

A Sampling Method for Predicting End-to-End Delay Performance in Wireless Sensor Networks

Yu Chen and Izzat Darwazeh

Department of Electronic and Electrical Engineering, University College London (UCL), London, the U.K.

Email: {y.chen, i.darwazeh}@ee.ucl.ac.uk

Abstract—End-to-end delay performance is an important Quality of Service (QoS) metric in 5G communication systems and wireless sensor networks (WSNs). Recently, a multi-hop effective capacity model was proposed to provide accurate characterization of end-to-end delay performance in wireless multi-hop environments. However, this model was developed in the continuous-time domain, which accounts for a discrepancy in digital/discrete-time systems. In this paper, we extend such a model into the discrete-time domain and derive new mathematical formulae for tail probabilities of delay, delay mean and jitter in multi-hop cases. Furthermore, we propose a simple algorithm for end-to-end delay performance prediction based on the sampling method. By using publicly-available real traces from a wireless sensor network, we recreate these field experiments in a simulation platform to validate the algorithm. The results show that the algorithm gives satisfactory prediction.

I. INTRODUCTION

It is envisioned that the fifth-generation (5G) communication networks will support ubiquitous connectivity so the networks will seamlessly integrate different types of networks, such as the third- and fourth- generation telecommunication networks, wireless sensor networks (WSNs) and so on [1]. A WSN consists of a number of inter-connected wireless sensor nodes; a typical wireless sensor node has two modules, namely, a sensor module responsible for sensing specific data information and a communication module, ensuring the inter-connectivity between sensor nodes. Nowadays, the applications of this technology includes environmental monitoring [2], health care [3], logistics [4], video transmission [5] and many others.

End-to-end delay has been identified as an important Quality-of-Service (QoS) metric in 5G networks [6]. Moreover, such delay is also important in WSNs [7] because information that sensors acquired can sometimes be delay-sensitive, for example, information of a fire being detected is usually urgent and delay-sensitive. End-to-end delay characterization is an old and extensively studied topic in the classic queueing theory [8]. However, due to the high complexity of queues behavior in different nodes within wireless multi-hop networks, the realistic analysis of delay performance using the classic queueing theory becomes intractable [9]. Apart from the classic queueing theory, the effective capacity theory is a relatively recent-developing theory, which reveals a new mathematical relation between throughputs and tail probabilities of delay [10] in wireless links. The advent of this theory has led to extensive work on QoS-driven resource

allocation schemes, such as power and rate allocation [11], [12], [13] and link scheduling [14], [15], to name a few. In 2010, a multi-hop mathematical model was developed based on the effective capacity theory [16]. However, such a model was developed based on continuous-time wireless links, i.e., sampling intervals are infinitesimally small. Although this assumption may arguably be valid for high-speed transmission, it may not hold well in WSNs when sampling intervals are relatively long due to energy constraints.

The work carried out in this paper continues the work of the multi-hop effective capacity model in [16] and the discrete-time effective capacity model in [17]. To the best of our knowledge, it is the first time in the literature to consider the multi-hop effective capacity model in the discrete-time domain. New mathematical formulae of tail probabilities of delay, delay mean and jitter are developed. We further propose a sampling method for predicting end-to-end delay and test the method in a cross-layer simulation platform. Moreover, the simulation settings of each layer is based on the description in [18] and real traces from [19].

The rest of the paper is structured as follows: Section II explains the system model. Section III gives the analysis of delay performance in single-hop and multi-hop effective capacity models in the discrete-time domain. The sampling method is proposed in Section III. The evaluation of real traces and simulation results are presented in Section V. Section VI concludes the paper.

II. SYSTEM MODEL

An example for the system model with a three-hop routing path is shown in Fig. 1. The first node, Node 1, is termed the source node and the last node is the sink, which aggregates all data generated from sensors in the whole network. The systems inside wireless sensor nodes are identical to the one presented in the box of Fig. 1. Packets from the previous node are buffered in a single queue and are fairly served based on the first-in first-out (FIFO) discipline, and the buffer size in each node is assumed to be infinite. Such a system model is called series or tandem queues in the classic queueing theory [8]. We further assume that

- 1) adjacent nodes use specific time slots to transmit and receive signals so the channel collision is eliminated;
- 2) packet sizes generated by the source are the same;
- 3) every sensor node is sampled by a common sampling interval, T_s , making each node a discrete-time system.

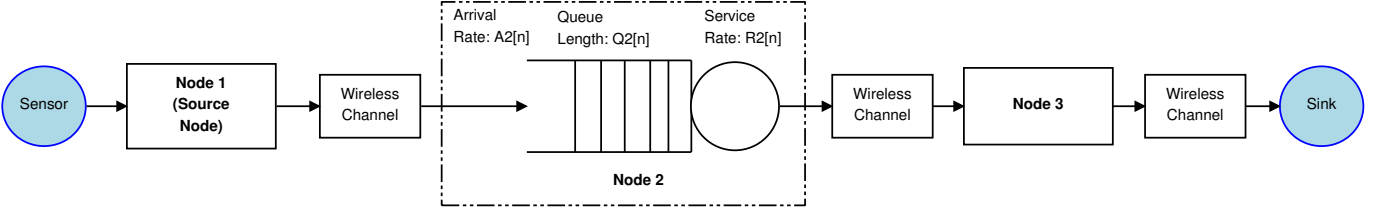


Fig. 1: Three-hop System Model in a Wireless Sensor Network

Fig. 1 also shows a snapshot of a sensor node (Node 2) at slot n . $A_2[n]$ denotes the number of arrival packets (the subscript “2” indicates the node is the second one from the source node), $R_2[n]$ denotes the number of packet(s) that the server is capable of serving, and $Q_2[n]$ is the queue length (number of packets in the queue). Finally, our interest in this paper is to understand the static behavior of end-to-end packet delays (packets generated at Node 1 and destined to the sink node).

III. MULTI-HOP DELAY PERFORMANCE ANALYSIS BASED ON THE EFFECTIVE CAPACITY THEORY

Since single-hop scenarios are special cases of multi-hop scenarios, the single-hop delay performance analysis is explained in Section III-A, followed by the multi-hop delay performance analysis in Section III-B.

A. Single-Hop Delay Performance Analysis

In a single sensor node, let D_n denote the delay value of the n^{th} packet and is a random variable. It has been shown in [10] that under some mild conditions, the delay process $\{D_1, D_2, D_3, \dots\}$ converges in distribution to a random variable D_∞ when n goes to ∞ and eventually agrees with the stationary and ergodic process. It has been further shown in [17] that for discrete-time system models, D_n is measured in slots and can only take non-negative integers, i.e.,

$$D_n \in \{0, 1, 2, 3, \dots\}. \quad (1)$$

The tail probability and probability mass function (PMF) of D_∞ in such a model may be approximated by

$$P(D_\infty > D_{\max}) \approx \gamma(1 - p^{D_{\max}}) \quad (D_{\max} \in \{0, 1, 2, \dots\}) \quad (2)$$

and

$$p(x) = P(D_\infty = x) = P(D_\infty > x) - P(D_\infty > x - 1) \approx \begin{cases} 1 - \gamma, & x = 0 \\ \gamma(1 - p^{x-1})p, & x > 0, \end{cases} \quad (3)$$

respectively. In [17], γ is termed the non-zero delay probability because the following equation holds:

$$\gamma = P(D_\infty > 0); \quad (4)$$

p is termed the success probability of a connection. Suppose the service time is negligible, the following expression shows a simple relation between γ and p via $E[Q[n]]$ and $E[A[n]]$ ($E[\cdot]$ is the expectation operator of an event $\{\cdot\}$):

$$E[Q[n]] = \frac{\gamma}{p} E[A[n]]. \quad (5)$$

(5) is the direct result from Little’s law [20].

B. Multi-Hop Delay Performance Analysis

In an H -hop scenario, after we order sensor nodes by the sequence in which a packet traverses from the source to the sink, and number them from 1 to H , the PMF of D_∞ at i -th ($1 \leq i \leq H$) node, D_i , is expressed as follows by using (3).

$$p_i(x) = P(D_i = x) \approx \begin{cases} 1 - \gamma, & x = 0 \\ \gamma(1 - p_i^{x-1})p_i, & x > 0 \end{cases} \quad (6)$$

In a tandem network, for a specific packet D_k (e.g., k -th packet), the delays that it experienced in different nodes are usually correlated [8]. To simplify our calculation, we follow the method in [16] to assume $D_k^1, D_k^2, \dots, D_k^H$ are independent. Then we have the following proposition:

Proposition 3.1: The H -hop tail probabilities of end-to-end delay can be approximated by

$$P\left(\sum_{i=1}^H D_i > D_{\max}\right) = 1 - \sum_{j=0}^{D_{\max}} p_1(x) * p_2(x) \cdots p_H(x) \quad (7)$$

where “*” stands for convolution.

The proposition below introduces a discrete-form equation that speeds up the numerical computation.

Proposition 3.2: when $\forall i, j \in \{1, 2, \dots, H\}$, if $i \neq j$, then $p_i \neq p_j$, (7) is equivalent to

$$P\left(\sum_{i=1}^H D_i > D_{\max}\right) = \sum_{i=1}^H \prod_{j=1, j \neq i}^H \left(1 + \frac{\gamma_j p_i}{p_i - p_j}\right) \gamma_i (1 - p_i)^{D_{\max}}. \quad (8)$$

The Propositions 3.1 and 3.2 can be proved by the mathematical induction and their proofs are similar to [16, Propositions 3.2 and 3.3]. However, the detailed proof is omitted for lack of space. Note that rare situations may cause the problem of $p_i = p_j$ when $i \neq j$ that is a singular point of (8), but we can use a trick of slightly changing either the value of p_i or the value of p_j to mitigate this issue. Finally, delay mean and

jitter are expressed in the following formulas:

$$\begin{aligned} \text{delay mean} &= \sum_{i=1}^H E[D_i] = \sum_{i=1}^H \frac{\gamma_i}{p_i} \\ \text{jitter} &= \sqrt{\sum_{i=1}^H E[D_i - \frac{\gamma_i}{p_i}]^2} \\ &= \sqrt{\sum_{i=1}^H \left(\frac{2(1-p_i)\gamma_i}{(p_i)^2} + \frac{\gamma_i}{p_i} - \left(\frac{\gamma_i}{p_i} \right)^2 \right)} \end{aligned} \quad (9)$$

For a proof of (9) and (10), see Appendix A.

IV. SAMPLING METHOD FOR PREDICTING END-TO-END DELAY PERFORMANCE

According to (8), (9) and (10), we need the $\{\gamma, \theta\}$ information of each node along the routing path in order to predict the end-to-end queueing delay. In a distributed network, two steps are involved, these are:

- 1) Sensor nodes extract/estimate their own information of $\{\gamma, \theta\}$ by themselves;
- 2) Sensor nodes share this information to either the sink node or other nodes that need such information by broadcasting these two values.

Step 2 is related to the standard broadcast implementation and is outside the scope of work in this paper.

The sampling method is a common technique in digital signal processing for data recovery at the receiver side. In our work, the basic idea is to use a small number of samples (the first K slots when the system starts) to represent/predict the end-to-end delay performance in a routing path. On the basis of (8), the $\{\gamma, p\}$ information from each node are the first quantities we need and these two values in the i -th node can be computed as follows:

$$\hat{\gamma}_i = 1 - \frac{\sum_{n=1}^K S_i[n]}{\sum_{n=1}^K A_i[n]} \quad (11)$$

and

$$\hat{p}_i = \hat{\gamma} \frac{\sum_{n=1}^K A_i[n]}{\sum_{n=1}^K Q_i[n]}, \quad (12)$$

where (12) can be derived from (5), $S[n]$ stands for the number of packets that arrived at slot n and have zero delays and its value is calculated by the simple Algorithm 1 shown below:

Algorithm 1 Obtaining S_n at slot n

- 1: X = sum packets of input at previous slot
 - 2: Y = sum packets of output at current slot
 - 3: **if** Y > X **then**
 - 4: S = Y - X
 - 5: **end if**
-

Finally, by following (8), (9) and (10), the tail probability of end-to-end delay, delay mean and jitter in an H -hop routing

path can be predicted by

$$P\left(\sum_{i=1}^H D_i > D_{\max}\right) = \sum_{i=1}^H \prod_{j=1, j \neq i}^H \left(1 + \frac{\hat{\gamma}_j \hat{p}_i}{\hat{p}_i - \hat{p}_j}\right) \hat{\gamma}_i (1 - \hat{p}_i)^{D_{\max}} \quad (13)$$

$$\text{predicted delay mean} = \sum_{i=1}^H \frac{\hat{\gamma}_i}{\hat{p}_i} \quad (14)$$

and

$$\text{predicted jitter} = \sqrt{\sum_{i=1}^H \left(\frac{2(1 - \hat{p}_i) \hat{\gamma}_i}{\hat{p}_i^2} + \frac{\hat{\gamma}_i}{\hat{p}_i} - \left(\frac{\hat{\gamma}_i}{\hat{p}_i} \right)^2 \right)} \quad (15)$$

It worth noting that in distributed networks, sharing the $\{\gamma, \theta\}$ information of one node with other nodes along a routing path is a multicast problem but such a topic is outside the scope of work in this paper.

V. EVALUATION AND RESULTS

In this section, we first investigate the wireless environment from publicly available real traces [19] and the PHY-layer behavior based on [18] in Section V-A. In Section V-B, we develop a simulation platform based on these behaviors above and a MAC layer. Our prediction algorithm (Section IV) is implemented in the simulation platform and its performance is presented in Section V-C.

A. Analysis of Real Traces

In [19], there are three types of real traces from three deployments: 1) an operational road tunnel; 2) a non-operational tunnel; 3) a vineyard. We chose real traces from the non-operational tunnel because the wireless channel condition was not affected by either the traffic as in case (1), or the weather as in case (2), resulting in relatively clean data. Further information regarding the geographical deployment, PHY- and MAC-layer behavior are summarized as follow:

1) *Deployment Scenario*: The network, which is depicted in Fig. 2 and used in the experiments [18], includes 20 battery-powered WSN nodes. Nodes were placed along two parallel lines on opposite walls and skewed so that a node on one wall is never directly opposite another node, minimizing the interference between nodes.

2) *Physical Layer (PHY)*: In the experiment, sensor nodes used the IEEE 802.15.4 standard, providing a nominal bandwidth of 250 Kbit/s. Packet Delivery Ratio (PDR) between different pairs i, j were recorded. PDR is a pure physical layer parameter, because

- 1) Each sensor is allocated a specific time slot for transmission to avoid collision, and
- 2) there is no re-transmission mechanism used to ensure the correctness of packets at the receiver.

For the entire duration of an experiment, each node broadcasts a packet every $N\delta$ seconds, where N is the number of participating nodes (20 Nodes in this case) and δ is a known time interval (δ has the same meaning as the sampling interval T_s in Section II). In [18], the authors used $\delta = 300$ ms so

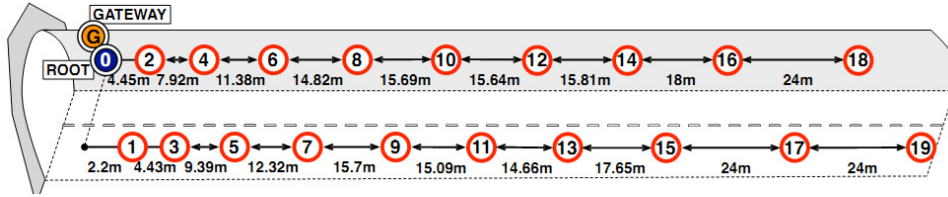


Fig. 2: Deployment Configuration in a non-operational tunnel [18]

TABLE I: Simulation Parameters

Parameter	Value
Service rate (kbps):	250
Average traffic load (kbps): μ	1.4
Packet size (Byte):	105
Hop number, H :	3
Time interval in PHY layer (ms) ($=T_s$): δ	100
Time interval in MAC layer (ms): ρ	300
PDR of the pair ③ \rightarrow ① \pm CI (%)	91.28 ± 2.96
PDR of the pair ⑤ \rightarrow ③ \pm CI (%)	69.30 ± 4.84
PDR of the pair ⑦ \rightarrow ⑤ \pm CI (%)	77.55 ± 4.38
Observation duration (used in our algorithm) (min.):	10
Total simulation time (min.):	60

transmission interval is six seconds ($20 \times 300 = 6000$ ms). Furthermore, channel coherence can be safely ignored due to this long transmission interval and it is safe to assume the packet loss probability at one transmission slot to be Bernoulli distributed with p being a specified PDR value.

B. Simulation Settings

We chose experiment 65 of [19] as the trusted source of real traces and only consider a 3-hop transmission path (⑦ \rightarrow ⑤ \rightarrow ③ \rightarrow ①).

In the MAC layer of our simulation platform, the source node constantly generates fixed-length packets at fixed interval, which is considered as Constant Bit Rate (CBR) traffic model. The packet size is 105 bytes. Two MAC-layer implementations are realized. The first one is the standard MAC-layer implementation, as shown in Fig. 3(a). Sensors are time synchronized and are programmed to transmit packets at the beginning of every time interval. The second one is the random-based MAC-layer implementation (used in [18]), as shown in Fig. 3(b). Specifically, a time interval ρ was decided before the experiment and should be integer times of δ . within ρ , a node decides randomly when to send a packet. Furthermore, we introduce a feedback mechanism that guarantees 100% reliability of communications, i.e., failed packets will be retransmitted (by receiver sending negative acknowledgment to the transmitter) until they are received successfully. The rest of simulation parameters are given in Table 1.

C. Results

The prediction results are obtained in this way: in an overall 60-min simulation run, for the first 10-min simulation, each node along the path extracts their own $\{\hat{\theta}, \hat{\gamma}\}$ (described in Section IV). We use these information to predict the delay

TABLE II: Delay Mean and Jitter

	Delay Mean (ms)	Jitter (ms)
Settings	Standard, Multi-hop and Exp 65	
Simulation/Prediction	474.71 / 437.26	656.42 / 547.96
Settings	Random-based, Multi-hop and Exp 65	
Simulation/Prediction	846.60 / 895.30	971.43 / 780.42

performance in the rest 50-min simulation by using (13), (14) and (15). Furthermore, since delay values are usually measured in time units, i.e., *seconds*, rather than *slots*. A simple translation is first performed: n slots equals to $n\delta (= nT_s)$ seconds.

Fig. 4 shows simulation and prediction results of tail probabilities under two different MAC-layer implementations. The X-coordinates are Delay Bounds (the unit is milli seconds), and the Y-coordinates are DBVP. The simulation results are shown in points with error bars (the length of error bar denotes the 99% confidence interval), while the prediction results are from (13) and are shown in dash lines. Table II lists simulation and prediction results of delay mean and jitter under different conditions. Prediction results are obtained from (14) and (15). As seen from the figure and the table, all results show sufficient accuracy of our sampling method in Section IV.

VI. CONCLUSION

In this paper, we extended the existing multi-hop effective capacity model from the continuous-time domain into the discrete-time domain. Mathematical formulae including tail probabilities of delay, delay mean and jitter over multi-hop wireless paths were derived. Furthermore, we used these formulae to develop a simple algorithm for predicting end-to-end delay based on the sampling method. The algorithm was tested via extensive simulations based on real traces data; the simulation results showed that our algorithm is capable of providing satisfactory prediction of end-to-end delay performance and hence gives good insight into QoS provisioning in WSNs. The techniques discussed in this work may be easily extended to other multi-hop and relay communication systems, where end-to-end delay performance is important.

REFERENCES

- [1] "5g: A technology vision," Huawei Technologies Co. Ltd., Tech. Rep., 2013.
- [2] S. Mansour, N. Nasser, L. Karim, and A. Ali, "Wireless sensor network-based air quality monitoring system," in *Computing, Networking and Communications (ICNC), 2014 International Conference on*, Feb 2014, pp. 545–550.

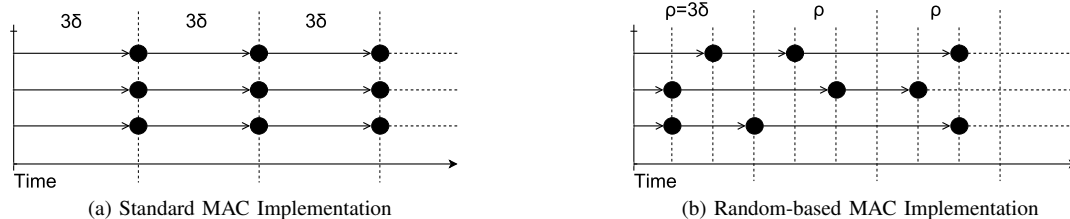


Fig. 3: Two different MAC-layer implementations

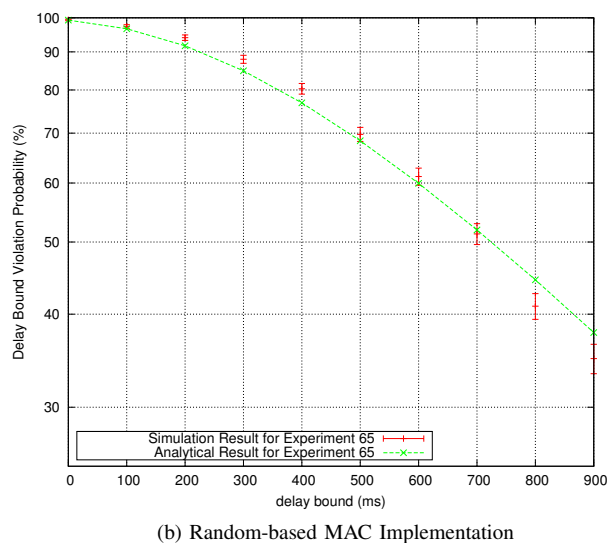
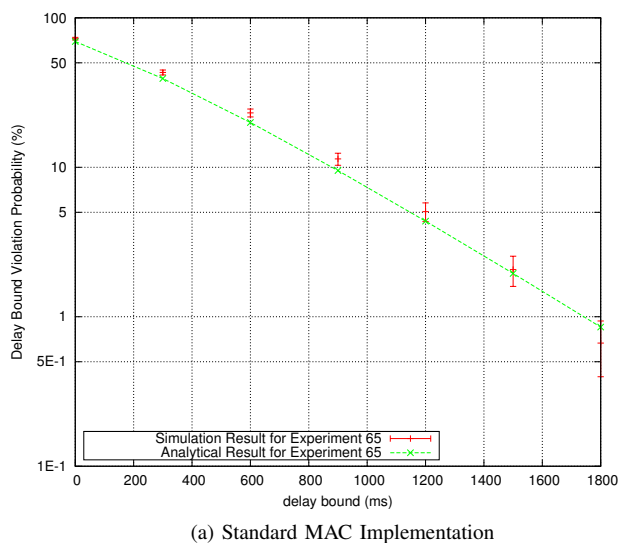


Fig. 4: Tail probabilities of delay with different MAC-layer implementations

- [3] Z. Zhang and X. Hu, "Zigbee based wireless sensor networks and their use in medical and health care domain," in *Sensing Technology (ICST), 2013 Seventh International Conference on*, Dec 2013, pp. 756–761.
- [4] T. Weisbach, A. Hurzig, T. Keutel, K. Nendel, E. Muller, and O. Kanoun, "Requirements for wireless sensors networks in production and logistic," in *Multi-Conference on Systems, Signals Devices (SSD), 2014 11th International*, Feb 2014, pp. 1–4.
- [5] A. Fallahi and E. Hossain, "A dynamic programming approach for qos-aware power management in wireless video sensor networks," *Vehicular Technology, IEEE Transactions on*, vol. 58, no. 2, pp. 843–854, Feb 2009.
- [6] T. Janevski, "5g mobile phone concept," in *Consumer Communications and Networking Conference, 2009. CCNC 2009. 6th IEEE*, Jan 2009, pp. 1–2.
- [7] S. Fang, L. Yu, J. Zheng, and Q. Zhu, "Stochastic qos performance analysis of diffserv-based wireless sensor network," in *Communications in China (ICCC), 2013 IEEE/CIC International Conference on*, Aug 2013, pp. 270–274.
- [8] D. Gross, J. F. Shortle, J. M. Thompson, and C. M. Harris, *Fundamentals of Queueing Theory*, 4th ed. New York, NY, USA: Wiley-Interscience, 2008.
- [9] D. Bertsekas and R. Gallager, *Data Networks*. Prentice-Hall, 1987.
- [10] D. Wu and R. Negi, "Effective capacity: a wireless link model for support of quality of service," *Wireless Communications, IEEE Transactions on*, vol. 2, no. 4, pp. 630 – 643, July 2003.
- [11] J. Tang and X. Zhang, "Quality-of-service driven power and rate adaptation for multichannel communications over wireless links," vol. 6, no. 12, pp. 4349–4360, 2007.
- [12] L. Musavian, S. Aissa, and S. Lambotharan, "Adaptive modulation in spectrum-sharing channels under delay quality-of-service constraints," vol. 60, no. 3, pp. 901–911, 2011.
- [13] S. Akin and M. C. Gursoy, "Effective capacity analysis of cognitive radio channels for quality of service provisioning," vol. 9, no. 11, pp. 3354–3364, 2010.
- [14] Q. Wang, D. O. Wu, and P. Fan, "Delay-constrained optimal link scheduling in wireless sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 9, pp. 4564–4577, Nov. 2010.
- [15] D. Wu and R. Negi, "Downlink scheduling in a cellular network for quality-of-service assurance," vol. 53, no. 5, pp. 1547–1557, 2004.
- [16] Y. Chen, Y. Yang, and I. Darwazeh, "A cross-layer analytical model of end-to-end delay performance for wireless multi-hop environments," in *Proc. IEEE Global Telecommunications Conf. GLOBECOM 2010*, 2010, pp. 1–6.
- [17] Y. Chen and I. Darwazeh, "An estimator for delay distributions in packet-based wireless digital communication systems," in *Wireless Communications and Networking Conference (WCNC)*, 2013.
- [18] L. Mottola, G. P. Picco, M. Ceriotti, S. Guna, and A. L. Murphy, "Not all wireless sensor networks are created equal: A comparative study on tunnels," *TOSN*, vol. 7, no. 2, 2010.
- [19] (2010) Publicly available real connectivity traces from tunnels and vineyards. <http://d3s.disi.unitn.it/tunnelvineyard>.
- [20] J. D. C. Little, "A proof of the queuing formula: $L=\lambda W$," *Operations Research*, vol. 9 (3), p. 383387, 1961.

APPENDIX A
PROOF OF (9) AND (10)

In an i -th node, the probability generating function (PGF) of D_i is given by

$$\begin{aligned}\Pi_{D_i}(s) &= E[s^{D_i}] = \sum_{k=0}^{\infty} s^k p_i(k) \\ &= -\gamma_i + \frac{\gamma_i p_i}{1-p_i} \frac{1}{1-(1-p_i)s}\end{aligned}\quad (16)$$

According to the attributes of PGF, the delay mean and jitter of D_i are given

$$\text{delay mean} = \Pi'_{D_i}(1) = \frac{\gamma_i}{p_i}\quad (17)$$

$$\begin{aligned}\text{jitter} &= \sqrt{\Pi''_{D_i}(1) + \frac{\gamma_i}{p_i} - \frac{\gamma_i^2}{p_i^2}} \\ &= \sqrt{\left(\frac{2(1-p_i)\gamma_i}{(p_i)^2} + \frac{\gamma_i}{p_i} - \left(\frac{\gamma_i}{p_i}\right)^2\right)}\end{aligned}\quad (18)$$

For an H -hop routing path, since $D_\infty^1, D_\infty^2 \dots D_\infty^H$ are assumed to be independent in Section III-B, we have (9) and (10) for end-to-end delay mean and jitter.