Performance of Distributed Coordination Function for Serving IoT Applications

Ruki Harwahyu¹, Ray-Guang Cheng², Riri Fitri Sari¹

¹ Department of Electrical Engineering, Faculty of Engineering, Universitas Indonesia, Indonesia.
² Department of Electronic and Computer Engineering, National Taiwan University of Science and Technology, Taiwan.

Abstract. The Internet of Things (IoT) interconnects “things” and autonomously exchanges data between them. It is expected that there are more than seven billion of connections interconnected by low-power wide-area (LPWA) networks to support IoT services before 2025. Several LPWA technologies are gaining popularity recently, such as LoRa and NB-IoT. Despite of their advertised advantages over conventional technologies such as low power, long range, and cheap module, several studies and field tests shows that they cannot communicate in long-range, the DIY supporting modules are not really power efficient, and the development are rather expensive. While commercial IoT networks have yet to be widely deployed, many smart connected devices use WiFi-based connectivity. This motivates us to study and evaluate the performance of WiFi’s medium access protocol for this new use case of IoT. In the experiment, our proposed analytical model is evaluated against a custom-made Monte Carlo simulation. The result shows that our model is accurate in estimating the system performance in term of normalized throughput.

1. Introduction
The Internet of Things (IoT) is starting to be used in more and more cases. It typically consists of autonomous and simple “things” which can exchanges data with the respective network aggregator and server. The “things: typically take form of sensors in various machines/appliances or wearables. This connected architecture allows us to gather many digital representations, i.e. data, out of physical phenomenon. This capability of IoT is in fact one of the reason behind the advancement of big data analysis in many fields such as healthcare, environment monitoring, security, utility metering, asset management, transportation, farming, etc. IoT has been envisioned to be the major drive for industrial 4.0. It is also anticipated to be the dominant source for data traffic over the internet by 2025.

The emerging IoT technology yields new potential market of telecommunication network services. Following the trend of IoT nowadays, the wireless technology to serve IoT devices is called low-power wide-area (LPWA) networks (LPWAN). There are LPWAN technologies which uses licensed and unlicensed spectrums. Those which use licensed spectrum are called cellular IoT (CIoT). Meanwhile those which uses unlicensed spectrum are more common to be found for fast adoption or prototyping, such as LoRa/LoRaWAN, SigFox, and RPMA. These unlicensed LPWAN usually employ 8-900 MHz industrial-scientific-medical (ISM) bands. Both of the LPWAN technologies are differ from the existing wireless technology as they focus to accomplish the following key requirements of IoT connectivity [1]:
1. Low end-device cost
2. Support for a massive number of devices
3. Long battery life
4. Extended coverage

To meet the first requirement from the above, LPWAN technology has lower complexity compared to the traditional wireless systems. The simplification typically includes the simplified communication protocol and smaller header and trailer for payload encapsulation. To fulfill the second requirement, LPWAN technologies have to include re-engineered medium access control (MAC) protocol. Meanwhile, the third and the fourth requirements seem to be contradicting each other. That is why they are the most challenging and fundamental characteristics of LPWAN technology. LPWAN needs to have better link budget (i.e. extended transmission coverage) while maintaining low energy consumption. For this aspect, new modulation and Layer 1 techniques are incorporated, such as proprietary modulation in LoRa and SigFox \cite{2} and repetition techniques in NB-IoT \cite{3}.

Regardless of the promoted and promised features of new radio technologies for connecting IoT devices, the readiness of these technologies to be adopted at the current stage is still not mature enough. For example, several studies and field tests of LoRa cannot realize the advertised long-range transmission \cite{4}. Meanwhile, in the current period when IoT is beginning to implemented, most of the IoT devices are for smart infrastructure (e.g. smart home, smart building, etc.) which typically has no issue of battery power consumption, resides within coverage of existing network of cellular or WiFi (which may significantly decrease the cost of deployment and operation), and does not bear stringent timing requirements. While commercial IoT networks have yet to be widely deployed, many smart connected devices use WiFi-based connectivity \cite{4}. This motivates us to study and evaluate the performance of WiFi’s medium access protocol for this new use case of IoT.

The distributed coordination function (DCF) principle used in WiFi is studied herein. Several related literatures are discussed in Section 2. The standard mechanism of DCF is elaborated in Section 3. An analytical model is constructed in Section 4, which will be used as the main tool in the experiment. Our analytical model considers a more realistic timing and slot duration which in practice can be vary depending on channel condition. We consider this as our strong contribution since the existing literatures normally only consider a simplified model with fixed slot duration \cite{3}. A custom-made DCF simulator is also prepared to crosscheck the correctness of our analytical model. The analytical model is then used for estimating the system performance in term of access success probability. Section 5 describes our evaluation and discusses the results. Section 6 delivers the conclusion of this study.

2. Distributed Coordination Function (DCF)

DCF is the MAC layer technique which is popularized by IEEE 802.11 protocol of WiFi. In this technique, different devices can use the same radio frequency band without prior synchronization. This technique mandates the devices to detect activity in the channel prior to its transmission to decrease the probability of the transmission being collided with other’s transmission. This method is called Carrier Sense Multiple Access (CSMA). Additional inter-frame interval is used to further ensure that there is no transmission from other devices in the vicinity (which is affected by receiver’s sensitivity threshold). This method is called CSMA with Collision Avoidance (CSMA/CA).

In wireless channel, CSMA/CA is preferred over CSMA/CD since the transmitting devices (i.e. those who are involved in a collision) cannot listen to the channel while doing the transmission themself. In WiFi, carrier sensing is done mainly by detecting the energy in the common channel, and is backed up with virtual sensing with the help of network allocation vector (NAV) to minimize the chance of hidden terminal problem to occur.
When required, this asynchronous operation can be paused temporarily to favour some devices (if any) which requires a more reliable collision-free transmission. In WiFi, this mode of operation can be activated with technique called Point Control Function (PCF), which operates based on polling from the gateway/access point.

In DCF, the CSMA/CA is used with a binary exponential backoff over slotted time. The backoff is performed by devices which found that the channel is busy when they sensed the channel prior to transmitting their backlogged packet. In a backoff, the device deferred their transmission for a certain duration. The duration is determined by the backoff timer which is randomly chosen from the range of \([0,CW]\). CW represents the contention window. The backoff timer is decremented in each elapsed slot time when the medium is idle (i.e. no transmission found in the channel) and is not decremented when the medium is busy. When the backoff timer reaches zero, the device check the channel again. If busy channel is found, backoff is conducted again. CW is doubled in the next backoff. Hence it is called binary exponential backoff.

Several studies focusing on the behaviour of DCF and CSMA/CA has been presented. In these studies, multiple models for analysing the effect of various parameters involved in the protocol are proposed. The upper bound of the throughput of p-persistent variant of 802.11 has been analysed theoretically in [5]. This literature does not consider the exponential increase of CW and the slotted nature of the system. In addition to that study, Markov process are adopted in [6,7] to analyse the throughput of 802.11 under saturated traffic load. While both of these studies employ similar basics, the model in [7] extends the model in [6] by taking into account the retransmission limit in each transmission turn.

3. Analytical Model

Following Bianchi’s popular approach in [4], the backoff duration is a discrete-time with the unit called slot. In this work, slot is indexed with \(i, 0\leq i\). Normally, a slot is equal to 10 microseconds. A user which is backing off decrements its backoff timer each time a slot elapses. However, the backoff timer must be frozen (not decremented) if the channel is busy. In this case, the length of the slot includes the duration of busy channel.

Upon the packet arrival, the node must sense idle channel for DIFS period. Otherwise (if DIFS is interrupted), it has to wait for DIFS again. Only if the channel is clear for DIFS period, the node starts counting down of its backoff counter. The count-down is frozen when the channel is busy and continued after uninterrupted clear channel for DIFS period.

This work considers a condition where all users arrive and start listening to the channel simultaneously. Each user can conduct up to \(N\) transmission attempts, where the attempt is indexed by \(n, 0\leq n\leq N\).

The elaboration will first go through the process in slot 1 and slot 2 to construct the basic probabilistic model. After that, a general model is formulated for all slots to estimate the system’s performance metrics of access success probability.

Let us analyse the condition in slot 1. At this initial condition, no one using the channel and the channel is idle for DIFS period. All nodes realize it and they do the initial backoff (backoff stage 1) with window of \(W_1\) to decide the time for their 1st transmission attempt. In this situation, \(M\) users are contending for slot 1, 2, \ldots, \(W_1\).

Let \(P_{i,k}\) be the probability that slot \(i\) chosen by \(k\) users. For this initial condition, probability that slot 1 is empty (chosen by 0 user) is

\[
P_{1,0} = \left(1 - \frac{1}{W_1}\right)^M
\]  

(1)
Let $S_i$ be the expected total successful user at slot $i$. In general for every slot $i>0$, if it is chosen by only 1 user, it yields 1 successful user. Thus, calculating the probability that slot $i$ is chosen by 1 user is similar to calculating the expected total successful user from that slot. Hence, let us replace the notation $P_{i,1}$ with $S_i$.

In general for every slot $i>0$, there can be several users conducting its 1st, 2nd, 3rd ... $N$th transmission attempt. Total successful user in a slot represents the summation of total successful user in that slot from all attempt (or backoff stage). Let $S_i[n]$ be the expected total successful user which conduct its $n$th transmission attempt at slot $i$. We can express

$$S_i = \sum_{n=1}^{N} S_i[n] \quad \text{for } i \geq 1$$

(2)

where $S_i[n]=0$ for $a>n>i$.

For slot 1, since it contains only 1st-attempt transmissions, $S_1=S_1[1]$, which can be calculated as

$$S_1[1] = \frac{M}{1} \left( 1 - \frac{1}{W_1} \right)^{M-1}$$

(3)

Equation (3) consists of 2 parts. The first part denotes the number of way for choosing 1 out of $M$ users. The second part indicates probability that a user would end up alone in slot 1. For the follows, since the combinational computation is only used for $\binom{x}{1}$, let us replace it with $x$ for any real number $x \geq 0$.

Subsequently, probability that slot 1 is collide, $P_{1,>1}$ can be calculated as

$$P_{1,>1} = 1 - P_{1,0} - S_1$$

(4)

In the simulation, $P_{1,1}$ is also similar to total number of experiments that yields a condition of “slot 1 is chosen by 1 out of $M$ users" divided by total number of experiments.

Let $C_i$ be the expected total collided user at slot $i$ and $C_i[n]$ be the expected total collided user which conduct its $n$th transmission attempt at slot $i$. Their relation can be written as

$$C_i = \sum_{n=0}^{N} C_i[n]$$

(5)

where $C_i[n]=0$ for $a>n>i$.

For slot 1, the expected total collided user, which only comes from 1st-attempt transmissions, can be calculated as

$$C_1[1] = \frac{M/W_1 - S_1}{P_{1,>1}}$$

(6)

Obtaining $C_1[1]$ resembles solving classic statistic problem of recalculating average point obtained by students in a class excluding students with known point and students who are absent during the exam, when the average point of all students is known.
Notice that in this analysis, in a slot, total successful user and total collided user CANNOT be summed together to yields total contending user, i.e. \( S_t + C_t \neq M/W_t \). It is because a slot cannot yields successful users and collided users at the same time.

\( C_t \) users that collide in slot 1 must conduct backoff for their 2nd transmission attempt. These users are possible to conduct their 2nd transmission attempt at slot \([2, 1+W_2]\).

Let us now analyze slot 2. The users who transmit at slot 2 can come from the following two groups: those who conduct their 1st attempt (since slot 2 is within the 1st backoff window) or those who are collided at slot 1 and then conduct their 2nd attempt. Slot 2 is empty if no user from those groups choose \( i=2 \). The probability of slot 2 to be empty, \( P_{2,0} \), is

\[
P_{2,0} = \left(1 - \frac{1}{W_1}\right)^M \left(1 \frac{1}{W_2}\right)^{C_t[1]} \tag{7}\]

The exponent for the first part, \( M \), denotes that it is sure some of \( M \) users will choose slot 2. On the other hand, the exponent for the second part, \( C_t[1] \), denotes that some of \( C_t[1] \) users will choose slot 2 if slot 1 is collide.

Before proceeding any further, note that \( C_t[n] \) also represents total unfinished user that has conducted its \( n \)th transmission attempt at \( i \)th slot. For notational convenience, let \( C_t[0] \) be total unfinished user at 0th slot that has not conduct any transmission attempt, and is equal to \( M \). Then we can rewrite Equation (6) as

\[
P_{2,0} = \left(1 - \frac{1}{W_1}\right)^{C_t[0]} \left(1 \frac{1}{W_2}\right)^{C_t[1]} \tag{8}\]

and slot 2 yields successful user if there is only 1 user choose \( i=2 \), either from the first or the second group. Hence, \( S_2 = S_2[1] + S_2[2] \), where

\[
S_2[1] = \frac{C_t[0]}{W_1} \left(1 - \frac{1}{W_1}\right)^{C_t[0]-1} \left(1 \frac{1}{W_2}\right)^{C_t[1]} \tag{9}\]

\[
S_2[2] = \left(1 - \frac{1}{W_1}\right)^{C_t[0]} \frac{C_t[1]P_{1,>1}}{W_2} \left(1 \frac{1}{W_2}\right)^{C_t[1]-1} \tag{10}\]

By the same manner of calculating \( C_1 \) in Equation (6), \( C_2 \) can be calculated as

\[
C_2 = \frac{(C_t[0]/W_1 + C_t[1]P_{1,>1}/W_2) - S_2}{P_{2,>1}} \tag{11}\]

And proportionally we can calculate \( C_t[1] \) and \( C_t[2] \) from \( C_2 \) as follows

\[
C_t[1] = C_2 \frac{C_t[0]/W_1}{C_t[0]/W_1 + C_t[1]P_{1,>1}/W_2} \tag{12}\]

\[
C_t[2] = C_2 \frac{C_t[1]/W_2}{C_t[0]/W_1 + C_t[1]P_{1,>1}/W_2} \tag{13}\]
while \( C_3[n] = 0 \) for \( 0 > n \geq 3 \).

Up until this point, notice that the approaches that has been conducted to reveal the mechanisms in the first and second slot can also be used extendedly to the subsequent slots. Finally, we can summarize the number of successful user from each slot and attempt to obtain the total number of successful users. Subsequently, the access success probability can also be calculated. Hence,

\[
S = \sum_{i=1}^{N} \sum_{n=i}^{\infty} S_i[n]
\]

\[
P_s = \frac{\sum_{i=1}^{N} \sum_{n=i}^{\infty} S_i[n]}{M}
\]

4. Evaluation and discussion

Evaluation is conducted to verify the correctness and accuracy of our proposed analytical model. Hence, a custom-made DCF Monte Carlo simulator is prepared to crosscheck the correctness of our analytical model. In the following result figures, the results generated by our analytical model are shown by lines, while the results generated by our simulation are shown by dot markers. Additionally, the simulation is performed with \( 10^4 \) iterations to obtain valid average values to be compared with the results from analytical model.

In this evaluation, let us consider a DCF system with RTS/CTS being activated, and serving \( M \) users which arrives simultaneously at the first slot. In this system, each user is allowed to transmit up to 5 times, and after collision the retransmission is conducted after a uniform backoff with window of 3 slots.

Figure 1 shows the probability of a slot to be empty (not used by any user), success (i.e. used by only 1 user), and collided (i.e. used by more than 1 user). These probability are notated as \( P[0] \), \( P[1] \) and \( P[>1] \), respectively. From this figure, we can observe that our analytical model can accurately estimate the performance of the system. This is denoted by the fact that all of the results from the simulation are coincident with the results from the proposed analytical model. Additionally, we can also observe that the collision in DCF system will be more likely to happen when the number of users increases. Meanwhile, lowering the number of users may also decrease the chance of a slot to be occupied. In fact, the more the slots are unoccupied, the lower the system’s utilization, which is also undesired.

![Figure 1](image-url)
Figure 2 plots the number of successful users for various number of loads (i.e. total number of users). Similar to the previous figure, this figure is also displaying consistent result between simulation and analytical model, which emphasis that our analytical model is accurate. In this figure, it is observed that with the help of RTS/CTS mechanism, number of successful users are almost equal to the increasing total number of users. Under the maximum number of load of $M=100$, there are about 84 users that are successful. This denotes the success probability of 84\%.

![Figure 2. Number of successful user under varying number of total users](image)

To better understand the dynamics of the system, Figure 3 and Figure 4 plot the statistics of number of transmission and number of successful transmission in each slot, which are represented in the form of cumulative distribution function (CDF). Since the retry limit is 5, these figures each show 10 lines in total (each attempt has its analytical and simulation results). From these figures, we can observe that in fact the simulation results are not exactly in agreement with the result obtained from our analytical model. It is especially shown for $n=5$. This is because in the 5th attempt, there are less retransmitting users. In simulation, this requires more iteration to produce smooth line plot. However, combined with the 1st to 4th attempt, these small errors are not noticeable in the overall result depicted in Figure 2. Notice that while allowing up to 5 transmission attempts, this system may yield high delay. However, noticing that there are many IoT applications that are delay-insensitive, DCF is still practical solution for serving IoT application.

![Figure 3. Cumulative distribution function (CDF) of number of transmission in each slot for $M=100$](image)
Figure 4. Cumulative distribution function (CDF) of number of successful transmission in each slot for $M=100$

5. Conclusion
This paper studies the performance of distributed coordination function (DCF) for serving IoT application. This study is important since WiFi is still contending solution for serving IoT application due to the reasons discussed above. An iterative-based analytical model is proposed to easily estimate the performance of the system in term of access success probability. The proposed analytical model is then evaluated against a compatible simulation for its correctness and accuracy. The results show that our analytical model can estimate the said performance metric accurately under various load condition. The result also demonstrates the usage of our proposed model to study and evaluate per-slot condition and overall performance of the system.

6. Acknowledgements
This research / article’s publication is partially supported by the United States Agency for International Development (USAID) through the Sustainable Higher Education Research Alliance (SHERA) Program for Universitas Indonesia’s Scientific Modeling, Application, Research and Training for City-centered Innovation and Technology (SMART CITY) Project, Grant #AID-497-A-1600004, Sub Grant #IIE-00000078-UI-1.

7. References