information about queueing

¹The definitions of these ARQ protocols are given in Section III-C.

and channel access delay is available. This statistics can be used to set the transport control protocol timeout value at the source node with a view to improving the end-to-end performance. Finally, since the proposed model is based on a generic link error process, the impact of different MAC and PHY-layer parameters on the end-to-end performance under different hop-level error control strategies can be analyzed, and hence cross-layer design and engineering can be performed.

The rest of the paper is organized as follows. Section II provides a summary of the mathematical preliminaries used for the development of the analytical framework. The system model and assumptions (specifically the packet error model and the different ARQ protocols) are presented in Section III. The Markov-based analytical model is presented in Section IV. The numerical and the simulation results as well as their useful implications are presented in Section V. Finally, conclusions are stated in Section VI.

II. MATHEMATICAL PRELIMINARIES

A. Absorbing Markov Process

An *absorbing Markov process* is a Markov process² which will finally stop at a particular state [10]. The states at which the process might stop are called *absorbing states*. If state i is an absorbing state, the transition probability from state i to j (p_{ij}) will be zero for all $j \neq i$ and will be 1 for $j = i$. Consider an $(S_0 + S)$ -state absorbing Markov process with S_0 absorbing states. The complete description of this Markov process is $((\alpha_0, \alpha), \mathbf{P})^3$, where the initial probability row vector $(\alpha_0, \alpha) \in [0, 1]^{1, S_0+S}$ refers to a set of probabilities corresponding to different states where the process might start and the transition probability matrix (TPM) $\mathbf{P} \in [0, 1]^{S_0+S, S_0+S}$ can be written in the form as follows:

$$\mathbf{P} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{t} & \mathbf{T} \end{pmatrix}. \quad (1)$$

The matrices \mathbf{T} and \mathbf{t} represent sets of probabilities with which the process is moving within transient states and moving from transient states to absorbing states, respectively. Hereafter, we denote all-zero, all-one, and identity matrices by $\mathbf{0}$, $\mathbf{1}$, and \mathbf{I} , respectively. Both $\mathbf{0}$ and \mathbf{I} imply that it is impossible to leave the absorbing states.

B. Phase Type Distribution

Phase type distribution is the distribution of time to absorption in an absorbing Markov process with one absorbing state. The distribution of time (in terms of number of transitions (k)) the process requires to reach the absorbing state (f_k) is given by (2), where α_0 is the probability that the process starts at the absorbing state [11]. The cumulative distribution function (F_k) and the expected time to absorption ($E[k]$) can be calculated

²All Markov processes in this paper are *discrete-time* Markov chain (DTMC).

³Throughout this paper, we use regular and bold-face type letters to represent scalar values and matrices, respectively. The notation $\mathbf{A} \in [b, c]^{d,e}$ denotes a $d \times e$ matrix (\mathbf{A}) whose entries are in $[b, c]$.

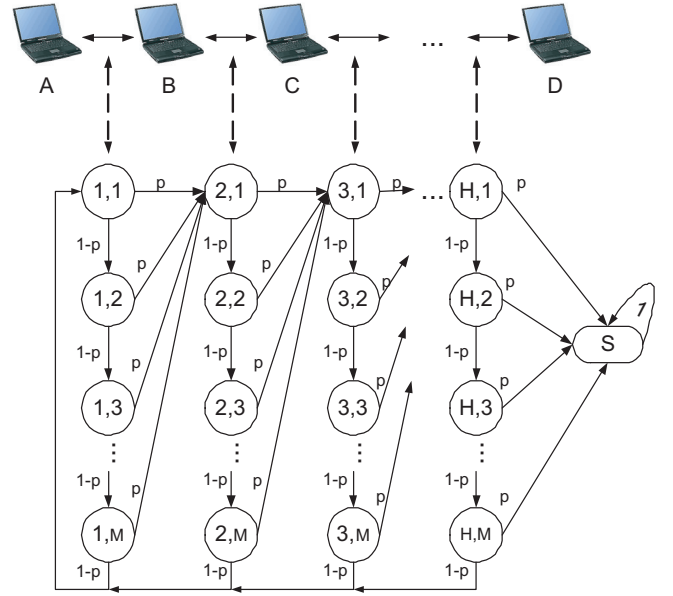


Fig. 1. An example of a chain topology and the corresponding Markov process for ARQ^F.

using (3) and (4), respectively.

$$f_k = \begin{cases} \alpha_0, & k = 0 \\ \alpha \mathbf{T}^{k-1} \mathbf{t}, & k \geq 1. \end{cases} \quad (2)$$

$$F_k = \sum_{i=0}^k f_i = \begin{cases} \alpha_0, & k = 0 \\ \alpha_0 + \mathbf{1} - \alpha \mathbf{T}^k \mathbf{1}, & k \geq 1. \end{cases} \quad (3)$$

$$E[k] = \alpha (\mathbf{I} - \mathbf{T})^{-2} \mathbf{t} = \alpha (\mathbf{I} - \mathbf{T})^{-1} \mathbf{1}. \quad (4)$$

III. SYSTEM MODEL AND ASSUMPTIONS

A. Multi-Hop Network Model

We consider a data packet transmission scenario from a source node (A) to a destination node (D) over a multi-hop wireless path (Fig. 1). We use a *chain topology* to represent the flow under consideration and regard all other active flows as background traffic⁴. We derive the statistics for the total number of hop-level transmissions required for successful end-to-end delivery of a particular packet.

After transmitting a packet, each node can determine whether the transmission has been successful or not. If the transmission has failed, the node will invoke a retransmission procedure based on the ARQ policy being used at that node. If the packet is dropped (e.g., in case of *zero-retransmission policy*), the source node will retransmit the dropped packet.

B. Packet Error Model

Transmission of a packet in a link may fail due to data collision (when a distributed MAC protocol is used) and/or channel fading (independent or time-correlated). Data collision depends primarily on the underlying MAC layer. For example, the steady-state collision probability (q_c) for an IEEE 802.11 DCF-based MAC was analyzed in [12]. In a channel with independent channel fading, packet transmission is assumed

⁴Similar topology is used in a wireless backhaul network [2].

to be in error with probability q_f . Correlated channel fading can be modeled by the *two state Markov* channel [13], where the channel state of a specific link during a particular period of time is either good (g) or bad (b), and is characterized by the transition probability matrix \mathbf{Q} as follows:

$$\mathbf{Q} = \left(\begin{array}{c|cc} \text{State} & g & b \\ \hline g & v & 1-v \\ b & 1-w & w \end{array} \right) \quad (5)$$

where v and w are the probabilities with which the channel stays in good and bad states, respectively. Given that the packet error probability when the channel is in the good state and in the bad state is $q_f^{(g)}$ and $q_f^{(b)}$, respectively, the steady state packet error probability due to fading (q_f) can be calculated from (6) below

$$q_f = \frac{q_f^{(g)}(1-w) + q_f^{(b)}(1-v)}{2-v-w}. \quad (6)$$

The values of $q_f^{(g)}$ and $q_f^{(b)}$ above depend on the SNR (Signal-to-Noise Ratio) in each channel state, and the parameters v and w can be calculated based on fading margin and normalized Doppler bandwidth.

In general, collision and fading are independent of each other. A methodology to calculate successful transmission probability (p) for an IEEE 802.11 DCF-based MAC under Rayleigh fading channel was given in [14]. For a general MAC under independent fading channel, p can be calculated from $(1-q_f) \cdot (1-q_c)$. On the other hand, the successful transmission probability in a good and bad state of a correlated fading channel can be calculated as $p^{(g)} = (1-q_f^{(g)}) \cdot (1-q_c)$ and $p^{(b)} = (1-q_f^{(b)}) \cdot (1-q_c)$, respectively. By replacing $q_f^{(g)}$ and $q_f^{(b)}$ in (6) with $1-p^{(g)}$ and $1-p^{(b)}$, respectively, the steady state packet error probability for a correlated fading channel can be determined.

C. Hop-Level Automatic Repeat ReQuest (ARQ) Model

If a node fails to deliver a packet to the next node, it will retransmit the packet according to one of the following hop-level ARQ policies:

- **Zero retransmission** (ARQ^0): If a transmission fails, the packet will be dropped immediately.
- **Infinite retransmission** (ARQ^∞): The packet is retransmitted repeatedly until the transmission is successful.
- **Finite retransmission** (ARQ^F): The maximum number of retransmissions allowed is $M-1$, and the packet will be dropped if the $(M-1)^{\text{th}}$ retransmission fails.
- **Probabilistic retransmission with infinite persistence** (ARQ^P): If a transmission fails, the packet will be dropped with probability d and retransmitted with probability $(1-d)$.
- **Probabilistic retransmission with finite persistence** (ARQ^{FP}): This is similar to ARQ^P with the following constraints: $d < 1$ for first $M-2$ unsuccessful retransmissions and $d = 1$ for the $(M-1)^{\text{th}}$ unsuccessful retransmission.

We define *link⁵ reliability* for a specific ARQ policy (r_{link}^{ARQ}) as the unconditional probability that a packet is successfully transmitted (before being dropped).

THEOREM 1: For different ARQ policies, r_{link}^{ARQ} can be calculated as follows:

$$r_{link}^{\text{ARQ}^0} = p, \quad (7)$$

$$r_{link}^{\text{ARQ}^\infty} = 1, \quad (8)$$

$$r_{link}^{\text{ARQ}^F} = 1 - (1-p)^M, \quad (9)$$

$$r_{link}^{\text{ARQ}^P} = \frac{p}{p+d-pd}, \quad (10)$$

$$r_{link}^{\text{ARQ}^{FP}} = \frac{p(1-(1-p)^M(1-d)^M)}{p+d-pd} \quad (11)$$

where p is the probability of successful packet transmission over a link, M is the maximum number of allowable transmissions in case of ARQ^F , and d is the packet dropping probability for ARQ^P . \square

Proof: See Appendix I. \blacksquare

We observe that ARQ^0 and ARQ^∞ provide lower-bound and upper-bound for link reliability with $r_{link}^{\text{ARQ}^0} = p$ and $r_{link}^{\text{ARQ}^\infty} = 1$, respectively. To achieve a target link reliability r_{link}^* , the minimum number of transmissions M^* (in case of ARQ^F) or the maximum dropping probability d^* (in case of ARQ^P) can be obtained as follows:

$$M^* = \left\lceil \frac{\log(1-r_{link}^*)}{\log(1-p)} \right\rceil, \quad d^* = \frac{p(1-r_{link}^*)}{r_{link}^*(1-p)}. \quad (12)$$

In Section V-G, we will show that THEOREM 1 would also be useful in estimating the end-to-end latency for a packet in a multi-hop route.

IV. MARKOV MODEL AND ANALYSIS

A. Markov Modeling Under Independent Packet Error Process

We model a multi-hop wireless path by an absorbing DTMC (Fig. 1). The matrices (α_0, α) , \mathbf{T} , and \mathbf{t} are formulated for ARQ^{FP} only, since it is the most general ARQ. Let $X^{(t)} \in \{SC, (h^{(t)}, m^{(t)})\}$ be the state of a tagged packet at service opportunity t . At opportunity t , the packet either reaches the destination ($X^{(t)} = SC$) or is being transmitted/waiting to be transmitted in one of the nodes along the route ($X^{(t)} = (h^{(t)}, m^{(t)})$), where the hop number ($h^{(t)} \in \{1, 2, \dots, H\}$) corresponds to the link where the packet is being transmitted and $m^{(t)} \in \{1, 2, \dots, M\}$ is the number of transmissions in the corresponding link.

The Markov process always starts when the packet is fed to the first node and finishes when the packet traverses the last hop. Therefore, we set $X^{(0)} = (1, 1)$ to be the initial state and $X^{(t)} = SC$ to be the absorbing state. Correspondingly, $\alpha_0 = 0$ and $\alpha = [1 \ 0 \ 0 \ \dots] \in [0, 1]^{1, (M \cdot H)}$.

We now construct \mathbf{T} and \mathbf{t} for the above DTMC. For ARQ^{FP} , $S_0 = 1$, $\mathbf{I} = 1$, and the matrices \mathbf{t} and \mathbf{T} are formulated as in (13). Both of the matrices consist of blocks of sub-matrices. A transition among these blocks is equivalent to a change in the hop number (h), while a transition within a particular block represents the change in number

⁵We use the terms *link* and *hop* interchangeably in this paper.

TABLE I
 IMPLICATIONS OF THE SUB-MATRICES

Matrix	Event at time n	(h_{n+1}, m_{n+1})	Implication
Unsuccessful (U)	$m_n < M$ and transmission fails.	$(h_n, m_n + 1)$	Transmission fails and the current node retransmits the packet.
Successful (S)	Transmission is successful.	$(h_n + 1, 1)$	Successful transmission. Start transmission in the next hop. Reset m to 1.
Restart (R)	$m_n = M$ and transmission fails.	(1,1)	Packet is dropped. Retransmit from the hop $h = 1$.
Zero (0)	Not possible	Not available	Not possible

of unsuccessful transmissions (m) in the same hop. The implications of all the sub-matrices are explained in Table I. The elements in row i and column j of the sub-matrices $\mathbf{U}_h \in [0, 1]^{M, M}$, $\mathbf{S}_h \in [0, 1]^{M, M}$, $\mathbf{R}_h \in [0, 1]^{M, M}$, and $\mathbf{s}' \in [0, 1]^{M, 1}$ (i.e., $u_h(i, j)$, $s_h(i, j)$, $r_h(i, j)$, $s'(i, 1)$, respectively) can be obtained from (14)-(17) below

$$(\mathbf{t}|\mathbf{T}) = \left(\begin{array}{c|cccccc} \mathbf{0} & \mathbf{U}_1 + \mathbf{R}_1 & \mathbf{S}_1 & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{R}_2 & \mathbf{U}_2 & \mathbf{S}_2 & \mathbf{0} & \cdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \\ \mathbf{0} & \mathbf{R}_{H-1} & \mathbf{0} & \cdots & \mathbf{U}_{H-1} & \mathbf{S}_{H-1} \\ \mathbf{s}' & \mathbf{R}_H & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{U}_H \end{array} \right). \quad (13)$$

$$u_h(i, j) = \begin{cases} (1 - p(i, h)) \cdot (1 - d(i, h)), & j < M, j = i + 1 \\ 0, & \text{otherwise} \end{cases} \quad (14)$$

$$s_h(i, j) = \begin{cases} p(i, h), & j = 1 \quad \forall j \\ 0, & \text{otherwise} \end{cases} \quad (15)$$

$$r_h(i, j) = \begin{cases} (1 - p(i, h)) \cdot d(i, h), & j = 1 \\ 0, & \text{otherwise} \end{cases} \quad (16)$$

$$s'(i, 1) = p(i, H) \quad (17)$$

where $p(m, h)$ denotes probability of successful transmission corresponding to the m^{th} transmission in hop h , and $d(m, h)$ denotes dropping probability when the m^{th} transmission in hop h fails. By applying the formulated matrices (i.e., $\alpha_0 = 0$, α , \mathbf{T} , \mathbf{t}) to (2)-(4), the pmf (f_k), cdf (F_k), and the expected value of the number of transmissions ($E[k]$) required for successful end-to-end delivery can be calculated.

B. Markov Modeling Under Correlated Packet Error Process

The analyses above can be easily extended for the packet error model under correlated fading. At service opportunity t , a channel with correlated error (due to fading) can be in either good or bad state and the corresponding successful transmission probability is $p^{(g)}$ and $p^{(b)}$, respectively. For simplicity, we assume a homogeneous link condition across all the links. The analyses for heterogeneous link conditions can be performed in a similar manner. Hereafter, we denote the parameters for a correlated-error channel by superscript *corr*.

Assuming $p(m, h) \in \{p^{(g)}, p^{(b)}\}$, $\forall m, h$, we modify the matrices α , \mathbf{T} , and \mathbf{t} to support correlated error. We embed one more variable into the Markov process to keep track of the channel state. Therefore, the TPM becomes

$$(\mathbf{t}^{\text{corr}}|\mathbf{T}^{\text{corr}}) = \left(\begin{array}{c|cccccc} \mathbf{0} & \mathbf{U}_1^{\text{corr}} + \mathbf{R}_1^{\text{corr}} & \mathbf{S}_1^{\text{corr}} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{R}_2^{\text{corr}} & \mathbf{U}_2^{\text{corr}} & \mathbf{S}_2^{\text{corr}} & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{s}^{\text{corr}'} & \mathbf{R}_H^{\text{corr}} & \mathbf{0} & \mathbf{0} & \mathbf{U}_H^{\text{corr}} \end{array} \right) \quad (18)$$

$$\mathbf{U}_i^{\text{corr}} = \frac{\mathbf{U}_i \otimes ((\mathbf{I} - \mathbf{G}) \cdot \mathbf{Q})}{1 - p}, \quad \mathbf{R}_i^{\text{corr}} = \frac{\mathbf{R}_i \otimes ((\mathbf{I} - \mathbf{G}) \cdot \mathbf{Q})}{1 - p}. \quad (19)$$

$$\mathbf{S}_i^{\text{corr}} = \frac{\mathbf{S}_i \otimes (\mathbf{G} \cdot \mathbf{Q})}{p}, \quad \mathbf{s}_i^{\text{corr}'} = \frac{\mathbf{s}' \otimes (\mathbf{G} \cdot \mathbf{1})}{p}. \quad (20)$$

where \otimes denotes Kronecker product, $\mathbf{G} = \begin{pmatrix} p^{(g)} & 0 \\ 0 & p^{(b)} \end{pmatrix}$, and \mathbf{Q} is defined in (5). After constructing the matrices, the statistics for correlated-error channel can be obtained by using α^{corr} , \mathbf{T}^{corr} , and \mathbf{t}^{corr} in (2)-(4).

C. Phase Type Representations for the Different ARQ Models

Assuming independent packet error process, we reduce entries of \mathbf{T} in (13) for the following special cases with $M = 1$ to scalar values.

1) ARQ^0 : In this case, $d = 1$ and $p(m, h) = p$ ($\forall m, h$). Therefore, $(\mathbf{U}_i, \mathbf{S}_i, \mathbf{R}_i) = (0, p, 1 - p)$, $\forall i$.

2) ARQ^∞ : For ARQ^∞ , $d = 0$ and $p(m, h) = p$ ($\forall m, h$). Therefore, $(\mathbf{U}_i, \mathbf{S}_i, \mathbf{R}_i) = (1 - p, p, 0)$, $\forall i$. Since \mathbf{T} is diagonal dominant, a closed-form expression for the pmf (f_k) can be obtained from THEOREM 2.

THEOREM 2: For ARQ^∞ , the pmf (f_k) of the number of transmissions required for a reliable end-to-end delivery can be calculated from

$$f_k = \begin{cases} \binom{k-1}{H-1} p^H (1-p)^{k-H}, & k \geq H \\ 0, & \text{otherwise.} \end{cases} \quad (21) \quad \square$$

Proof: See Appendix II. \blacksquare

In (21), f_k can be also regarded as a *negative binomial* or *Pascal* distribution which corresponds to the probability that there are $H - 1$ successes among $k - 1$ trials and another success at the k^{th} trial [10].

3) ARQ^P : In this case, $p(m, h) = p$ ($\forall m, h$), and $(\mathbf{U}_i, \mathbf{S}_i, \mathbf{R}_i) = ((1 - d) \cdot (1 - p), p, d \cdot (1 - p))$, $\forall i$.

V. NUMERICAL AND SIMULATION RESULTS: MODEL VALIDATION AND USEFUL IMPLICATIONS

For brevity, we present results only for the independent packet error process. The results for a correlated packet error

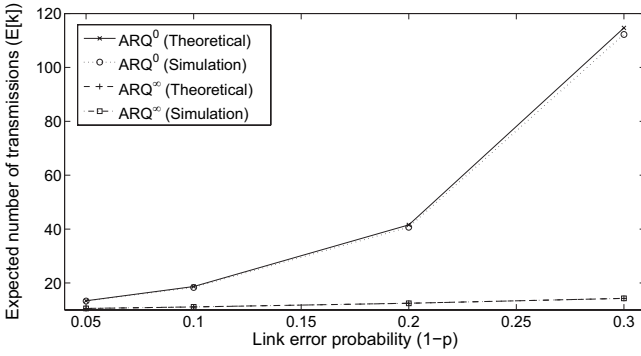


Fig. 2. Comparison between analytical and simulation results.

process can also be generated by using methodology provided in Section IV-B. Unless otherwise specified, we assume that all nodes implement the same type of ARQ and the probability of successful packet transmission is the same for all the links.

A. Model Validation

We verify the accuracy of our model by simulations using ns-2 [15]. We establish a 10-hop chain topology and insert a link-loss module as well as a specific ARQ model between each node and the connecting link. We run FTP (File Transfer Protocol) over TCP and plot the expected number of transmissions $E[k]$ as a function of packet error probability in each link (Fig. 2). During simulation, we measure the error probability in each link. The simulation terminates when the difference between input and measured link error probability is less than 10^{-6} . From the simulation, we calculate the average number of measured hop-level transmissions associated with each TCP packet, and compare it with that obtained from (4).

As expected, $E[k]$ increases as the link error probability increases and the results from the simulations are fairly close to those obtained from the analytical model (Fig. 2). When the packet error probability is high, for ARQ^0 , the packet rarely reaches the nodes closer to the destination node. In such a case, during simulation, the link-loss module for the corresponding hops is rarely invoked. For this reason, at high link error rate, the analytical results on $E[k]$ deviate slightly from the simulation results.

B. Impact of Hop-Level ARQ Policies on Expected Number of Transmissions

Fig. 3 plots $E[k]$ as a function of the number of hops (H) under different ARQ policies when the packet error probability in each link is 0.3. As expected, $E[k]$ increases as the number of hops in the route increases. Since with ARQ^0 each unsuccessful transmission at any intermediate node requires retransmission from the source node, the upper bound of $E[k]$ is observed for this ARQ policy. On the other hand, the lower bound of $E[k]$ is achieved with ARQ^∞ because in this case each intermediate node retransmits the lost packet until the transmission is successful.

The expected total number of transmissions increases as the maximum number of allowable transmissions in each hop (M) decreases (for ARQ^F) and/or the dropping probability (d) increases (for ARQ^P). In any case, the value of $E[k]$ is

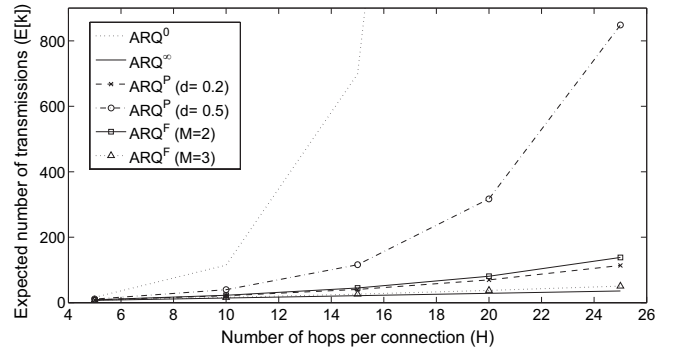


Fig. 3. Variations in the expected number of transmissions with number of hops in the route.

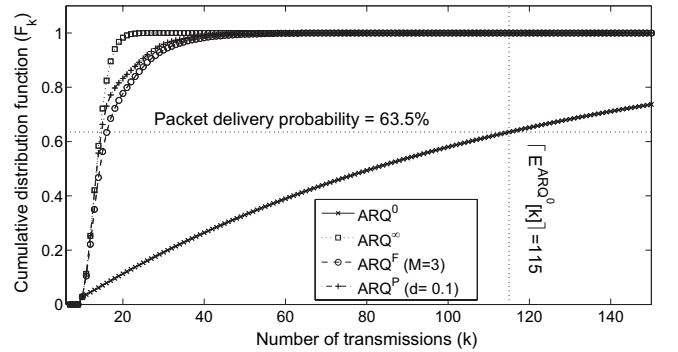


Fig. 4. Cumulative distribution function of the required number of transmissions (F_k) for different ARQ policies.

bounded by those for ARQ^0 and ARQ^∞ . Both ARQ^F and ARQ^P policies are complementary to each other in that they yield the same $E[k]$ when the parameters M and d are properly adjusted. Note that, $E[k]$ for ARQ^F converges to the lower-bound fairly fast. From Fig. 3, the required total number of transmissions becomes very close to the lower-bound when $M = 3$.

In general, the average amount of energy spent per hop-level transmission is a function of parameters such as transmission range, modulation techniques, and packet size. Given the average amount of energy consumption for a packet per hop-level transmission \bar{e} , the average energy spent for successful end-to-end delivery of a packet can be calculated as $E[k] \times \bar{e}$. Since $E[k]$ depends solely on the hop-level reliability and the number of hops, but neither on queuing delay nor on channel access delay, the expected number of transmissions required to deliver a window of N packets is $N \times E[k]$. Also, the corresponding *pmf* for a window of N packets can be obtained by convoluting f_k in (2) for N times.

C. Distribution of the Total Required Number of Transmissions

Fig. 4 illustrates the *cdf* (F_k) of the required number of transmissions for a 10-hop connection when the packet error probability in each hop is 0.3. With the same argument, ARQ^0 and ARQ^∞ converge to 1 at the lowest and highest rate. Again, the rate of convergence of ARQ^F and ARQ^P fall in between the above two.

We observe that the expected number of transmissions in

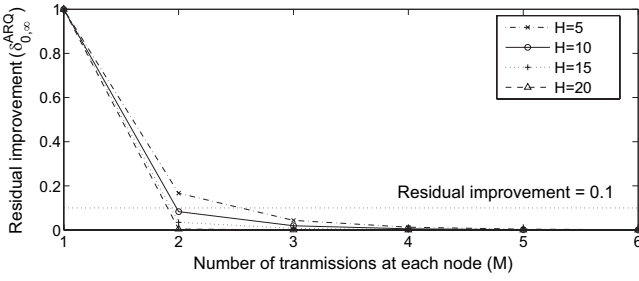


Fig. 5. Variations in residual improvement with maximum number of allowable transmissions at each node.

ARQ⁰ ensures less than 63.5% of packet delivery. Since F_k represents the probability that a packet will be delivered to the destination within k transmissions, using F_k , the tradeoff between reliability and energy consumption can be analyzed.

D. Residual Improvement

Although ARQ[∞] is the best ARQ policy in terms of $E[k]$, it may cause head-of-line (HOL) blocking, and consequently, result in large hop-level delay and/or buffer overflow. Therefore, ARQ^F and ARQ^P might be preferable to ARQ[∞] in some scenarios. For these ARQ policies, there exist values of M or d further increase or decrease of which do not result in significant improvement in the total required number of transmissions but would rather cause the HOL blocking problem.

To quantify the improvement for a particular ARQ policy (compared to the improvement from lower-bound to upper-bound corresponding to ARQ^{lb} and ARQ^{ub}, respectively), we define a parameter *Residual Improvement* (δ) as follows:

$$\delta_{(lb,up)}^{ARQ} \triangleq \frac{\sum_{k=0}^{\infty} (F_k^{ub} - F_k^{ARQ})}{\sum_{k=0}^{\infty} (F_k^{ub} - F_k^{lb})}. \quad (22)$$

For two ARQ policies, namely, ARQ1 and ARQ2, it can be shown that (Appendix III)

$$\sum_{k=0}^{\infty} F_k^{ARQ1} - \sum_{k=0}^{\infty} F_k^{ARQ2} = E^{ARQ2}[k] - E^{ARQ1}[k] \quad (23)$$

where F_k^{ARQ} and $E^{ARQ}[k]$ refer to *cdf* and expectation of the number of transmissions for successful end-to-end delivery for a particular ARQ policy. Therefore,

$$\delta_{(lb,ub)}^{ARQ} = \frac{E^{ARQ}[k] - E^{ub}[k]}{E^{lb}[k] - E^{ub}[k]}. \quad (24)$$

In this section, we use ARQ⁰ and ARQ[∞] as the lower-bound and the upper-bound, respectively. As can be observed from Fig. 4, $\delta_{(lb,ub)}^{ARQ}$ is in fact the ratio of two areas—the area between the *cdf* for a certain ARQ and the *cdf* for ARQ[∞] and the area between the *cdf* for ARQ⁰ and the *cdf* for ARQ[∞]. Since $E[k]^{ARQ\infty} \leq E[k]^{ARQ} \leq E[k]^{ARQ^0}$, $\delta_{(0,\infty)}^{ARQ}$ lies between 0 and 1, where $\delta_{(0,\infty)}^{ARQ^0} = 1$, and $\delta_{(0,\infty)}^{ARQ\infty} = 0$.

Figs. 5-6 show typical variations in the residual improvement ($\delta_{0,\infty}^{ARQ^F}$ and $\delta_{0,\infty}^{ARQ^P}$) as a function of the maximum number of allowable transmissions (M) and the dropping

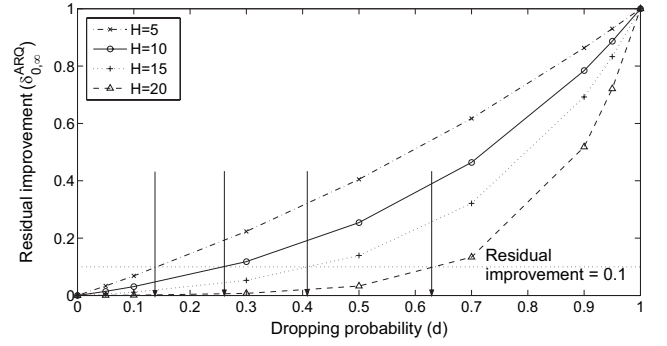


Fig. 6. Variations in residual improvement with packet dropping probability at each node.

probability (d) in each hop, when the packet error probability in each hop is fixed to 0.3 and the number of hops varies from 5, 10, 15, to 20. As expected, for a particular H , $\delta_{0,\infty}^{ARQ^F}$ and $\delta_{0,\infty}^{ARQ^P}$ decrease with increasing M and decreasing d , implying that the corresponding *cdf* becomes closer to the *cdf* for ARQ[∞] (upper-bound). The performance of both ARQ^F and ARQ^P converge to the upper and the lower bounds when $M = \infty$ and 1 and $d = 0$ and 1, respectively.

We can achieve a certain level of residual improvement by adjusting M and d . For example, for a target value of $\delta_{0,\infty}^{ARQ^F} \leq 0.1$ and $\delta_{0,\infty}^{ARQ^P} \leq 0.1$, M must be chosen to be 3 for a 5-hop connection (Fig. 5) and d must be set to the values located at the end of the arrows in Fig. 6. As H increases, both $E^{lb}[k] - E^{ub}[k]$ and $E^{ARQ}[k] - E^{ub}[k]$ increase. However, the rate of increase of $E^{lb}[k] - E^{ub}[k]$ (the denominator in (24)) is higher than that of $E^{ARQ}[k] - E^{ub}[k]$. Therefore, as H increases, the residual improvement relative to the $E^{lb}[k] - E^{ub}[k]$ becomes smaller. In effect, for larger H , smaller and larger values of M and d , respectively, can still satisfy the same constraint on residual improvement.

E. Heterogeneous Wireless Links

In general, packet error probability in each link in a multi-hop route can be different. This section presents typical numerical results under heterogeneous wireless links. We assume that each link can be either ‘good’ with $p = 0.9$ or ‘bad’ with $p = 0.7$, and show the *cdf* for the required number of transmissions for successful end-to-end delivery (F_k) under ARQ[∞] in Fig. 7. The legends in this figure correspond to cases where all links are ‘good’, all links are ‘bad’, ‘first bad’ (where all except the first link are ‘good’), and ‘last bad’ (where all except the last link are ‘good’).

With ARQ[∞], since all the nodes persist on transmitting the packet until the transmission is successful, the *cdf* in this case depends only on the number of ‘bad’ links but not on the location of the ‘bad’ link(s). However, the location of a ‘bad’ link does affect the performance of all other ARQ policies. Using the ‘all good’ and the ‘all bad’ cases for ARQ⁰ as the upper-bound and the lower-bound, respectively, we denote the residual improvements in case of ARQ⁰ for the ‘first bad’ and the ‘last bad’ cases by $\delta_{(all\ bad, all\ good)}^{first\ bad}$ and $\delta_{(all\ bad, all\ good)}^{last\ bad}$, respectively. For the ‘last bad’ case, it is more likely that the packet will be dropped at the last hop. If this happens,

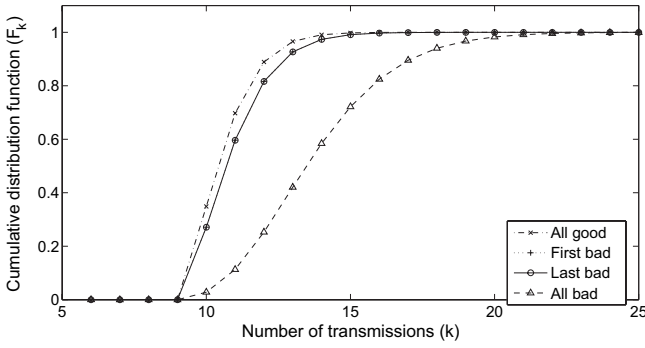


Fig. 7. Cumulative distribution function of the total required number of transmissions (F_k) for ARQ $^\infty$ under non-homogeneous wireless links.

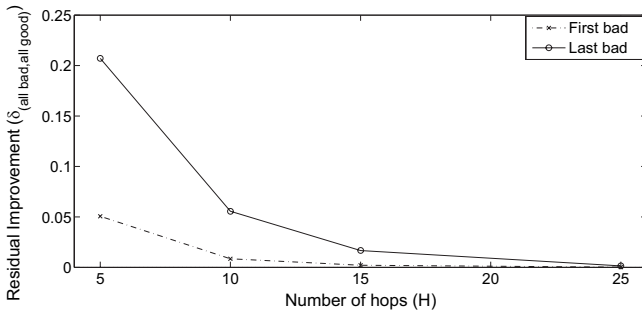


Fig. 8. Variations in residual improvement with number of hops under non-homogeneous wireless links.

the transmissions which have already been succeeded earlier will be useless, since the source will have to retransmit the packet. We observe that $\delta_{(all\ bad, all\ good)}^{first\ bad}$ is always less than $\delta_{(all\ bad, all\ good)}^{last\ bad}$ (Fig. 8). That is, the system performance drops when location of the ‘bad’ link moves towards the destination.

As the number of hops increases, the required number of transmissions in all the cases increases. With respect to the ‘all good’ case, the rate of relative increase in the ‘all bad’ case is the highest because it has more number of ‘bad’ links. In fact, this relative increase is the denominator in (24). Therefore, the residual improvement always decreases as H increases. Although $\delta_{(all\ bad, all\ good)}^{first\ good} \leq \delta_{(all\ bad, all\ good)}^{first\ bad}$, the difference becomes smaller as H increases because the denominator becomes more dominant.

F. Impact of MAC Protocols: Typical Results for IEEE 802.11 DCF

We run an FTP/TCP flow over a four hop linear chain topology. Implementing IEEE 802.11 MAC, each node has a transmission range of 250 m. The distance between any two nodes is also 250 m (we will refer to these nodes as chain nodes, hereafter). We place background nodes (each with transmission range of 100 m) within a distance of r meters from each chain node, where r is randomly chosen between 0 to 100. We also set the retry limit in IEEE 802.11 DCF to ∞ and compare the results with ARQ $^\infty$. This setting is necessary to prevent route failure along the chain of nodes.

Under a two-ray ground reflection propagation model, every node generates CBR (Constant Bit Rate) traffic at the rate of 10, 20, 30, 40, and 50 kbps, destined to the closest chain node.

Consisting of chain nodes as well as background nodes, this topology is very similar to that of wireless backhaul networks [2], where the chain nodes are analogous to the TAPs.

The simulation is run for 20 times. In each simulation, the first chain node sends out 1000 TCP segments destined to the last chain node. We measure the collision probability at each hop along the chain of nodes as well as the required number of hop-level transmissions for successful delivery of each TCP packet. Based on the measured collision probabilities, we calculate $E[k]$ from (4) and compare it with that obtained from simulation. Note that, this paper does not focus on modeling the collision probability q_c . Therefore, we use the measured value of q_c to calculate $E[k]$. The value for q_c can be obtained by estimating the number of interfering mobiles following the approach in [16] and then using the method presented in [12].

In the absence of channel fading, collision is the major cause of transmission failure. Fig. 9 (a) shows typical variations in collision probability in the first hop of the chain route, where the number of interfering nodes refers to the number of mobiles which can cause collision at each node in the chain route. Similar results have been observed in the other hops.

We observe that IEEE 802.11 suffers greatly from frequent data collisions due partly to the CBR background traffic, hence resulting in very low successful transmission probability. This result is in fact not surprising, since the collision probability for only 7 mobile nodes in the same neighborhood is expected to be greater than 20% [12]. In our simulations, this probability becomes even higher due to *hidden terminal jamming* problem [17].

Fig. 9 (b) compares $E[k]$ obtained from simulation (shown by symbols) with that obtained from (4) (shown by solid line). When collision probabilities are known, our model is extremely accurate in that all the symbols (simulation) fall very close to the line generated by (4).

G. Estimation of End-to-End Transmission Latency

The end-to-end latency (D) for a packet depends primarily on queuing delay (D_q), channel access delay (D_{acc}) associated with each hop-level transmission, total number of hop-level transmissions (k), and packet transmission time (D_{tx}). In each hop, the queuing delay for a tagged packet is the time that the packet waits before it reaches the head of the queue. The channel access delay is the time required for a packet to be transmitted after it reaches the head of the queue. The queuing delay depends on the buffer management and scheduling policies, while the channel access delay depends on the MAC scheme used by the nodes. The transmission time depends on the packet size and the link speed.

Assuming that the statistics for the average queuing delay ($E[D_q]$) and the average channel access delay ($E[D_{acc}]$) are available, the expected end-to-end latency for a packet can be calculated from $E[D] = E[n_q] \cdot E[D_q] + E[k] \cdot (E[D_{acc}] + D_{tx})$, where n_q is the number of times the packet is enqueued in the queues across the links in the multi-hop route before it reaches the destination node. Note that, the expected end-to-end throughput can be calculated as $1/E[D]$. The expected value of n_q ($E[n_q]$) for the different hop-level ARQ policies can be calculated as follows.

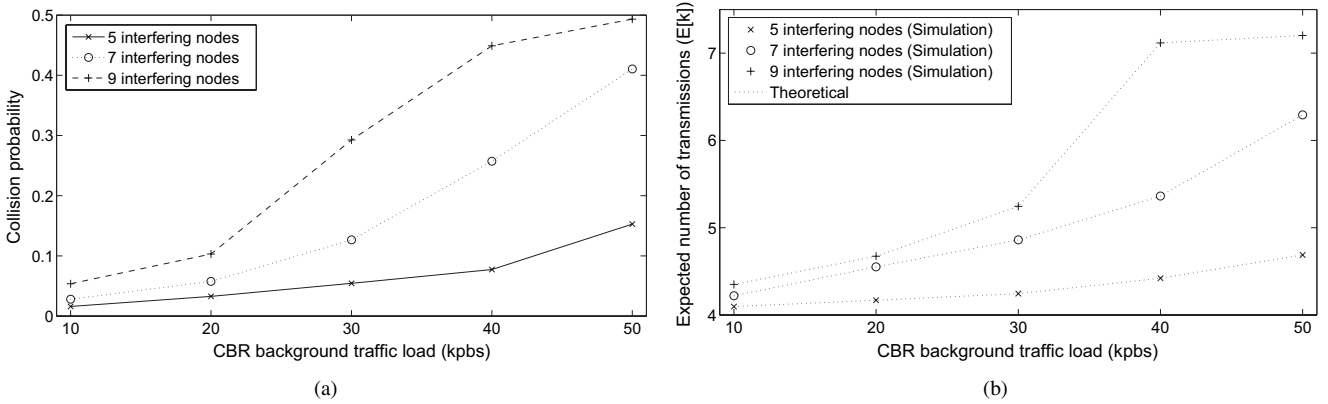


Fig. 9. Impact of 802.11 DCF MAC on (a) collision probability and on (b) expected number of transmissions.

First, evaluate link reliability (r_{link}^{ARQ}) for the implemented ARQ (in each hop) using THEOREM 1. Secondly, formulate \mathbf{U}_i , \mathbf{S}_i , and \mathbf{R}_i for ARQ⁰ (Section IV-C.1). Thirdly, replace p in the formulated matrices with the calculated r_{link}^{ARQ} . This substitution is equivalent to the use of ARQ⁰ at each hop along the route with the unconditional packet dropping probability at each node being equal to the one for the implemented ARQ. Finally, calculate $E[k]$, which in this case is equivalent to $E[n_q]$, by using (4).

VI. CONCLUSIONS

We have presented a methodology for analyzing the impacts of hop-level ARQ policies on end-to-end performance statistics in a multi-hop wireless network. Both the hop-level and the end-to-end performances are upper-bounded and lower-bounded by ARQ policies with infinite and zero retransmissions, respectively. Performances of all the ARQ policies, except that for ARQ with infinite retransmissions, degrade as the location of a weak wireless link moves towards the destination.

Simulation results have validated the analytical results. The proposed framework can be used to estimate the total amount of energy consumption and to analyze the tradeoff between reliability and energy for reliable end-to-end transmission. Also, it would be useful in estimating the end-to-end latency (hence throughput) and improving the end-to-end flow control mechanism. Consideration of wireless channel parameters and medium access control schemes would extend the use this model for analyzing cross-layer protocol performance. After all, the proposed model can be used to effectively study the inter-relationship among link-level packet error probability, hop-level ARQ policy and parameters, and the end-to-end performance in a multi-hop wireless network.

APPENDIX I PROOF OF THEOREM 1

- 1) ARQ⁰: Since the packet is dropped after one transmission, $r_{link}^{ARQ^0} = p$.
- 2) ARQ[∞]: Since the packet will never be dropped, $r_{link}^{ARQ^\infty} = 1$.
- 3) ARQ^F: $r_{link}^{ARQ^F} = \sum_{i=1}^M p(1-p)^{i-1} = 1 - (1-p)^M$.

- 4) ARQ^P: $r_{link}^{ARQ^P} = \sum_{i=1}^{\infty} p((1-p)(1-d))^{i-1} = \frac{p}{p+d-pd}$.
- 5) ARQ^{FP}: $r_{link}^{ARQ^{FP}} = \sum_{i=1}^M p((1-p)(1-d))^{i-1} = \frac{p(1-(1-p)^M(1-d)^M)}{p+d-pd}$.

APPENDIX II PROOF OF THEOREM 2

LEMMA 1: The element in row i column h of \mathbf{T}^k or the transition probability from state i to h in k steps denoted by $T_{i,h}^{(k)}$ in ARQ[∞] can be calculated from

$$T_{i,h}^{(k)} = \begin{cases} \binom{k}{h-i} p^{h-i} (1-p)^{k-(h-i)}, & k \geq h-i \\ 0, & \text{otherwise.} \end{cases} \quad (25) \quad \square$$

Proof: Since \mathbf{T} is nil-potent, $T_{i,h}^{(k)} = 0$ for $k < h-i$. We prove LEMMA 1 for $k \geq h-i$ by induction. From Section IV-C.2,

$$T_{i,j}^{(k)} = \begin{cases} p, & j = i+1 \text{ and } (j \neq h) \\ 1-p, & (j = i) \\ 0, & \text{otherwise.} \end{cases} \quad (26)$$

Using Chapman-Kolmogorov equation, for ARQ[∞]:

$$T_{i,h}^{(k)} = \begin{cases} (1-p) \cdot T_{i,h}^{(k-1)} + p \cdot T_{i+1,h}^{(k-1)}, & i \leq h \\ (1-p) \cdot T_{i,h}^{(k-1)}, & i = h \\ 0, & \text{otherwise.} \end{cases} \quad (27)$$

STEP 1: Obviously, LEMMA 1 is true for $k = 2$.

STEP 2: Assume that LEMMA 1 is true for $k > 0$.

STEP 3: We prove that LEMMA 1 is true for $k+1$. From (27),

- CASE I: $i < h$

$$\begin{aligned} T_{i,h}^{(k+1)} &= (1-p) \cdot T_{i,h}^{(k)} + p \cdot T_{i+1,h}^{(k)} \\ &= \binom{k+1}{h-i} p^{h-i} (1-p)^{(k+1)-(h-i)}. \end{aligned} \quad (28)$$

- CASE II: $i = h$

$$T_{i,h}^{(k+1)} = (1-p) \cdot T_{i,h}^{(k)} = (1-p)^{k+1}. \quad (29)$$

We observe that (28) and (29) are the same as those provided by LEMMA 1, and therefore, the proof is complete. ■

We now prove *Theorem 2*. From (2),

$$\begin{aligned} f_k &= \alpha \mathbf{T}^{k-1} \mathbf{t} \\ &= [1 \ 0 \ 0 \dots] \mathbf{T}^{k-1} [0 \ 0 \dots 0 \ p]^T = p \cdot T_{1,H}^{(k-1)}. \end{aligned} \quad (30)$$

By applying LEMMA 1 to (30), *Theorem 2* is proven.

APPENDIX III PROOF OF EQ. (23)

$$\begin{aligned} &\sum_{k=0}^{\infty} (F_k^{ARQ1} - F_k^{ARQ2}) \\ &= \sum_{k=1}^{\infty} \{1 - \alpha \cdot (\mathbf{T}^{ARQ1})^k \cdot \mathbf{1}\} - \{1 - \alpha \cdot (\mathbf{T}^{ARQ2})^k \cdot \mathbf{1}\} \\ &= \alpha \cdot \sum_{k=1}^{\infty} (\mathbf{T}^{ARQ2})^k \cdot \mathbf{1} - \alpha \cdot \sum_{k=1}^{\infty} (\mathbf{T}^{ARQ1})^k \cdot \mathbf{1} \\ &= E^{ARQ2}[k] - E^{ARQ1}[k]. \end{aligned} \quad (31)$$

ACKNOWLEDGEMENTS

This work was supported in part by a scholarship from the TRILabs, Winnipeg, Canada, in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada, and in part by University of Manitoba Graduate Fellowship (UMGF). We are thankful to the anonymous reviewers for their comments.

REFERENCES

- [1] H.-Y. Hsieh and R. Sivakumar, "On using peer-to-peer communication in cellular wireless data networks," *IEEE Trans. Mobile Comput.*, vol. 3, no. 1, pp. 57–72, Jan.-Mar. 2004.
- [2] V. Gamberoza, B. Sadeghi, and E. W. Knightly, "End-to-end performance and fairness in multihop wireless backhaul networks," in *Proc. ACM MobiCom'04*.
- [3] S. Bansal *et al.*, "Energy efficiency and throughput for TCP traffic in multi-hop wireless networks," in *Proc. IEEE INFOCOM'02*, pp. 210–219.
- [4] S. Xu and T. Saadawi, "Does the IEEE 802.11 MAC protocol work well in multihop wireless ad hoc networks?" *IEEE Commun. Mag.*, vol. 39, no. 6, pp. 130–137, Jun. 2001.
- [5] Z. Fu *et al.*, "The impact of multihop wireless channel on TCP throughput and loss," in *Proc. IEEE INFOCOM'03*.
- [6] A. A. Kherani and R. Shorey, "Throughput analysis of TCP in multi-hop wireless networks with IEEE 802.11 MAC," in *Proc. IEEE WCNC'04*.
- [7] T. Issariyakul, E. Hossain, and A. S. Alfa, "Markov-based analysis of end-to-end batch transmission in a multi-hop wireless network," in *Proc. IEEE ICC'05*.
- [8] —, "Analysis of latency for reliable end-to-end batch transmission in multi-rate multi-hop wireless networks," in *Proc. IEEE ICC'05*.
- [9] S. Banerjee and A. Misra, "Minimum energy paths for reliable communication in multi-hop wireless networks," in *Proc. ACM MobiHoc'02*, pp. 146–156.

- [10] K. S. Trivedi, *Probability and Statistics with Reliability, Queuing and Computer Science Applications*. New York: Wiley & Sons, Inc., 2002.
- [11] M. F. Neuts, *Matrix-Geometric Solutions in Stochastic Models*, The John Hopkins University Press, 1981.
- [12] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 535–547, Mar. 2000.
- [13] A. Chockalingam *et al.*, "Performance of a wireless access protocol on correlated Rayleigh-fading channel with capture," *IEEE Trans. Commun.*, vol. 46, no. 5, pp. 644–655, May 1998.
- [14] L.-C. Wang, S.-Y. Huang, and A. Chen, "On the throughput performance of CSMA-based wireless local area network with directional antennas and capture effect: a cross-layer analytical approach," in *Proc. IEEE WCNC'04*.
- [15] The Network Simulator - ns-2. [Online]. Available: <http://www.isi.edu/nsnam/ns/>
- [16] G. Bianchi and I. Tinnirello, "Kalman filter estimation of the number of competing terminals in an IEEE 802.11 network," in *Proc. IEEE INFOCOM'03*.
- [17] C. Ware, T. Wysocki, and J. Chicharo, "Hidden terminal jamming problems in IEEE 802.11 mobile ad hoc networks," in *Proc. IEEE ICC'01*, pp. 261–265.



Teerawat Issariyakul (S'02) received his bachelor in Electrical Engineering from Thammasat University, Thailand in 1997. In 1999, he received M.Eng in telecommunications from Asian Institute of Technology (AIT), Thailand, where he had been awarded with Telephone Organization of Thailand Scholarship. He obtained his Ph.D. from the department of Electrical and Computer Engineering at the University of Manitoba. He was a researcher at TRILabs, Winnipeg, Canada. He held a scholarship from TRILabs as well as a University of Manitoba Graduate Fellowship (UMGF). His main research interests are in the area of modeling, analysis, and optimization of protocols for wireless networks.



Ekram Hossain (S'98-M'01-SM'06) is an Associate Professor in the Department of Electrical and Computer Engineering at University of Manitoba, Winnipeg, Canada. He received his Ph.D. in Electrical Engineering from University of Victoria, Canada, in 2000. He was a University of Victoria Fellow and also a recipient of the British Columbia Advanced Systems Institute (ASI) graduate student award. Dr. Hossain's research interests include design, analysis, and optimization of 'wireless' protocols, cognitive wireless networks, and mobile computing. Currently

he serves as an Editor for the *IEEE Transactions on Wireless Communications*, the *IEEE Transactions on Vehicular Technology*, the *KICS/IEEE Journal of Communications and Networks*, the *Wireless Communications and Mobile Computing Journal* (Wiley InterScience), the *International Journal of Vehicular Technology (IJVT)* (Hindawi Publishing), and the *International Journal of Sensor Networks (IJSNet)* (Inderscience Publishers). Dr. Hossain served as a technical program co-chair for the Symposium on "Next Generation Mobile Networks (NGMN'06)" held in conjunction with *ACM International Wireless Communications and Mobile Computing Conference (IWCMC'06)* during July 3–6, 2006 in Vancouver, Canada. One of the papers co-authored by him won the *Best Student Paper Award in ACM IWCMC'06*. Dr. Hossain served as a technical program committee member for the *IEEE WCNC'07*, *Globecom'06*, *ICC'06*, *ICC'05*, *WCNC'05*, *WCNC'04*, *Globecom'04*, *Globecom'03* and *IFIP Networking'05*. He was a recipient of the Lucent Technologies, Inc. research award for his contribution to the *IEEE International Conference on Personal Wireless Communications (ICPWC)*, 1997.