Abstract—In this paper, we consider the problem of modelling the average delay experienced by an application packets of variable length in a single cell IEEE 802.11 DCF wireless local area network. The packet arrival process at each node is assumed to be a stationary and independent increment random process with mean $a_i$ and second moment $a_i^2$. The packet lengths at node $i$ are assumed to be i.i.d random variables $l_i$ with finite mean and second moment. A closed form expression has been derived for the same. We assume the input arrival process across queues to be uncorrelated Poison processes. As the nodes share a single channel, they have to contend with one another for a successful transmission. The mean delay for a packet has been approximated by modelling the system as a 1-limited Random Polling system with zero switchover times. Extensive simulations are conducted to verify the analytical results.

Index Terms—Delay Modelling ; Single Cell WLAN ; Random Polling Systems ; Variable Packet Length

I. INTRODUCTION AND RELATED WORK

The IEEE 802.11 has become ubiquitous and gained widespread popularity as a protocol for wireless networks. As a result, various models have been proposed to analyze and model the parameters of interest.

Since the seminal paper by Bianchi [1], throughput analysis of IEEE 802.11 DCF has come under much scrutiny. In [1], the main feature of the analysis is the 2-dimensional Markov model, which captures the back-off phenomenon of IEEE 802.11, given a transmission attempt rate for each node. In [2], the authors give an analytical model for throughput analysis of DCF using average back-off state as compared to the Markovian model being proposed by Bianchi. In [3], the authors study the fixed point solution and performance measure in a more generalized framework.

Delay analysis of IEEE 802.11 DCF is limited in comparison to the throughput studies. In [4], the authors propose System Centric and User Centric Queuing Models for IEEE 802.11 based Wireless LANs. In the System Centric Model, the arrivals are assumed to be Poisson, thus the resource sharing model takes the form of an M/G/1/PS system with the mean delay being the same as that in an equivalent M/M/1 system. In the User Centric Model, each user queue is modeled as a separate G/G/1 queue.

In [5], the authors provide an analysis of the coupled queue process by studying a lower dimensional process and by introducing a certain conditional independence approximation. The authors in [5], provide an analytical framework to model the delay only for the case of homogeneous Poisson arrivals. In [6], we have analyzed the mean delay for single hop wireless mesh networks under light aggregate traffic by introducing a decoupling assumption. But as load increases, interactions between the queues appear and our modelling assumption ceases to be valid. In [7], the authors analyze delay under homogeneous arrivals assuming packet lengths to be i.i.d across all the queues. In our previous work [8], we have addressed delay modelling for nonhomogeneous Poisson arrivals under the assumption of fixed packet length.

In this paper, we model the system as a 1-limited random polling system with zero switchover times. We provide a simple model to obtain the mean delay for the variable length packet arrival process at node $i$. This enables us to use the mean delay expressions from [9] to analyze the delay in a single cell wireless local area network. We remark that the user traffic delay is not merely the Head-Of-Line (HOL) packet delay that has been analyzed in [1], [2] and [3]; it includes the delay from the time a user packet arrives at the queue, till the packet reaches the destination. Thus, both queuing delay and HOL delay are included.

Our objective is to explore the use of known results for the saturated network and random polling systems in analyzing the mean delay experienced by a packet. Our main contributions can be outlined as follows,

- We propose a random polling system framework to analyze mean delay in a single cell wireless local area network with variable packet lengths.
- We obtain closed form expression using a novel approach for mean delay by applying results from [2] to our random polling system framework, for the case when the packet lengths at node $i$ are i.i.d random variables.
- We show though simulations that our random polling framework can be used to estimate the mean delay in a single cell IEEE 802.11 wireless local area network.

The rest of the paper is organized as follows. In Section II, the system model is described in detail. In Section III, we discuss the delay modelling framework in detail. In section IV, we obtain closed form expressions for application level mean delay under Poisson packet arrivals and random packet lengths. In Section IV, the proposed framework is validated against simulation results. Finally, Section V presents conclusion and remarks regarding future work.
II. SYSTEM MODEL

We consider a single cell wireless local area network with $n$ nodes and no hidden nodes. We assume that the packet arrival process at node $i$ is a stationary and independent increment random process $X_i(t)$ with mean $a_i$ and second moment $a_i^{(2)}$. The lengths of packets at node $i$, are assumed to be i.i.d random variables represented by $P_i$, having finite mean and second moment. Each node in the network shares the medium and uses the IEEE 802.11 DCF to exchange data with one another. A packet transmitted by a node is destined for any of the other nodes. Each node is assumed to have a single network output queue. All wireless links are assumed to operate at the same date rate. We assume the network layer to employ fragmentation to break down application packets that have length greater than the MTU, since the maximum size of data that can be transmitted in 802.11 at a time is bounded above by the size of the MTU.

From II, III and IV, it is known that the average aggregate rate of data transfer is dependent on the number of nodes contending. We also note that each node has equal probability of success. As in IV, we model the network as a multiple queue single shared server system, where the service rate of the server is dependent on the number of non-empty queues and the server selects a non-empty queue uniformly at random. We assume that the destination can push data out of the queue instantaneously. In this system, we are interested in quantifying the average delay between a packet’s arrival to a queue and its departure from the system.

III. ANALYSIS FRAMEWORK

A. Computing aggregate service rate in 802.11 DCF

In our model, at most $n$ nodes contend for access to the wireless medium. Let $\beta_n$ be the probability that a node attempts transmission, when $n$ nodes are contending for access to the wireless medium. From II, $\beta_n$ can be expressed in terms of the conditional collision probability $p$ in two ways as follows.

\[
\beta_n(p) = \frac{2 \cdot (1 - 2p)}{(W + 1) \cdot (1 - 2p) + pW \cdot (1 - (2p)^n)} \quad (1)
\]

\[
\beta_n(p) = 1 - (1 - p)^\frac{1}{n} \quad (2)
\]

Now, the RHS of Equation (1) is monotonically increasing from $\frac{2}{W + 1}$ to $\frac{2}{n}$, for $p \in [0, 1]$, and the RHS of Equation (2) is monotonically increasing from 0 to 1, for $p \in [0, 1]$. We can use fixed point analysis to obtain $\beta$ for a given $n$. Similar approach has been followed in [5]. Let $T_S$, $T_I$ and $T_C$ be the durations of success, idle and collision slots, respectively. Now, $T_S$ can be represented as $\frac{H + E[P]}{R}$, where $H$ and $E[P]$ denote the header and expected packet lengths respectively with $R$ being the rate of transmission. Let us define the probability of a successful transmission ($p_s^{(n)}$), collision ($p_c^{(n)}$) and idle ($p_i^{(n)}$) as

\[
p_s^{(n)} = n\beta_n \cdot (1 - \beta_n)^{n-1} ; \quad p_i^{(n)} = (1 - \beta_n)^n
\]

By applying the Renewal Reward theorem, we define the aggregate throughput as follows

\[
S(n) = \frac{p_s^{(n)} \cdot E[P]}{p_i^{(n)}T_I + p_s^{(n)}T_S + p_c^{(n)}T_C}
\]

B. Random polling system framework (RPS)

From Figure 2 it can be seen that the aggregate service rate is nearly constant. This suggests that we can imagine a constant-rate server serving the collection of queues. Accordingly, we consider the system as $n$ infinite-buffer queues being served by a single server. When the server visits a non-empty queue, it serves one packet and moves on to the next queue. This type of polling system is referred to as a \textit{Limited Random Polling System}. The service time of customers in queue $i$ is a nonnegative random variable with mean $p_i$ and second moment $p_i^{(2)}$. We define the expected offered load to the system due to queue $i$ as $\rho_i = a_i p_i$. The expected system utilization factor is defined as $\rho = \sum_{i=1}^{n} a_i p_i$. Clearly, a necessary condition for the system to reach a stable state is

\[
\rho < 1
\]

Now we restate the notations and assumptions used in [9], merely for the purpose of understanding our approach. A memoryless polling policy is assumed, such that the next queue, say queue $j$, is selected for service with probability $\gamma_j$, where $0 < \gamma_j < 1$. The service time provided (if any)
and the switchover time that follows are collectively defined as a period. The switchover time after the server visits queue \( i \) is a non-negative random variable with mean \( s_i \) and second moment \( s_i^2 \). Let \( b_i \) and \( b_i^2 \) be the first and second moment of the batch size. Let \( s = \sum_{j=1}^{\infty} s_j \gamma_j \).

Let \( Q_i(k) \) denote the queue length at node \( i \), at the beginning of the \( k \)-th period. Then we define \( E[Q_i] = \lim_{k \to \infty} E[Q_i(k)] \). From Theorem 12 of \([9]\), we can obtain a closed form expression for the average waiting time of the packets at node \( i \) as

\[
E[W_i] = \frac{E[Q_i]}{a_i \cdot P\{Q_i \geq 1\}} - \left( 1 - \frac{\rho s_i}{a_i} \right) - \frac{b_i^2 - b_i}{a_i b_i}
\]

(3)

where \( \chi_i = 1 - \frac{sa_i}{\gamma_i} - \frac{\rho sa_i}{a_i} \)

\[
\nabla_{ij} = a_i a_j \sum_{j=1}^{\infty} \left[ \gamma_j s_j (p_j^2) + (p_j^2) + 2 s_j p_j \right] \frac{a_j s}{(1 - \rho)}
\]

(5)

\[
\psi_i = \sum_{i=1}^{\infty} \frac{a_i}{2 \gamma_i} + \sum_{i=1}^{\infty} \frac{p_i \nabla_{il}}{(1 - \rho)}
\]

The steady state probability of a queue being non-empty is given by

\[
P\{Q_i \geq 1\} = \frac{sa_i}{\gamma_i (1 - \rho)}
\]

(6)

The above analysis applies only for a system with non-zero switch over times.

**C. Application of RPS to a Single Cell 802.11 WLAN**

According to our modelling assumptions, if the server visits a node with a nonempty queue, it will serve exactly one packet and will move onto the next queue in zero switchover times. Else, if the server visits a node with an empty queue, we assume that it will move onto the next queue with zero switchover times. As 802.11 DCF follows fair server allocation policy, we can say that \( \{\gamma_i = \frac{1}{n}, 1 \leq i \leq n\} \). This emulates the process of nonempty queues succeeding with equal probability. So we are interested in simplifying Equation (3) to reflect the case of zero length switchover period.

In a different context, the author in \([10]\) has proposed that the expression for mean delay with zero switchover times can be obtained from the expression for non-zero switchover times by proper application of limits to the distribution of the switchover times. Motivated by this, we follow a similar approach to arrive at the expression for zero switchover times by defining the switchover times after servicing queue \( i \) (i.e. \( s_i \)) as a constant \( \epsilon \). The expression for average delay when switch over times are zero (\( E[W_i^{(0)}] \)) can be obtained by letting \( \epsilon \) go to 0. Thus we have

\[
E[W_i^{(0)}] = \frac{\rho}{a_i} - \frac{1}{a_i} + \lim_{\epsilon \to 0} \frac{E[Q_i]}{a_i \cdot P\{Q_i \geq 1\}} - \frac{b_i^2 - b_i}{2 a_i b_i}
\]

(7)

Since we have uncorrelated arrival processes, we have \( \{e_{ij} = 0, \forall i \neq j\} \). Now, after evaluating the limit in Equation (7) and through some algebraic manipulation, we obtain

\[
E[W_i^{(0)}] = \sum_{i=1}^{\infty} \left( \frac{p_i^2 - p_i^2 a_i}{2(1 - \rho)} \right) \frac{\rho}{a_i} + \frac{1 + e_{ii}}{a_i} \frac{b_i + b_i^2}{a_i b_i}
\]

(8)

**D. Computation of \( e_{ii} \) and \( a_i \) for a Batch Poisson process**

The expression in \([9]\) is valid only for class of processes that have the stationary and independent increment property. Since the batch Poisson process is a member of the same class, we use batch Poisson process to model variable packet lengths. Therefore, in order to model the delay for a packet of variable length, we consider as a batch of packets that arrive at the node at the same time.

We invoke the linear quadratic model of uncertainty from \([9]\) to compute \( e_{ii} \) and \( a_i \). The motivation and the example for the same have been well described in the same paper. The model is as follows,

\[
E[A_i(t)] = a_i t
\]

(9)

\[
E[A_i(t)^2] = c_{ii} t + a_i^2 t^2
\]

(10)

where \( A_i(t) \) is the number of packets that arrive at node \( i \) during the time interval \([0, t]\). Since the inter-arrival times between the batches are exponentially distributed random variables and the batch size is independent of the inter-arrival times, hence \( A_i(t) \) satisfies stationary independent increment property. Let \( N_i(t) \) be the number of batches that have arrived during the time interval \([0, t]\) at node \( i \). Let \( B_{ij} \) be the random variable representing the batch size of the \( j \)-th batch of this arrival process with mean \( b_i \) and second moment \( b_i^2 \). Let the batch arrival process at node \( i \) be Poisson with rate \( \lambda_i \). Then, the moment generating function (MGF) for the batch Poisson Process is

\[
M_{A_i(t)}(z) = E[z^{A_i(t)}] = E[z^{\sum_{j=i}^{N_i(t)} B_{ij}}]
\]

The random variables \( \{B_{ij}, j \geq 0\} \) are i.i.d. Using the properties of moment generating function, we obtain

\[
a_i = \lambda_i M_{B_{i}(1)}(1) = \lambda_i b_i
\]

(11)

\[
e_{ii} = \lambda_i M_{B_{i}(1)}(1) = \lambda_i b_i^2
\]

(12)

**IV. APPLICATION DELAY MODELLING**

In this section we apply the framework laid in the previous section to arrive at closed form expression for the mean delay of an application packet. In our analysis, we allow the application packet length to vary from zero to any arbitrary length. But in practice, packets whose lengths are larger than
the Maximum Transmission Unit (MTU), are split and sent as different entities. We are interested in the overall delay of a packet, which may compose of different smaller MTUs. Thus we quantify the mean packet delay, by analyzing two separate cases.

A. Application packet size is bounded above by the MTU

Let us assume that the MTU has a length of \( P \) bits. All other packet lengths are expressed in terms of the MTU. At node \( i \), we define random variable \( \Omega_i = \frac{\lambda_i}{C} \) with mean \( \omega_i \) and second moment \( \omega_i^2 \). We would like to emphasis the fact that the random variable \( \Omega_i \) will be a fraction.

In order to obtain the expression for delay when the packet size is bounded above by the size of the MTU, we set

\[
p_i = \frac{\omega_i P}{C} \quad \text{and} \quad p_i^2 = \frac{\omega_i^2 (P^2)}{C^2}
\]

Since the packet size equals a MTU at most, we also set the batch size to a a fixed value 1, thus we get

\[
b_i = 1 \quad \text{and} \quad b_i^2 = 1
\]

Setting the above variables in Equation (8), we get the expression for average delay as

\[
d_{\text{avg}} = \frac{\sum_{i=1}^{n} (\omega_i^2 - \omega_i^2) \lambda_i}{2 (\frac{C}{H})^2 (1 - \rho)} + \frac{\omega_i}{2 (\frac{C}{H})} \left( \frac{2 - \rho}{1 - \rho} \right)
\]

B. Application packet size is bounded below by the MTU

As in the previous section, we assume the unit packet size to be \( P \) bits and that every other packet length is expressed in terms of this unit packet. In this section, we take an alternate approach as compared to the one in the previous section. We model the large packet as an aggregation of smaller units of packet, each of which has size lesser than or equal to \( P \) bits. For a node \( i \), we model this aggregated packet as a batch Poisson process, whose batch size has mean \( \omega_i \) and second moment \( \omega_i^2 \); we do not restrict the batch size to be discrete. As the packet size is greater than \( MTU \), the packet is fragmented at the source and reassembled at the destination. We note that the packet is constructed only after the reception of its fragments. Or on other words, we can say that the mean packet delay is equal to the mean delay of the last fragment. In this section, we are interested in quantifying the total delay of all the fragments of a packet.

Since the packet size is greater than the \( MTU \), the packet is fragmented at the source. Thus, this effect had to be factorized into the computation of aggregate throughput (\( C \) bits/s). Due to lack of space, we are not presenting the modified expression in this paper. From the arguments in the section \( III-A \), we assume that an \( MTU \) at node \( i \) is serviced at a constant time of \( \frac{p_i}{C} \) sec.1 By adopting the throughput computation by \( II \), we have abstracted the idle and the collision slots into the constant service time for a packet of node \( i \), given by \( \frac{p_i}{C} \) sec. We model our system using the RPS framework by taking a server which serves at fixed rate. Thus,

\[
\forall i, \quad p_i = \frac{P}{C} \quad \text{and} \quad p_i^2 = \left( \frac{P}{C} \right)^2
\]

By substituting the various values in (8), as in the previous section, we get the mean delay for a fragment of the packet as

\[
d_i = \frac{P}{2C} \left( 1 + \frac{\omega_i^2}{\omega_i (1 - \rho)} \right)
\]

This fragment is held at the network buffer at the destination node, until the remaining fragment arrive. Now the we approximate the average cycle time of the server at node \( i \) as

\[
w_i = d_i - \frac{P}{C}
\]

It can be shown that the mean packet delay is related to the mean fragment delay and the mean batch size as follows,

\[
d_{\text{avg}} = \left( \frac{\omega_i + 1}{2} \right) d_i + \left( \frac{\omega_i - 1}{2} \right) \frac{P}{C}
\]

After substituting the value of \( d_i \), we obtain the average delay per packet as

\[
d_{\text{avg}} = \frac{P}{4C} \cdot \left( 3 - \omega_i + \frac{\omega_i^2 (1 + \omega_i)}{\omega_i (1 - \rho)} \right)
\]

V. VALIDATION

In this section, we show tables which compare the analytical against the simulated values of mean delay. The simulation was done using Quanet 4.5, which is a discrete event simulation system. In order to obtain accurate estimates of the mean delay, the simulation was run long enough so that the average delay at the destinations are within \( 1 \mu s \) interval. This process was repeated for 30 simulation. In the tables, one packet is equivalent to 1500 \( \text{Bytes} \) (i.e \( MTU = 1500 \text{Bytes} \)). Data Rate is 1 \( \text{Mbps} \) and 4 nodes contend for the WLAN medium.

| \( \lambda_1 \) | \( \lambda_2 \) | \( \lambda_3 \) | \( \lambda_4 \) | \( d^{(1)}_{\text{avg}} \) | \( d^{(2)}_{\text{avg}} \) | \( d^{(3)}_{\text{avg}} \) | \( d^{(4)}_{\text{avg}} \) | \( d_{\text{avg}} \) |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 10.0   | 10.0   | 10.0   | 10.0   | 15.0   | 15.0   | 14.9   | 14.9   | 14.9   |
| 2.0    | 2.0    | 3.0    | 3.0    | 3.45   | 4.44   | 4.44   | 4.44   | 4.44   |

1We assume that fragments having size lesser than MTU are also transmitted at same rate, thus our expression is in-fact an upper bound for the mean delay experienced by a packet. Simulations validate our assumption by showing that the bound is in-fact tight.

In Tables I, II, III and IV, we compares the analytical and simulation result for various arrival rate and type of distribution for the packet size for the case when the packets size is less than the \( MTU \). We note that for lower aggregate rate, the mean delay across the queues are invariant and matches with the analytical results. For higher load, our analysis is able to approximate the maximum delay experienced by a packet in the network.
In Tables II, VI, VII and VIII, we compare the analytical and simulation result for various arrival rate and type of distribution for the packet size for the case when the packets size is greater than the MTU. We note that, for the case of homogeneous arrival, the analysis is matching with the simulation result under high loads as well.

### VI. CONCLUSION

In this paper, we have abstracted the mechanism of IEEE 802.11 MAC enabling us to find simple closed form expression for application level packet delay in a single cell IEEE 802.11 wireless area network. Our analysis enables to approximate the application level mean delay for variable packet length using a closed form expressions. For low loads, the delay across the queues is invariant and is closely approximated by the analytical formula. It can be seen that, even for the scenario of nonhomogeneous packet arrivals, the analytical delays is closest to the mean delay of a packet in the queue with the highest data rate. Simulations indicate that the proposed framework is able to model the maximum mean delay in single hop wireless mesh network with reasonable accuracy, provided the system operates within the capacity region.

### REFERENCES

[1] G. Bianchi, “Performance analysis of the IEEE 802.11 distributed coordination function,” in JSAC Wireless Series, vol. 18, no. 3, 2000.

[2] Y. Tay and K. Chua, “A capacity analysis for the IEEE 802.11 MAC protocol,” in Wireless Networks, vol. 7, no. 2, pp. 159-171, March 2001.

[3] A. Kumar, E. Altman, D. Miorandi, and M. Goyal, “New insights to the analysis of single cell IEEE 802.11 WLANs,” in Proc. of 22nd International Conference on Queueing Networks and their Applications, 2015.

[4] K. Medepalli and F.A. Tobagi, “System centric and user centric queueing models for IEEE 802.11 based wireless LANs,” in Proc. of IEEE 1st International Conference on Wireless Networks, vol. 7, no. 2, pp. 159-171, June 2005.

[5] Manoj K. Panda and Anurag Kumar, “State dependent attempt rate modeling of single cell IEEE 802.11 WLANs with homogeneous nodes and poisson arrivals,” in Proceedings of IEEE COMSNET, 2009.

[6] Albert Sunny, Joy Kuri, and Saurabh Aggarwal, “Delay modelling for a single-hop wireless mesh network under light aggregate traffic,” online version available at http://www.arxiv.org/abs/1009.3468.

[7] M. Tickoo and B. Sikdar, “Queueing analysis and delay mitigation in IEEE802.11 random access mac based wireless networks,” in Proc. of IEEE INFOCOM, 2004.

[8] Albert Sunny, Joy Kuri, and Saurabh Aggarwal, “Delay modelling for single cell IEEE 802.11 wlan using a random polling system,” online version available at http://www.arxiv.org/abs/1009.3468.

[9] Thomas (Yew Sing) Lee, “A closed form solution for the asymmetric random polling system with correlated levy input process,” in Mathematics of Operations Research, Vol. 22, No. 2 (May, 1997), pp. 432-457.

[10] Leonard Kleinrock and Hanoch Levy, “The analysis of random polling systems,” in Operations Research, Vol. 36, No. 5 (Sep - Oct) , pp. 716-732, 1988.