Modelling Co-operative MAC Layer Misbehaviour in IEEE 802.11 Ad Hoc Networks with Heterogeneous Loads

Rohith Dwarakanath Vallam, A. Antony Franklin, and C. Siva Ram Murthy *
Department of Computer Science and Engineering,
Indian Institute of Technology Madras, Chennai - 600036, India
{rohith, antony}@cse.iitm.ernet.in, murthy@iitm.ac.in

***Abstract*— Misbehaviour due to back-off distribution manipulation has been one of the significant problems faced in IEEE 802.11 wireless ad hoc networks which has been explored recently by the research community. In addition, collusion between misbehaving nodes adds another dimension to this security problem. We examine this problem in a three-node network scenario wherein two nodes are assumed to be malicious colluding adversaries causing unfair channel access to the other legitimate node. The misbehaving nodes, through back-off manipulation, will try to minimize the channel access share got by the legitimate node and at the same time maximize the detection delay to detect such an attack. We explore this problem and its solution, analytically, in a non-saturated setting, by modelling a single IEEE 802.11 node as a Discrete Time Markov Chain (DTMC) and suggest a measure for evaluating fairness in the network. We then propose an attacker-detector non-linear optimization model through which the joint optimal attacker distribution is evaluated by applying results from the area of variational calculus. We finally use the Sequential Probability Ratio Test (SPRT) for estimating the average number of samples for detecting colluding adversaries in the network. We validate all the models using MATLAB and verify the model results by sampling values from the evaluated optimal attacker distribution using a robust statistical library called UNU.RAN.**

I. Introduction

As we enter the age of ubiquitous wireless networks, the issue of security in such networks is growingly becoming a pervasive problem. New vulnerabilities in these networks have emerged in these networks and thus, solutions related to these security issues in wireless networks has been explored by the research community of late. Security issues in wireless ad hoc networks pose significant challenges due to the unpredictable nature of the wireless medium and independent behaviour of the nodes in the network. These challenges make security a very interesting area of research. There have been significant progress towards addressing issues related to this area. The broad area of security in wireless networks encompass issues like privacy protection, naming and addressing, secure neighbour discovery and secure routing. Also, there are issues regards to trust evaluation, secure localisation, behaviour enforcement, selfishness in packet forwarding, and selfish behaviour at MAC layer.

This paper is related to the issue of security at the Medium Access Control (MAC) layer and in particular, the problem of thwarting misbehaviour by malicious IEEE 802.11 nodes whose objective is to cause unfair channel access to legitimate nodes in its transmission range. We model the problem from an analytical perspective and try to understand the various aspects of the problem and use mathematical analysis i.e., Markov chain modelling, variational calculus based non-linear optimization theory, and statistical estimations to get insights into detecting misbehaviour of the monitored nodes. The main contributions of this paper can be put forth as follows.

- The problem of back-off manipulation has been addressed in a non-saturated scenario where each of the nodes in the network have different data arrival rates following the Poisson distribution.
- A detailed model of a single IEEE 802.11 non-saturated node based on Discrete Time Markov Chain (DTMC) is proposed and steady state probabilities are evaluated.
- The global system state of the three-node network is then modelled probabilistically and a new fairness measure is proposed for the nodes in the network.
- The problem of collusion of two nodes is taken and a non-linear optimization model is developed to depict the colluding attackers-detector scenario.
- By using principles from the area of variational calculus, the optimal joint probability density function of the colluding attackers is found and the average sample size for detecting such an optimal attack is evaluated using a statistical approach namely the Sequential Probability Ratio Test (SPRT).

*Author for correspondence.

The rest of the paper is organised as follows. Section II discusses related work in the area of security in IEEE 802.11 MAC layer. Section III explains the problem setting and formulates the problem addressed in this paper. Section IV gives an overview of the approach followed and then, explains, in detail, the modelling aspects of the various aspects of this problem. Section V validates the models by plotting results from MATLAB. Section VI discusses the conclusions and future work.

II. Related Work and Motivation

The ad hoc network community has tried to understand and address issues related to attack resistance at MAC layer in recent times. In [1], the authors study simple DoS attacks at the MAC layer, show their dependence on attacker traffic patterns, and deduce that the use of MAC layer fairness can mitigate the effect of such attacks. In [2], the focus is also on DoS attacks against the IEEE 802.11 MAC protocol. They describe vulnerabilities of IEEE 802.11 and show ways of exploiting them by tampering with normal operation of device firmware.

There has been some significant work on detecting MAC layer misbehaviour in Wireless LANs [3]. A modification to the IEEE 802.11 MAC protocol is proposed to facilitate the detection of selfish and misbehaving nodes. The approach assumes a trustworthy receiver, since the receiver assigns to the sender the back-off value to be used. The receiver can detect misbehaviour of the sender and accordingly penalize it by providing less favourable access conditions through higher back-off values for future transmissions. A decision about protocol deviation is reached if the observed number of idle slots of the sender is smaller than a pre-specified fraction of the allocated back-off. The sender is labelled as misbehaving if it turns out to deviate continuously based on a cumulative metric over a sliding window. This work also presents techniques for handling potential false positives due to the hidden terminal problem and the different channel quality perceived by the sender and the receiver. However, our work differs from [3] and [4], as we consider an ad hoc environment wherein no trusted centralized Access Point (AP) can be assumed.

Also, there have been recent approaches like [4],[5], and [6], that have addressed the problem of back-off manipulation at MAC layer. The authors in [4], focus on MAC layer misbehaviour in wireless hot-spot communities. They propose a sequence of conditions on some available observations for testing the extent to which MAC protocol parameters have been manipulated. The advantage of the scheme is its simplicity and easiness of implementation, although in some cases the method can be deceived by cheating peers, as the authors point out. Greedy behaviour by the nodes is considered in [4] and not malicious behaviour by nodes as considered in this paper.

Detecting MAC layer back-off timer violations in ad hoc networks have been studied in [5]. They exchange the state of the random number generator of each of the neighbours by modifying the IEEE 802.11 protocol and then, using Wilcoxon rank sum test, which uses fixed sample size, compare difference between analytically computed samples with observed samples and detect misbehaviour exists or not. However it does not handle collusion between nodes. Also, they have an approach wherein the number of samples required for detection is fixed. Our work uses an optimal statistical method, SPRT [7], for adaptive estimation of number of samples for misbehaviour detection.

The problem of determining the attacker distribution in the saturated case (i.e., all nodes have data to send in every time slot) has been addressed in [6]. The work considers the case of colluding attackers, but in a network where all nodes have always data to send i.e., they are saturated. In real IEEE 802.11 networks, data and multimedia traffic (for eg., traffic due to e-mail, Internet, audio and video) is inherently bursty [8] in nature. The demanded transmission rate for most real traffic varies with significant idle periods and hence, nodes are usually far from being saturated. So, we study the effect of back-off manipulation at the MAC layer in a network where nodes may have different data rates. In addition, we explore this misbehaviour when there is co-operation between misbehaving nodes to jointly cause unfairness to the legitimate nodes. This collusion between adversarial nodes makes the detection of such an attack harder and hence, in order to aid the detection mechanism, there is a need to evaluate the worst case attack that can be caused by this collusion. We consider this problem of back-off attack when there are colluding nodes and the network is non-saturated.

III. Problem Setting and Formulation

This paper addresses a vulnerability in the IEEE 802.11 Distributed Co-ordination Function (DCF) Medium Access Control (MAC) protocol [9] namely the back-off attack. In a back-off attack, nodes will not follow the uniform distribution for choosing a waiting time (back-off) after successful packet transmission. They will choose smaller waiting times from a different non-uniform distribution resulting in unfairness in channel access. Consider a three-node (referred to as Node 1, Node 2, and Node 3) ad hoc wireless network where each node is in the wireless range of the other. The primary

objective of the adversarial nodes is to cause unfairness, with respect to channel access, to the legitimate node. Consider two nodes (Node 2 and Node 3) in the network as colluding malicious nodes whose aim is to disrupt the channel access of the other legitimate node (Node 1). The objective, from the colluding attackers' point of view, is to determine their back-off values in such a way that it causes maximum unfairness in the network. In other words, the attackers will not follow the uniform distribution for selecting the back-off values between packet transmissions as specified by the IEEE 802.11 DCF protocol and thus, try to deny fair access to the channel by the legitimate node. The legitimate node, on the other hand, will be sampling the back-off values used by each of its neighbours by some mechanism like the statistical monitoring mechanism proposed in [5] and will test these samples to check if the neighbours are misbehaving or not. To perform this function, the legitimate node needs to fix the number of samples that needs to be collected after which it is enough to decide if the neighbours are misbehaving or not. Hence, a mechanism for estimating the average sample size required for detection is needed.

Assumptions:

- Each node is supposed to follow the IEEE 802.11 DCF protocol as the MAC layer protocol.
- The nodes are static or moving with a velocity such that they continue to be in the transmission range of each other.
- Each node has some data to be sent to any of the other nodes and the arrival process in Poisson. All the nodes are aware of the traffic loads at the other nodes in its vicinity.
- The higher layers of the network stack generates the data traffic. The delay occurring between the time the packets are generated to the time when they arrive at the MAC layer for transmission, which may be due to delay at the higher layers like TCP is not considered in this work.
- Each node has small buffers (as small as possible to avoid the effect of queueing dynamics) to hold the incoming packets from higher layers.
- The nodes are able to sense the channel at any time (using promiscuous mode) and hence detect if the medium is busy or idle. If busy, it can sniff the packets to know information about the headers in the packet like destination, duration of the transfer, sequence number, etc.
- Once a packet encounters a collision, it is dropped and no retransmissions are attempted. This is to simplify analysis as the main focus of this paper is to explore the collusion problem between two nodes under non-saturated conditions and not the way packet collision is handled.

IV. Proposed Solution

A. A Broad Overview

The problem under consideration can be approached in the following way:

- Develop a detailed mathematical model for understanding the behaviour of a single node following the IEEE 802.11 DCF mechanism considering the assumptions mentioned in Section III.
- Develop a model for the system as a whole and the various states that the system may be in. This will provide a global outlook of the system which will be useful to understand the Quality of Service (QoS) aspects, like fairness, of the three-node network .
- Formulate an optimization scenario, wherein, under the evaluated fairness conditions, the colluding attackers try to maximize the unfairness of the network, while at the same time, try to avoid detection by the legitimate node to the maximum extent possible.
- Estimate the optimal attacker back-off distribution from the optimization problem thus formulated and then derive the value for the expected sample size required to detect such an optimal attack. Due to the optimality of the attacker distribution, any other attack caused by the colluding nodes will be suboptimal and the sample size thus calculated for the optimal case will be enough to identify these suboptimal attacks, on an average.

B. A Single Node Model

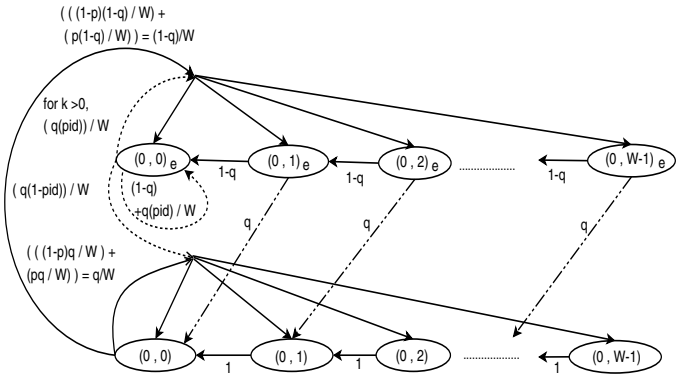

Fig. 1: Discrete Time Markov Chain Model of IEEE 802.11 DCF protocol

1) Preliminaries: Non-saturation in IEEE 802.11 networks was considered in [10]. They model a node based on the well known Bianchi [11] model of IEEE 802.11

DCF node. In this model, an IEEE 802.11 DCF node is modelled as a Discrete Time Markov Chain (DTMC) with each state being denoted by a pair of integers (s, k) where s is the back-off stage and k is the back-off counter value at that stage as shown in Fig. 1. We use this model as the basis of our work, but as mentioned earlier, add a restriction that the collision of a packet is handled by dropping the packet thereby not considering it for retransmission i.e., the node is always in back-off stage 0.

2) Analytical evaluation of the model: As in [10], due to the non-saturation assumption, idle states can be present wherein there is no data for transmission and they are represented by the states $(0, k)_e$ (known as the post-back-off states). Hence the DTMC will either be in any of the $(0, k)$ states if there is packet to be transmitted or in any of the $(0, k)_e$ states if the node is idle. Transmission of a packet is attempted either in $b(0, 0)$ state after the back-off counter (which is selected uniformly from $[0, W_0 - 1]$ where W_0 is the maximum contention window size at stage 0) reaches zero or when a packet arrival occurs in the $b(0, 0)_e$ state. Note that W_0 is shown as W in Fig. 1. This DTMC is solved for steady state probabilities (represented by the stationary distribution b) analytically by first formulating the one-step state transition probabilities (as shown in Fig. 1) and then finding the expressions for the steady state probability of each state. The notations followed are:

- p - probability of collision given the node is attempting transmission.
- $(1 - q)$ - probability that the node's buffer has no packets awaiting transmission at the start of each counter decrement.
- $b(0, k)_e$ - steady state probability of being in state $(0, k)_e$ of the DTMC where $k \in [0, W_0 - 1]$.
- $b(0, k)$ - steady state probability of being in state $(0, k)$ of the DTMC where $k \in [0, W_0 - 1]$.
- P_{idle} (shown as pid in Fig. 1) - probability that the medium is sensed idle during a typical slot.

Determining one-step state transition probabilities:
If the node is in $(0, 0)$ state, two things can happen. It might get a packet to send or not. In case of packet arrival (with probability q), the node will choose a back-off uniformly and transition into any one of the $(0, k)$ states. If no data arrives, the node will choose a uniform back-off in the range $[0, W_0 - 1]$ and move into any one of the $(0, k)_e$ states. To put it in terms of state transitions,

$$P((0, k)|(0, 0)) = q/W_0$$
$$P((0, k)_e|(0, 0)) = (1 - q)/W_0$$

If the MAC is in $(0, 0)_e$ state, three things can occur. (i) A packet may arrive in which case, the medium is sensed and if it is idle, the packet is transmitted. Due to our assumption, after the packet is transmitted, the MAC enters back to the $(0, k)_e$ chain irrespective of whether the packet collided or not as the case of retransmission of the packet (in case of collision) is not handled in this work.

$$P((0, 0)_e|(0, 0)_e) = (1 - q) + qP_{idle}/W_0$$

(ii) If the medium is busy, then the MAC enters stage-0 back-off by choosing a uniformly distributed back-off in the range $[0, W_0]$.

$$P((0, k)|(0, 0)_e) = q(1 - P_{idle})/W_0$$

(iii) If no packet arrives in the considered slot, then the MAC will loop in the $(0, 0)_e$ state.

$$k > 0,\ P((0, k)_e|(0, 0)_e) = qP_{idle}/W_0$$

By similar reasoning, the other one step transition probabilities can be described as below.

$$\begin{aligned} P((0, k-1)|(0, k)) &= 1 \\ P((0, k-1)_e|(0, k)_e) &= 1 - q \\ P((0, k-1)|(0, k)_e) &= q \end{aligned}$$

In order to evaluate the steady state probabilities in the setup described above, we make the following observations. With $b(i, k)$ and $b(0, k)_e$ denoting the steady state probabilities of being in states (i, k) and $(0, k)_e$, we have

$$\sum_{k=0}^{W_0-1} b(0, k) + \sum_{k=0}^{W_0-1} b(0, k)_e = 1 \qquad (1)$$

Eqn. (1) is a very important equation for our simplification. The objective of the following simplification is to reduce the two sums in Eqn. (1) in terms of a common term, $b(0, 0)_e$, and evaluate the rest of the steady state probabilities in terms of this value. We proceed as given below to achieve this simplification.
We know that

$$b(0, W_0 - 1)_e = \frac{b(0, 0)_e q(1 - p)P_{idle}}{W_0} + \frac{(1 - q)b(0, 0)}{W_0}$$

For, $(W_0 - 1) > k > 0$,

$$b(0, k)_e = (1 - q)b(0, k + 1)_e + b(0, W_0 - 1)_e$$

Simplifying recursively,

$$b(0, k)_e = \frac{qb(0, 0)_e - b(0, W_0 - 1)_e \left(\frac{1-(1-q)^k}{q}\right)}{(1 - q)^k}$$

Using the above expressions and following some basic simplification, we can evaluate

$$\frac{b(0, 0)_e}{b(0, 0)} = \frac{1 - q}{q}\left(\frac{1 - (1 - q)^{W_0}}{qW_0 - (1 - p)P_{idle}(1 - (1 - q)^{W_0})}\right)$$

We then derive the following,

$$\sum_{k=0}^{W_0-1} b(0,k)_e = \frac{b(0,0)_e}{(1-(1-q)^{W_0})}(qW_0) \quad (2)$$

To evaluate $\sum_{k=0}^{W_0-1} b(0,k)$, we know that

$$b(0,k) = b(0,0) - q\sum_{i=1}^{k} b(0,i)_e - kb(0,W_0-1)$$

By using the above equation, $b(0,i)_e$ can be simplified to

$$b(0,i)_e = ((q^2 b(0,0)_e - b(0,W_0-1)_e) \times (1/(1-q)^k - 1) + qi\ b(0,W_0-1)_e)/(q^2)$$

To sum it up, we start by splitting the terms as following.

$$\sum_{k=0}^{W_0-1} b(0,k) = b(0,0) + b(0,W_0-1) + \sum_{k=1}^{W_0-2} b(0,k) \quad (3)$$

Further, $b(0,k)$ can be written as below.

$$b(0,k) = b(0,0) - result_1 - result_2 \quad (4)$$

where $result_1$ and $result_2$ are got by further simplification as below.

$$result_1 = k\left(\frac{b(0,0)_e q}{W_0}(1-pP_{idle}) + \frac{b(0,0)}{W_0}(1-pq)\right)$$

$$result_2 = \left(\frac{1}{q}\left(\frac{1}{(1-q)^k} - 1\right)\right) \times \left(\frac{1}{q}\left(q^2 b(0,0)_e - b(0,W_0-1)_e\right)\right)$$

Since we need to get an expression for $\sum_{k=0}^{W_0-1} b(0,k)$, we use Eqn. (3) and Eqn. (4) and get the following expression.

$$\sum_{k=0}^{W_0-1} b(0,k) = b(0,0) + b(0,W_0-1) + \sum_{k=1}^{W_0-2} b(0,0) - \sum_{k=1}^{W_0-2} result_1 - \sum_{k=1}^{W_0-2} result_2$$

Simplifying, the following equation results

$$\sum_{k=0}^{W_0-1} b(0,k) = result_3 - result_4 + b(0,0) + b(0,W_0-1) \quad (5)$$

where $result_3$ and $result_4$ are as follows:

$$result_3 = \left[\left(\left(((W_0-2)/(2W_0))b(0,0)\right) \times \left(2W_0 - \left((W_0-1)(1-pq)\right)\right)\right) - \left(\left(\left(b(0,0)_e/W_0\right)q(1-pP_{idle})\right) \times \left((W_0-2)(W_0-1)/2\right)\right)\right]$$

$$result_4 = \left[\left(\left(b(0,0)_e q\right)\left((1/\gamma)-1\right)\right)\delta\right]$$

$$\delta = \left(\left((1/q) \times \left(1/((1-q)^{W_0-2}) - 1\right)\right) - \left(W_0 - 2\right)\right)$$

$$\gamma = \left(1 - \left((1-q)^{W_0}\right)\right)$$

As we can clearly observe from the Eqn. (2) and Eqn. (5), we have got the values of the two sums in terms of $b(0,0)_e$ and hence the normalisation equation of Eqn. (1) can now be used to determine $b(0,0)_e$ in terms of q, W_0 and P_{idle}. From this value, we can now evaluate values of $b(0,0)$, $\sum_{k=0}^{W_0-1} b(0,k)$, and $\sum_{k=0}^{W_0-1} b(0,k)_e$. These values are used in the next section to determine the state of the system as a whole and define fairness condition for the network.

C. A Global System Model

1) Details of the Model: Using the single node model proposed in Section IV-B, we define the following:

a_i = Probability that the node i is choosing a back-off value in a time slot. We can say that a node will start to choose a random back-off value in the current time slot in the following two scenarios.

- The node is in $b(0,0)$ state and a new packet awaits transmission at the end of the packet transmission.
- The node is in any of the $b(0,k)_e$ states and a new packet arrives for transmission in the current time slot.

Hence, from Fig. 1, we can deduce

$$a_i = \left(q_i \times \left(b(0,0) + \sum_{k=0}^{W_0-1} b(0,k)_e\right)\right) \quad (6)$$

where q_i is the probability that Node i has at least one packet to be sent at the start of each time slot.

b_i = Probability that the Node i is not choosing a back-off value in the considered time slot ($b_i = 1 - a_i$).

Now, in the current three-node network under consideration, in a time slot, there may be 0, 1, 2 or 3

nodes which are independently choosing a back-off. It should be noted here that we do not assume that all nodes choose their back-off timers at the same time. Each node can choose back-off timers independently and thus, a global system state needs to be defined, to determine the nodes which are choosing a back-off value in a particular slot.

We put forth the following definitions to capture the state of the system at the beginning of any time slot. Let X_i denote that Node i is choosing a back-off value in a time slot. Correspondingly, let Y_i denote that Node i is not choosing a back-off value in a time slot. Then the three-node network can be identified by any one of the eight states given in Table I. Correspondingly, their steady state probabilities are also evaluated in the table.

TABLE I: Global system model of the three node network

State ID	State Representation	Steady State probability
p_1	$(Y_1\, Y_2\, Y_3)$	$(b_1 \times b_2 \times b_3)$
p_2	$(X_1\, Y_2\, Y_3)$	$(a_1 \times b_2 \times b_3)$
p_3	$(Y_1\, X_2\, Y_3)$	$(b_1 \times a_2 \times b_3)$
p_4	$(X_1\, X_2\, Y_3)$	$(a_1 \times a_2 \times b_3)$
p_5	$(Y_1\, Y_2\, X_3)$	$(b_1 \times b_2 \times a_3)$
p_6	$(X_1\, Y_2\, X_3)$	$(a_1 \times b_2 \times a_3)$
p_7	$(Y_1\, X_2\, X_3)$	$(b_1 \times a_2 \times a_3)$
p_8	$(X_1\, X_2\, X_3)$	$(a_1 \times a_2 \times a_3)$

2) A Measure of Fairness: We know that, based on the problem definition given in Section III, Node 1 is the legitimate node. Also we know that Node 2 and Node 3 are potential misbehaving adversaries to Node 1. Before we go any further, we need to understand the meaning of misbehaviour in this context. Let us suppose that in a particular time slot, two nodes choose a back-off value. In the ideal case, if they obey the IEEE 802.11 DCF protocol, they should choose back-off uniformly from the range $[0, W_0]$ where W_0 is the maximum contention window size at back-off stage 0 (assuming both the nodes are in back-off stage 0). In this case, the node which gets access to the channel is the node which chooses the smaller back-off of the two nodes. So, the probability of a node's back-off being less than the other's node's back-off is $1/2$ as they are obeying the protocol. We claim that *this probability value* is an indicator of the *fairness* in the network and it is, at this point, that misbehaviour can occur which can lead to unfair channel access to the legitimate node. If one of the two nodes is misbehaving and not following the IEEE 802.11 protocol, then the probability of that node's back-off being smaller than that of the other node will definitely not be $1/2$. As the misbehaving node's objective is to reduce channel access to the other node, the probability that its (misbehaving node) back-off will be less than the other (legitimate) node will be more than $1/2$ and misbehaviour will reach maximum when the probability value becomes 1. So the misbehaving node will choose a random back-off value from such a distribution which will ensure that the probability of its back-off value being smaller than the legitimate node's back-off becomes 1. Such a distribution is known as the optimal attacker's distribution. Mathematically, in the three-node network considered in this problem, let r_{ik} represent the reward of Node i in state p_k where $k \in [1 \ldots 8]$. The reward r_{ik} is nothing but the probability that Node i has smaller back-off than other competing nodes in state p_k. Based on this, the combined channel access share of Node 2 and Node 3 (the potential colluding nodes) can be evaluated in generalised terms as

$$Combined\ Share\ of\ Node\ 2\ and\ Node\ 3 = \sum_{k=1}^{8}\left[p_k \times \left(\sum_{i=2}^{3} r_{ik}\right)\right] \quad (7)$$

If there is no misbehaviour by any of the nodes, then we use the following equation for calculating fair combined channel access share (denoted by ρ) of Node 2 and Node 3.

$$r_{ik} = \begin{cases} 1/n & \text{if Node } i \text{ is among the } n \text{ nodes} \\ & \text{choosing back-off in state } p_k \\ 0 & \text{if Node } i \text{ is not choosing a back-off} \\ & \text{in state } p_k \end{cases}$$

Similarly, if Node 2 and Node 3 are misbehaving, in all the states in which any of them (or both) are choosing back-off, namely the states in $[p_3, p_4, p_5, p_6, p_7, p_8]$, they will try to get maximum reward (i.e., 1) together in that state. So, we use the following equation in Eqn. (7) for calculating the maximum unfair combined channel access share (denoted by σ) of Node 2 and Node 3 (from a back-off selection point of view).

$$\sum_{i=2}^{3} r_{ik} = \begin{cases} 1 & \text{if either Node 2 or Node 3 (or both)} \\ & \text{is(are) choosing a back-off value} \\ & \text{in state } p_k \\ 0 & \text{if neither Node 2 nor Node 3 is} \\ & \text{choosing a back-off value} \\ & \text{in state } p_k \end{cases}$$

The colluding attackers can misbehave with different degrees and the share that they get by their misbehaviour can vary between the above two bounds given by ρ and σ. So, putting this together and applying some mathematical simplification, we formulate the colluding attackers' constraint. The colluding attackers' constraint can be formulated by using the reasoning described above as given below.

$$\sum_{k=1}^{8} [p_k \times Pr(min(X_2, X_3) < X_1)] \ \leq \ (\rho + \eta) \quad (8)$$
$$\eta = \delta * (\sigma - \rho) \ where \ \delta \in [0, 1]$$

where η represents the amount of misbehaviour and δ represents the misbehavior coefficient. Based on the global system model described earlier, we can setup the following conditions to evaluate Eqn. (8) for the state p_k.

$$Pr(min(X_2, X_3) < X_1) = \begin{cases} Pr(X_2 < X_1) & \text{if } k = 4 \\ Pr(X_3 < X_1) & \text{if } k = 6 \\ 1 & \text{if } k = 3, 5, 7 \\ 0 & \text{if } k = 1, 2 \\ Pr(min(X_2, X_3) < X_1) & \text{if } k = 8 \end{cases}$$

By applying probability rules (See Appendix), we can derive the following equation where Y_1 is an random variable with an unknown probability density function (pdf) and Y_2 is a uniform random variable.

$$Pr(Y_1 < Y_2) = (1 - E(Y_1)/W_0)$$

where $E(Y_1)$ is the expected value of random variable Y_1. Applying the above formula in Eqn. (8), we get

$$\sum_{k=3}^{8} p_k - p_4 * E(X_2)/W_0 - p_6 * E(X_3)/W_0 - p_8 * E(min(X_2, X_3))/W_0 \leq (\rho + \eta)$$

$$\Rightarrow (-p_4)E(X_2)/W_0 + (-p_6)E(X_3)/W_0 + (-p_8)E(min(X_2, X_3)/W_0) \leq ((\rho + \eta) - (1 - (p_1 + p_2)))$$

Setting $\alpha = (-p_4); \beta = (-p_6); \gamma = (-p_8)$, we get

$$\int_0^{W_0} \int_0^{W_0} (\alpha x_2 + \beta x_3 + \gamma min(x_2, x_3)) \times f_{23}(x_2, x_3) \ dx_2 \, dx_3 \leq W_0 \left((\rho + \eta) - (1 - (p_1 + p_2))\right) \quad (9)$$

Here, $f_{23}(x_2, x_3)$ is the joint attacker density function of Node 2 and Node 3, δ is the misbehaviour coefficient, η is the amount of misbehaviour and W_0 is the maximum contention window size as given in Section IV-B.

D. A Statistical Approach for Detecting Colluding Adversaries

We will use the Sequential Probability Ratio Test (SPRT) for determining if the Node 2 and Node 3 are colluding nodes or not. This test is a method of statistical inference whose characteristic feature is that the number of observations required by the procedure is not determined in advance of the experiment. The decision to terminate the experiment depends, at each stage, on the results of the observations previously made. Thus, the number of observations required by SPRT is not predetermined, but is a random variable. It has been shown in the literature that of all statistical tests with the same power, the SPRT requires fewest observations on the average [7]. In the context of our problem, we define two simple hypotheses H_0 and H_1 as follows:

$H_0 \longrightarrow$ Node 2 and Node 3 are legitimate nodes obeying the IEEE 802.11 DCF protocol for selecting back-off values i.e., the nodes follow $f_0(x_2)$ and $f_0(x_3)$ as their back-off distribution and f_0 represents the uniform distribution.

$H_1 \longrightarrow$ Node 2 and Node 3 are colluding adversaries following an optimal joint probability density function (pdf) $f_{23}(x_2, x_3)$ for selecting their back-off values.

From the mathematical point of view, it is important here to mention that, throughout this paper, we shall be considering only random variables which either admit a probability density function or have a discrete distribution. By the probability distribution, or more briefly distribution, $f(t)$, of a random variable X, we shall always mean the probability density function of X, if it exists. If X is a discrete random variable, $f(t)$ will denote the probability that $X = t$. We have assumed $f_{23}(x_2, x_3)$ as a continuous density function for the sake for analysis. In practical scenarios, as back-off values are usually discrete integers, we can think of nodes sampling from $f_{23}(x_2, x_3)$ and use the value after rounding off to the nearest integer.

The SPRT collects observations until significant evidence in favour of one of the two hypotheses is accumulated. The legitimate node (Node 1) is assumed to execute the SPRT and it periodically collects samples from each of its neighbours. These samples are an indication of the back-off values used by the neighbours. After collecting each sample, at the $i-th$ stage, Node 1 chooses between the following options: accept one or the other hypothesis and stop collecting samples, or defer decision for the moment and obtain observation $i + 1$. The SPRT has two thresholds a and b that aid the decision. The figure of merit at each step is the logarithm of the likelihood ratio of the accumulated sample vector until that stage. Let (x_i, y_i) represent the sample vector collected from Node 2 and Node 3 respectively at stage i. For any positive integral value i, the probability that a sample $(x_1, y_1), ..., (x_i, y_i)$ is obtained from Node 2 and Node 3 is given by

$$p_{1i} = f_{23}(x_1, y_1)...f_{23}(x_i, y_i)$$

when H_1 is true, and by

$$p_{0i} = (f_0(x_1, y_1)) ... (f_0(x_i, y_i))$$

when H_0 is true. The latter equation is due to the fact that, when H_0 is true, both the Node 2 and Node 3 follow the uniform distribution independently.

Let α be the probability that H_1 will be accepted when H_0 is true and β be the probability that H_0 will be accepted when H_1 is true These parameters are the strength parameters of the statistical test. So, in the context of our problem, α represents the probability of false alarms and β represents the probability of missing the detection. For the case of testing between hypotheses H_0 (normal behaviour) and H_1 (misbehaviour), that involve probability distributions f_0 and f_{23}, the logarithm of the likelihood ratio z_i at stage i with accumulated samples $(x_1, y_1), ..., (x_i, y_i)$ is calculated as follows.

$$z_i = log\left(\frac{f_{23}\left((x_1, y_1), ..., (x_i, y_i)\right)}{f_0((x_1, y_1), ..., f_0(x_i, y_i))}\right)$$

We assume the observation samples are statistically independent. Hence,

$$z_i = \sum_{j=1}^{i} log\left(\frac{f_{23}((x_j, y_j))}{f_0(x_j) * f_0(y_j)}\right)$$

Now, the estimated sample size, $E(N)$, for SPRT is calculated based on [12] where $L(\theta)$ is the Operating Characteristic (OC) function of the test.

$$\begin{aligned} \alpha = P_{FA} = 0.01;\ \beta = P_M = 0.01; \\ A = \frac{(1-\beta)}{\alpha};\ B = \frac{\beta}{(1-\alpha)} \\ z = log\left(\frac{f_{23}(x_i, y_i)}{f_0(x_i) * f_0(y_i)}\right) \\ L(\theta) = \left(\frac{A^{-1} - 1}{A^{-1} - B^{-1}}\right) \\ c = (L(\theta)\ logB\ +\ (1 - L(\theta))\ logA) \\ E(N) = \left(\frac{c}{E(z)}\right) \end{aligned} \quad (10)$$

At each stage i of the experiment (at each integral value of i), the cumulative sum $S = z_1 + z_2 + ... + z_i$ is computed. If $log\ B \leq S \leq log\ A$, the experiment is continued by taking additional observation, If $S \geq log\ A$, the experiment is terminated with the acceptance of H_1. If $S \leq log\ B$, the experiment is terminated with the acceptance of H_0.

Thus, the intelligent and adaptive attackers will try to maximize the number of samples required by the monitoring node (Node 1). As per Eqn. (10) given above, this can be achieved by minimizing $E(z)$.

E. The Attacker-Detector Model

1) Formulation of the Optimization Problem: Putting the above discussion together, we can formulate an optimization scenario wherein the attackers try to minimize the channel access share for the legitimate node in the network, and at the same time, maximize the detection delay by the monitoring node. Let $X2, X3$ be two random variables representing the back-off values chosen by Node 2 and Node 3 following a joint pdf $f_{23}(x_2, x_3)$. The basic problem is to find the optimal joint pdf $f_{23}(x_2, x_3)$ which minimizes the following objective function i.e.,

$$\begin{aligned} &\min_{f_{23}}\ g_{23}(x_2, x_3) = \\ &\int_0^{W_0}\int_0^{W_0} log(f_{23}(x_2, x_3))\, f_{23}(x_2, x_3)\, dx_2\, dx_3 \\ &\text{subject to} \int_0^{W_0}\int_0^{W_0} f_{23}(x_2, x_3) dx_2\, dx_3 \ = \ 1 \end{aligned} \quad (11)$$

In addition, Eqn. (9) needs to be satisfied as discussed in Section IV-C.2.

2) Determining the Optimal Attacker Distribution: We can see that this problem involves solving the optimization problem, given by Eqn. (9) and Eqn. (11), in the space of functions. We form the Lagrangian function converting the constrained optimization problem to an unconstrained problem by using the method of Lagrange multipliers. From variational calculus [13], we examine the problem of finding the extrema of the functional

$$v(z(x, y)) = \iint_D F\left(x, y, z, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right) dx\, dy \quad (12)$$

where x and y are independent variables and values of the functions $z(x, y)$ on the boundary C of the domain D are prescribed, i.e., there is given a curve C^* in the three-dimensional space, and every admissible surface $z(x, y)$ is supposed to pass through this curve. The Lagrangian function of the optimization problem is of the form given in Eqn. (12) where the solution to the optimization problem is the joint optimal attacker distribution f_{23}. It follows that each function $z(x, y)$ which gives an extremum to Eqn. (12) should satisfy the following second order partial differential equation, known as the Ostrogradski's equation [13].

$$F_z - \frac{\partial}{\partial x}(F_p) - \frac{\partial}{\partial y}(F_q) = 0$$

where

$$p = \frac{\partial z}{\partial x};\quad q = \frac{\partial z}{\partial y};$$

Here, F_z is interpreted as partial derivative of F with respect to z. Similar interpretation holds for F_p and F_q.

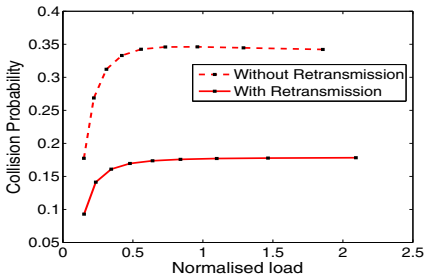

Fig. 2: Normalised Load versus collision probability for $W_0 = 8$

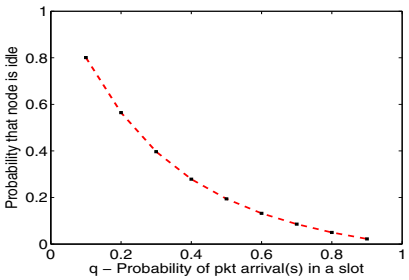

Fig. 3: Probability of queue being non-empty versus probability of node being in idle states for $W_0 = 8$

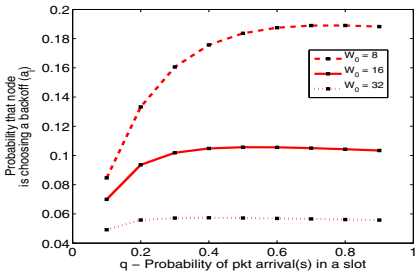

Fig. 4: Normalised Load versus Probability of choosing a back-off for $W_0 = 8, 16$ and 32

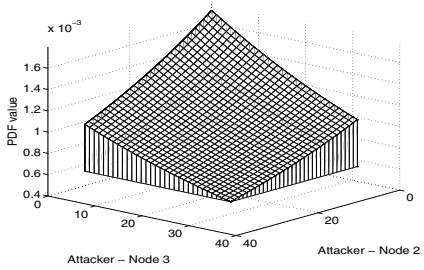

Fig. 5: 3-D mesh curtain plot of attackers' pdf for $W_0 = 32$ and $q_1 = q_2 = q_3 = 0.7$

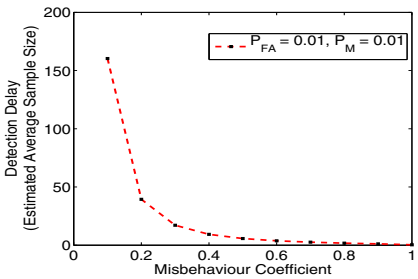

Fig. 6: Misbehaviour co-efficient versus detection delay at $q_1 = q_2 = q_3 = 0.7$

Applying correspondence between $z(x, y)$ and $f_{23}(x_2, x_3)$, we can get the optimal attacker distribution as

$$f_{23}(x_2, x_3) = e^{-1-\lambda-\mu(\alpha x_2+\beta x_3+\gamma min(x_2,x_3))} \quad (13)$$

where λ and μ are Lagrange multipliers which need to be evaluated using Eqn. (11) and Eqn. (9).

V. Validation and Results

A. Single Node DTMC model: We can see from the results of our MATLAB implementation of the model, i.e., in Fig. 2, that the collision probability increases linearly with load and reaches a saturation at the normalised load of about 0.5. Due to our assumption that a packet will not be retransmitted if it encounters collision, the nodes do not select back-off from a higher contention window and thus results in higher collision probabilities than in the scenario where packet collisions lead to contention window doubling. Also, from the model results, we can see, from Fig. 3, that as the value of q increases the node will spend lesser time in the idle states as it tends to have more data to transmit. Finally, it tends towards the saturation assumption as $q \rightarrow 1$ where the probability of the node being in idle state is almost zero. Based on these results, we claim that the proposed DTMC model behaves as expected and use it for modelling the collusion between two nodes.

B. Global System Model: We observe from Fig. 4, a_i increases linearly with q and reaches a saturation at around normalised load of 0.5 for $W_0 = 8$. a_i reaches saturation at lower loads as W_0 increases. Also, saturation value of a_i decreases with increase in W_0 as the node spends more time waiting for transmission or being idle.

C. Attacker-Detector Model: Fig. 5 depicts the estimated attackers' distribution (Eqn. (13)) as a 3-D mesh curtain plot for $W_0 = 32$ and $q_1 = q_2 = q_3 = 0.7$. We can clearly see that smaller back-off values of Node 2 and Node 3 are more probable than higher values which does not happen if both the attackers were obeying the IEEE 802.11 protocol. This clearly gives advantage for the attackers for using the estimated distribution. As we can see from Fig. 6, the degree of misbehaviour has a significant effect on the detection delay. It requires more samples and correspondingly, more time, to detect collusion for lower values of misbehaviour coefficient which conforms to the results obtained for the saturation case in [6].

D. Experimental Details:-

Based on the above analysis, the optimal attacker distribution is found and the expected value of the sample size is estimated. Table II reports the analytical results obtained using MATLAB. The values of λ and μ indicate the optimal attacker distribution for different loads.

TABLE II: Attackers' distribution parameters and estimated sample size for different loads for $W_0 = 8$. Note that E(N) indicates the estimated sample size

q_1	q_2	q_3	λ	μ	E(N)
0.2	0.2	0.2	2.609020	-4.472466	156.45
0.7	0.7	0.7	2.620430	-2.271771	159.95
0.8	0.1	0.1	2.599258	-4.706785	153.60
0.1	0.8	0.8	2.621764	-5.255232	160.40
0.1	0.1	0.8	2.701894	-5.631417	188.10
0.9	0.05	0.05	2.592955	-8.046555	151.80

To validate the estimate, we have generated samples from the attacker distribution using a statistics library called UNU.RAN [14]. Using this library, samples are generated according to the bivariate attacker distribution using the HITRO (HIT-and-run sampler with Ratio-Of-uniforms) Markov chain method of statistical sampling. The generated samples are input to SPRT, in MATLAB, which is setup with the parameters corresponding to attacker distribution under test. For different sets of samples, the experiment is conducted and on an average, the sample size needed approaches the estimated value.

VI. Conclusion and Future Work

To summarize, we have explored the problem of back-off manipulation at the MAC layer by malicious colluding adversaries in a non-saturated environment from an analytical perspective. We proposed a non-linear optimization model based on results from a sequential statistical inferencing test such as SPRT to understand the attacker-detector scenario. To evaluate the optimal joint attacker distribution, a measure for fairness in the network was provided by modelling each node as a Discrete Time Markov Chain (DTMC). The fairness condition was expressed as a constraint to the optimization model. We then got an analytical estimate of the sample size for different loads. We then presented validation and verification results got from simulations done in MATLAB. As possible extension of this work, the number of nodes can be increased and the scenario of collusion of any two attackers in this setting can be studied. Also, we plan to extend the analysis for handling finite number of retransmissions (contention window doubling) in our future work.

References

[1] V. Gupta, S. Krishnamurthy, and M. Faloutsos, "Denial of service attacks at the MAC layer in wireless ad hoc networks," in *Proceedings of MILCOM*, 2002.

[2] J. Bellardo and S. Savage, "802.11 denial-of-service attacks: real vulnerabilities and practical solutions," in *Proceedings of USENIX Security Symposium, San Antonio, TX*, June 2003.

[3] P. Kyasanur and N. H. Vaidya, "Detection and handling of MAC layer misbehavior in wireless networks," in *Proceedings of International Conference on Dependable Systems and Networks*, 2003.

[4] M. Raya, J.-P. Hubaux, and I. Aad, "DOMINO: A system to detect greedy behavior in IEEE 802.11 Hotspots," in *Proceedings of MobiSys*, 2004, pp. 84–97.

[5] V. N. Lolla, L. K. Law, S. V. Krishnamurthy, C. Ravishankar, and D. Manjunath, "Detecting MAC layer back-off timer violations in mobile ad hoc networks," in *Proceedings of International Conference on Distributed Computing Systems*, 2006.

[6] S. Radosavac, A. A. Cardenas, J. S. Baras, and G. V. Moustakides, "Detecting IEEE 802.11 MAC layer misbehavior in ad hoc networks: Robust strategies against individual and colluding attackers," *Journal of Computer Security*, vol. 15, pp. 103–128, 2007.

[7] A. Wald and J. Wolfowitz, "Optimum character of the sequential probability ratio test," *Annals of Mathematical Statistics*, vol. 19, no. 3, pp. 326–339, 1948.

[8] T. Kuang and C. Williamson, "Hierarchical analysis of RealMedia streaming traffic on an IEEE 802.11b wireless LAN," *Computer Communications, Elsevier*, vol. 27, pp. 538–548, 2004.

[9] *IEEE standard for wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications*, 1997, P802.11.

[10] D. Malone, K. Duffy, and D. Leith, "Modeling the 802.11 distributed coordination function in nonsaturated heterogeneous conditions," *IEEE/ACM Transactions on Networking*, vol. 15, no. 1, pp. 103–128, 2007.

[11] G. Bianchi, "Performance analysis of the IEEE 802.11 Distributed Coordination Function," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 535–547, 2000.

[12] A. Wald, *Sequential Analysis*. Wiley, New York, 1947.

[13] L. E. Elsgolc, *Calculus of Variations*. Pergamon Press, Oxford, London, 1961.

[14] *UNU.RAN - Universal Non-Uniform RANdom number generators*, http://statmath.wu-wien.ac.at/unuran/.

Appendix

Simplification of the expression $Pr(Y_1 < Y_2)$ where Y_1 and Y_2 are two random variables is given below. Y_1 is an unknown random variable with pdf $f(y_1)$ taking values in $[0, W_0]$ whereas Y_2 is a uniformly distributed in $[0, W_0]$, where W_0 is some constant.

$$\begin{aligned} Pr(Y_1 < Y_2) &= \int_0^{W_0} Pr(Y_2 > Y_1 | Y_1 = y_1) f(y_1) dy_1 \\ &= \int_0^{W_0} Pr(Y_2 > y_1) f(y_1) dy_1 \\ &= \int_0^{W_0} (1 - Pr(Y_2 \leq y_1) f(y_1) dy_1 \\ &= 1 - \int_0^{W_0} (y_1/W_0) f(y_1) dy_1 \\ &= 1 - E(Y_1)/W_0 \end{aligned}$$

In the context of this paper, Y_2 refers to the legitimate node, Y_1 may refer to any attacker node and W_0 represents the maximum back-off window size at stage 0 of the IEEE 802.11 DCF protocol.