Throughput of TCP over Cognitive Radio Channels

Sudheer Poojary*, Akash Agrawal†, Bhoomika Gupta‡, Archana Bura*, Vinod Sharma*
*{sudheer,vinod}@ece.iisc.ernet.in, Department of ECE, Indian Institute of Science, Bangalore, India
†akasha@arubanetworks.com, Aruba Networks, Bangalore, India
‡Electronics and Communication Engineering, Sardar Vallabhbhai National Institute of Technology, Surat, India

Abstract—We study TCP performance over ON-OFF channels, i.e., channels which are busy/available for transmissions intermittently. The data may also be subject to random errors. Such situations arise in cognitive radio networks when secondary users opportunistically use the channel when the licensed user is not using it. We use a Markov model to evaluate the TCP throughput and probability of retransmission timeout when the ON-OFF periods are exponentially distributed and then extend it to phase-type ON and phase-type OFF periods. We then generalize the Markovian model to a more general ON-OFF duration distributions and also to multiple TCP connections.

Keywords—TCP, ON-OFF channels, cognitive radio.

I. INTRODUCTION

TCP provides reliable, in-order, end-to-end data transfer along with network congestion and flow control and fair allocation of resources [1]. TCP along with UDP is the dominant transport protocol over the Internet. Typical TCP variants, viz., TCP Reno, TCP New Reno were designed for wired links and treat packet losses as an indication of congestion and reduce their transmission rate on detection of packet loss. With the increasing usage of wireless links, e.g., cellular communication and wifi, packet losses may be due to link errors or due to collision with other transmissions. The misinterpretation of these losses as congestion leads to severe underutilization of network capacity. Besides packet losses, in certain scenarios such as cognitive radio networks, the channel may not always be available [2], [3]. These lead to frequent timeouts and the TCP performance is adversely affected.

TCP performance has been thoroughly investigated. In [4], the authors study the performance of TCP with random losses in WANs. TCP Reno throughput expressions are derived in [5] for Bernoulli losses and in [6] for more general loss processes. In [7], the authors propose a network layer modification to improve TCP performance over wireless links. Though TCP performance has been extensively studied in a wide variety of environments, there are few analytical models of TCP with ON-OFF channel behaviour which arise in cognitive networks.

The primary motivation for this work is cognitive radio (CR) networks [8]. Cognitive radio is an emerging technology which intends to use the spectrum more efficiently. Huge portions of the spectrum are reserved and lay underutilized. Cognitive radio permits devices called secondary devices to utilize the channel when the primary user of the channel (the user to whom the channel is licensed) is not utilizing the channel. One can model this behaviour with an alternating ON-OFF channel, with ON periods corresponding to the primary being inactive and hence the channel being available for transmission and OFF durations correspond to primary busy periods where the channel is unavailable to the secondary for transmission. In CR networks, an example use case for TCP over ON-OFF channel is cellular data boost [9] in cellular networks where non-real time or delay tolerant traffic such as email, ftp etc, (which use TCP) can be offloaded from the cellular network to meet QoS requirements of delay sensitive traffic. Besides CR networks, there are other networks where the links can be modeled as ON-OFF channels. Intermittent loss of connectivity also happens in cellular networks due to hand-offs, in mobile ad-hoc networks [10] due to link failures, in satellite networks [11] and in 802.11 networks due to collisions.

The performance of secondary users in a CR network has been studied in [2] and [3]. In [2], the authors compute the channel availability probability and the throughput for CR nodes in MIMO CR networks. In [3], the authors use Markovian models to derive the stability condition for the secondary users in a multi-channel CR network. In [12], the authors discuss properties and research challenges posed by cognitive radio. They suggest changes that need to be incorporated into transport layer protocols for operation over CR channels. In [13], the authors compare the performance of TCP SACK, TCP New Reno and TCP Vegas over dynamic spectrum access links using ns2 simulations. In [14], the authors study the effect of spectrum sensing duration, primary user interference and channel bandwidth variation on different TCP variants using simulations. In [15], the authors use relay selection, power allocation and adaptive modulation and coding schemes to improve secondary user TCP performance over CR channels. In [16], the authors model the system as a M/G/1 queue with the primary users getting priority over the secondary users and provide expressions for throughput for data traffic and mean delay for voice traffic of secondary users. Transport layer protocols for cognitive radio have been developed in [17]–[20].

From above we see that there is considerable literature on performance analysis of secondary users in a CR network. However most analytical results address the problem at the MAC layer ignoring the impact of TCP dynamics and studies of TCP behaviour are mostly simulation-based. In this paper, we provide a theoretical model for TCP connections over a CR channel. We compute throughput and probability of retransmission timeout for a secondary TCP connection. Our work complements [16]. In [16], the authors consider the case where ON and OFF durations are of the order of round trip time (RTT) and hence they ignore TCP timeouts. In this paper we consider the case where the ON and OFF durations are larger than RTT where the effect of TCP timeouts cannot be ignored. Such a scenario can often happen in CR networks and then the timeouts can significantly affect the TCP throughput.

The paper is organized as follows. In Section II, we describe our system model. In Section III, we develop a Markovian model for TCP behaviour in an ON-OFF channel with exponential ON and OFF periods. In Section IV, we develop models for ON and OFF periods with more general
II. System Model

We consider a single TCP New Reno flow going through an ON-OFF channel. The TCP flow could be of a secondary device in a cognitive radio network using opportunistic spectrum sharing so that the OFF periods correspond to the primary using the channel and the ON periods correspond to the time when the channel is available for the secondary user. The overall RTT of the TCP flow is \( R \) sec. Any transmission (of a window load of packets) when the channel is ON goes through, although the packets may experience transmission error. The next window of packets will be transmitted after RTT, i.e., \( R \) secs. However a transmission attempted during an OFF period results in loss of all the packets of that window and hence causes a retransmission timeout (RTO) and the TCP source has to wait for its RTO timer to expire before it can attempt the next retransmission. It is possible that the channel becomes OFF while a secondary transmission is going on. Such an event is likely if the average ON and OFF periods are of the order of RTT or lesser. This case has been studied in [16]. Here we consider the case where the average ON and OFF periods are larger than the RTT of the flow. If during the ON period, the transmission error probability is small, after an error the next transmitted packets (i.e., packets in the window preceding the packet in error) may be received successfully causing transmission of duplicate ACKs. In response to these losses, the secondary TCP reduces its window size. In Figure 1, we show the window size evolution for a TCP flow over an ON-OFF channel. In the ON period, the TCP window size increases, however random losses over the channel cause reductions in the window size even during the ON periods.

Let \( S_k \in \{0, 1\} \) be the channel state at the \( k^{th} \) transmission of a window load of packets, where 1 corresponds to the channel being ON and 0 corresponds to the channel being OFF. Let \( D_k \in \mathcal{D} \) denote the duration between the \( k^{th} \) transmission and the \((k+1)^{st}\) transmission. The duration between successive transmissions in the ON period is equal to the RTT, \( R \), of the flow. However, on encountering a timeout, TCP uses a binary exponential back-off strategy. The first RTO is set to \( M = \max\{R, T_{\min}\} \), where \( T_{\min} \) is the minimum value that a timeout duration can be. If a failed transmission (i.e., one which leads to a timeout) is immediately followed by another, TCP doubles the timeout duration. The timeout duration is bounded above by \( T_{\max} \). Therefore \( \mathcal{D} = \{R, M, 2M, 4M, \ldots, T_{\max}\} \). The TCP timeout mechanism is illustrated in Figure 1.

Let \( W_k \) be the window size at the end of \( k^{th} \) transmission of a window of packets and \( T_k \) be the value for the corresponding slow start threshold of TCP. The window evolution of TCP new Reno is as given below. If there is no packet loss

\[
W_{k+1} = W_k + 1, \text{ if } W_k \geq T_k = 2W_k, \text{ if } W_k < T_k.
\]

If there is a packet loss, TCP retransmits the lost packet and reduces the window size. If TCP detects the loss through duplicate ACKs, then it reduces the window size by half and sets the slow start threshold to that value. This is called recovery through fast retransmit. In that case,

\[
W_{k+1} = \frac{W_k}{2}, \text{ } T_{k+1} = \frac{W_k}{2}.
\]

If the loss is detected through a timeout (this will happen probably because channel is OFF), we have

\[
W_{k+1} = 1, \text{ } T_{k+1} = \frac{W_k}{2}.
\]

The window size \( W_k \) and the threshold \( T_k \) are restricted to \( W_{\max} < \infty \).

In the rest of the paper, we develop theoretical models for TCP over an ON-OFF channel and compute probability of retransmission timeout and throughput and compare to ns2 simulations. We then show how these results can be used when there are multiple TCP connections.

III. Analysis for Exponential ON-OFF

We assume that the OFF and ON periods are iid exponential with parameters \( \lambda_0 \) and \( \lambda_1 \) respectively. The PU activity is usually modeled as exponential [12]. We will generalize these assumptions in the following sections.

We denote the state of the system at the beginning of the \( k^{th} \) transmission of a window workload by \( \{S_k, D_k, W_k, T_k\} \). Since we assume that the OFF and ON durations are exponential, the process, \( \{S_k, D_k, W_k, T_k\} \) forms a finite state, discrete time Markov chain. In Figure 2, we illustrate the single-step transitions from generic ON and OFF states.

Let \( S(t) \) be the state of the channel at time \( t \) with \( S(t) = 0 \) if channel is OFF and 1 if it is ON. In the ON period, packets can be dropped due to losses on the wireless channel with probability \( p \) independently of others. Let \( P_t(i, j) \) be the probability of channel state \( S(t) \) at time \( t \) being in state \( j \) given that it was in state \( i \) at time 0. Once we know \( P_t(i, j) \), we can find the transition probabilities for the Markov chain \( \{S_k, D_k, W_k, T_k\} \). Proposition 1 provides \( P_t(i, j) \) explicitly via renewal theory.

**Proposition 1.** We have for \( i \neq j \),

\[
P_t(i, i) = \frac{\lambda_j + \lambda_0 e^{-(\lambda_0+\lambda_1)t}}{\lambda_0 + \lambda_1},
\]

**Figure 1.** Cognitive radio channel with one secondary TCP source

**Figure 2.** Single-step transitions for the \( \{S_k, D_k, W_k, T_k\} \) Markov chain between the \( k^{th} \) transmission and the \((k+1)^{st}\) transmission of windows, we have
Proof: Let \(X_i\) be exponentially distributed with parameter \(\lambda_i, i = 0, 1\) and \(X_0\) independent of \(X_1\). Let \(Y = X_0 + X_1\). Let us denote the density function of \(Y\) by \(f_Y\). By renewal theory arguments [21], we have,

\[
P_t(0, 0) = P(X_0 > t) + \int_0^t P_{t-x}(0, 0) f_Y(x) dx.
\]  

(5)

Taking Laplace transform, we have

\[
\hat{P}_s(0, 0) = \frac{1}{\lambda_0 + s} + \hat{P}_s(0, 0) \hat{f}_Y(s).
\]  

(6)

The Laplace transform, \(\hat{f}_Y\) of \(f_Y(t)\) is given by,

\[
\hat{f}_Y(s) = \frac{\lambda_0 \lambda_1}{(\lambda_0 + s)(\lambda_1 + s)}.
\]  

(7)

From equations (6) and (7), we have

\[
\hat{P}_s(0, 0) = \frac{\lambda_1}{\lambda_0 + \lambda_1} s + \frac{\lambda_0}{\lambda_0 + \lambda_1} s + \frac{1}{\lambda_0 + \lambda_1}.
\]  

(8)

Taking inverse Laplace transform, we have

\[
P_t(0, 0) = \frac{\lambda_1}{\lambda_0 + \lambda_1} + \frac{\lambda_0}{\lambda_0 + \lambda_1} e^{-(\lambda_0 + \lambda_1)t}.
\]  

(9)

Similarly we can derive expression for \(P_t(1, 1)\).

The Markov chain, \(\{(S_k, D_k, W_k, T_k)\}\) is finite. The state \(1, R, 1\) can be reached from any state in the state space with positive probability. Therefore the state \(1, R, 1\) is positive recurrent and any state that can be reached from \(1, R, 1\) is positive recurrent and the remaining states are transient. Also if the probability of packet loss during the ON period is greater than 0, the state \(1, R, 1\) has a self loop. Therefore the Markov chain is aperiodic and has a unique stationary distribution, \(\pi\) which can be computed. Also, starting from any initial state the chain converges exponentially to the stationary distribution in total variation.

The probability of retransmission timeout, \(P_o\), i.e., the fraction of packets that are timed out is

\[
P_o = \frac{E_\pi[1(\Delta > 0)]}{E_\pi[W]},
\]  

(10)

where \(1(A)\) is an indicator function of set \(A\) and \(E_\pi\) denotes mean under stationarity. The throughput, \(\lambda\) (in packets/sec) of the TCP connection is computed using Palm calculus [22],

\[
\lambda = \frac{E_\pi[W 1(\Delta > 1)]}{E_\pi[\Delta]}.
\]  

(11)

A. Extension to channels with non-negligible queuing

The above analytical model assumes that the round trip time for the secondary TCP connection is constant. If the queuing is non-negligible, the model may not be accurate. When the queuing is non-negligible, we can use the above model with a minor modification. In the state space of \(\{S_k, D_k, W_k, T_k\}\) process, if the channel is ON at the end of \(k\)th transmission, i.e., if \(S_k = 1\), we set \(R_k = \max\{\Delta, \frac{W_k}{p}\}\), where \(\Delta\) is the constant component of the round trip time which includes the propagation delay and processing delays at the nodes and \(p\) is the link speed of the channel in packets/sec. Our simulation results justify approximations made in the analytical model. Such an approximation has been used before in [23], [24].

B. Simulation Results

We now compare the probability of timeout and throughput obtained from our model with ns2 simulations. We have modified ns2 code so as to simulate an ON-OFF channel. We generate a sequence of alternate ON and OFF periods and drop all packets that arrive in the OFF period. The packet sizes are 1050 bytes. We set the link speed of the ON-OFF channel to 5 Mbps. The other links that the flow traverses have link speeds 1 Gbps. In practice, the ON-OFF channel could be a wireless link connecting a wireless device to an access point or a base station. The base station/access points are then connected to the Internet through well-provisioned optical fiber links.

We denote the fraction of time that the channel is OFF by \(\alpha\). We plot the probability of retransmission timeout (RTO), \(P_o\) and the secondary TCP throughput in Figure 3. We set \(\alpha = 1/3\) and vary the average channel OFF duration, \(E[Y_{\text{off}}]\). The probability of packet transmission error in the idle duration is set to 0.01. The maximum window size, \(W_{\text{max}}\) is 100 packets. We see that as \(P_o\) increases, the probability of retransmission timeout increases and throughput decreases. Our model results match well with ns2 simulations and the errors are less than 5% in most cases. The errors are larger when the average ON and OFF durations are of the order of RTT. However, even for these cases the errors are less than 11%. Our analytical model results show that in an ON-OFF channel we can not approximate TCP throughput by multiplying the throughput expressions for TCP Reno (which may be found in [5]) by \((1.0 - \alpha)\) where \(\alpha\) is the fraction of time that the channel is busy.

In Figure 4, we see the effect of packet error probability \(p\) on the probability of retransmission timeout and the throughput of the flow. The round trip time is 0.1 sec. We see that for fixed \(\alpha\) and RTT, \(P_o\), the probability of retransmission timeout decreases and throughput increases with increase in the average OFF duration. Our model results match well with ns2 simulations and the errors are less than 8%.

In Figure 5, we show the effect of link speed on the probability of retransmission timeout and the throughput of the TCP flow. The RTT is 0.1 sec and CR channel link speeds are set to 5 Mbps and 1 Mbps. We see that our model
Fig. 5. Effect of link speed and state $S_X$ to the case where both ON and OFF periods are phase type distributions. The model can be extended to the beginning of the other. Also, as before the state of the system at the period and OFF periods are both i.i.d. and are independent denoted by $(S_k, D_k, W_k, T_k)$.

When ON and OFF periods are phase type, the channel state $S(t)$ can be taken as a finite state Markov chain with state space $\mathcal{X}_0 \cup \mathcal{X}_1$. When $S(t) \in \mathcal{X}_0$ then the channel is OFF and when $S(t) \in \mathcal{X}_1$ then it is in ON state. Also, the process $(S_k, D_k, W_k, T_k)$ is a discrete time Markov chain (embedded in $S(t)$). Let $P_i(j)$ be the probability of channel state $S(t)$ at time $t$ being $j$ given that it was in state $i$ at time 0. Once we know $P_i(i, j)$, we can find the transition probabilities for the Markov chain $\{(S_k, D_k, W_k, T_k)\}$.

Suppose $Q$ is the transition rate matrix for the CTMC $\{S(t)\}$, and let $\psi := \{\psi_0(i) : i \in \mathcal{X}_0 \cup \mathcal{X}_1\}$ denote the probability distribution of $S(t)$ at time $t > 0$. It satisfies $\psi_t = \psi_0 Q^t$, where $\psi_0$ is the initial distribution. Also, $P_i = e^{\psi t}$. It can be computed using the expm command in MATLAB or by using eigenvalue decomposition techniques [25]. When the state space of $S(t)$ is large, the exact computation of $e^{\psi t}$ may be difficult. We may then use some approximations [26].

Now the Markov chain $\{(S_k, D_k, W_k, T_k)\}$ is finite. We can show that the state $(i, R, 1, 1)$ for any $i \in \mathcal{X}_1$ is a positive recurrent state. If the probability of packet loss during the ON period is greater than 0, then the state $(i, R, 1, 1)$ also has a self-loop. Therefore the Markov chain is aperiodic and has a unique stationary distribution, $\pi$ which can be computed. Also, starting from any initial state the chain converges exponentially to the stationary distribution in total variation.

We now develop theoretical results for (a) phase type OFF and general ON periods under certain conditions and (b) phase type ON and general OFF periods under certain conditions. We will show stationarity of the process $\{(S_k, D_k, W_k, T_k)\}$ in these cases. The stationary behaviour can then be used to compute the throughput and the probability of timeout for a TCP flow on an ON-OFF channel.

A. Phase type OFF periods

In this section, we consider the case when OFF periods are i.i.d. phase type while the ON periods are more general. Let us denote by $U_n$ the time when the secondary TCP makes the $n^{th}$ transmission attempt of a window load of packets. Let $U_0 = 0$. Therefore $U_n = \sum_{n=1}^{N_t} D_n$, where $D_n$ is the time duration between the $n^{th}$ and $(n+1)^{th}$ transmission attempt. Let $\{X(t)\}$ be a CTMC with finite state space $\{1\} \cup \mathcal{X}_0$, where $0 \notin \mathcal{X}_0$. The set $\mathcal{X}_0$ is irreducible and state 1 is an absorbing state reachable from every state in $\mathcal{X}_0$. The sojourn times in each state $i \in \mathcal{X}_0$ are exponential with parameter $\lambda_i$, $0 < \lambda_i < \infty$. Let $\lambda = \max_{i \in \mathcal{X}_0} \lambda_i < \infty$. We define the channel state process $S(t)$ as follows. Suppose $S(t)$ enters ON state at time $t$ (denoted by $S(t) = 1$). It will stay in ON state with a general distribution ($Y_{off}$ will denote a random variable with that distribution). At time $t + Y_{off}$, it will enter the OFF period. The duration of the OFF period is given as follows. Let $X(0) = i_0 \in \mathcal{X}_0$ where $i_0$ is a fixed state of $\{X(t)\}$. The OFF period of $S(t)$ will equal the time $\{X(t), t \geq 0\}$ takes to reach state 1. Let this time be $X$. Then $S(t)$ stays in OFF state till time $t + Y_{on} + X$ and then switches to ON state. At time $t + Y_{on} + s$, $S$ takes the value $X(s)$, for $0 \leq s \leq X$. The ON-OFF periods of $S(t)$ alternate with durations independent of each other with distributions specified above.

Let $Q$ be the generator matrix for the CTMC $X(t)$ when in $\mathcal{X}_0$. By uniformization [27], before $X(t)$ gets absorbed, the transition function of $\{X(t)\}$,

$$P_i(i,j) = \sum_{n=0}^{\infty} e^{-\lambda_i} P_{ij}^n \frac{(\lambda_i)^n}{n!},$$

where $P_{ij} = Q_{ij}$ for $i \neq j$ and $P_{ii} = 1 - \sum_{j \neq i} Q_{ij}$. Consider the state $(S_k = i_0, D_k = 2M, W_k = 1, T_k = 1)$ where $i_0 \in \mathcal{X}_0$. Once the process $\{(S_k, D_k, W_k, T_k)\}$ hits the
state, \((i_0, 2M, 1, 1)\), the future process evolution is independent of the past. Thus, \(\{(S_k, D_k, W_k, T_k)\}\) is regenerative. Let \(\mathcal{N}\) denote the number of transmission attempts made by TCP in a regeneration cycle. We have the following result.

**Proposition 2.** If the ON periods are i.i.d. with

\[
P(Y_{on} \leq s + R|Y_{on} \geq s) \geq \epsilon_0 > 0
\]

and \(0 < P(Y_{on} \leq R) < 1\), then for all \(\alpha > 0\), \(E[N^\alpha] \leq \infty\) and \(N\) has a finite moment generating function (mgf) in a neighborhood of 0. Also \(\mathcal{N}\) is aperiodic and the stochastic process \(\{(S_k, D_k, W_k, T_k)\}\) converges in total variation, exponentially to its unique stationary distribution

\[
P_x((S_k, D_k, W_k, T_k) \in A) = \frac{E[\sum_{k=0}^{N} 1_A(S_k, D_k, W_k, T_k)]}{E[N]},
\]

where \(k = 0\) is a regeneration epoch.

**Proof:** We first prove that from any arbitrary state in the OFF period the process can visit the ON state in one step with probability \(\geq \epsilon_1 > 0\).

**Step 1:** Since \(X_0\) is irreducible and 1 is reachable from all states, for all \(i \in X_0\) and all \(j \in X_0 \cup \{1\}\), there exists \(n(i, j) > 0\) such that the probability \(P_{i,j}^{n(i,j)}\) of hitting state \(j\) in \(n(i, j)\) steps starting from state \(i\) for the DTMC of \(\{X(t)\}\) with transition matrix \(P_{ij}\), is strictly positive in equation (13). Let

\[
\epsilon(i, j) = e^{-\lambda x} P_{i,j}^{n(i,j)}(M_{ij})^{n(i,j)}.
\]

Then \(P_{i,j}(i, j) > \epsilon(i, j) > 0\) for all \(i \in [t_1, t_2]\) where \(0 < t_1 < R < T_{\text{max}} < t_2 < \infty\). Let \(\epsilon_1 = \min_{i}(\epsilon(i, 1)) > 0\), where \(i \in X_0\). Therefore \(P_{i,j}(i, 1) \geq \epsilon_1\) for all \(i\) and for all \(t \in [t_1, t_2]\).

Then,

\[
P(S_{k+1} = y, D_{k+1} = R|S_k = x, D_k = d) \geq \epsilon_1,
\]

for all states \((x, d)\) in the state space of the process \(\{(S_k, D_k)\}\) with \(x \in X_0\).

We note that the inequality (17) is true if we replace \((S_{k+1} = 1, D_{k+1} = R)\) by the term \((S_{k+1} = y, D_{k+1} = 2d)\) for any \(y \in X_0\) possibly with a different \(\epsilon_1(y) > 0\) as a lower bound. We will use this fact later in the proof.

**Step 2:** Let us denote by \(Y^t\) the age of the ON period at \(U_k\), i.e., the time elapsed since the ON period started, when \(S_k = 1, D_k = R\) at \(U_k\) and let \(F_{Y^t}\) be its cdf. Now

\[
\sum_{y \in X_0} P(S_{k+1} = y, D_{k+1} = M|S_k = 1, D_k = R) \geq \int_{\{s\}} \left(\int_{0}^{R} P(Y_{on} = s + u|Y^t = s) \right. \\
\left. P(Y_{off} > R - u|Y^t = s)\right)\, du\, dF_{Y^t}(s)\]

\[
\geq \int_{\{s\}} \left(\int_{0}^{R} P(Y_{on} = s + u|Y_{on} \geq s) \right. \\
\left. P(Y_{off} > R|Y_{on} \geq s)\right)\, du\, dF_{Y^t}(s)\]

\[
= P(Y_{off} > R) \int_{s} P(Y_{on} \leq s + R|Y_{on} \geq s)\, dF_{Y^t}(s)\]

\[
> \epsilon_0 P(Y_{off} > R) \geq \epsilon_0 > 0.
\]

Fig. 6. The regeneration epoch \((i_0, 2M, 1, 1)\) can be hit with non-zero probability in finite time from any state in process \((S_k, D_k, W_k, T_k)\).

where the last inequality follows from (14) and \(P(Y_{off} > R) > 0\) because \(Y_{off}\) is phase type.

Thus with probability \(> \epsilon_0\) the process exits the ON period to visit the OFF period in one step. The inequality (17) is true for the state \((S_{k+1} = i_0, D_{k+1} = 2M)\) with some \(\epsilon_2 > 0\) as the lower bound. Therefore, \(P(S_{k+1} = i_0, D_{k+1} = 2M|S_k = y, D_k = M) \geq \epsilon_2 > 0\) for any \(y \in X_0\). This shows that the process \(\{(S_k, D_k, W_k, T_k)\}\) visits the state \((i_0, 2M, 1, 1)\) in a sequence of two steps from the ON state with probability \(> \epsilon_0 \epsilon_2 > 0\) (subfigure 1 of Figure 6).

**Step 3:** From steps 1 and 2, the process can visit the regeneration epoch \((i_0, 2M, 1, 1)\) from any state in the state space in less than three steps with probability \(> \epsilon = \epsilon_0 \epsilon_2 > 0\). Thus in particular, \(P(N = 3) > 0\). Consider random variable \(Z\) with distribution \(P(Z = 3k) = (1 - \epsilon)^{k-1}\epsilon\), for \(k \geq 1\). The random variable \(Z\) is stochastically larger than the regeneration length \(N_i\), i.e., \(P(Z \geq \beta) \geq P(N \geq \beta)\) for all \(\beta > 0\). Therefore, for all \(\alpha > 0\), \(E[N^\alpha] \leq E[Z^\alpha] \leq E[Z^\alpha] < \infty\). Also, \(N\) has a finite mgf in a neighborhood of 0.

We have shown that \(P(N = 3) > 0\). In subfigure 2 of Figure 6, we show that \(P(N = 4) > 0\). This shows that \(N\) is aperiodic. Thus \(\{(S_k, D_k, W_k, T_k)\}\) has a unique stationary distribution, (15) and converges to it in total variation from any initial distribution exponentially (because of finiteness of mgf of \(N\) in a neighbourhood of 0 [28]).

The condition (14) is satisfied by a general class of distributions called New Better than Used (NBU) [29] and also by phase type distributions. The NBU distributions are useful in reliability theory. They are also relevant in our case, as one would typically expect that the primary busy period starting afresh is likely to last longer than an ongoing busy period.

**B. Phase type ON periods**

We now develop theoretical results for phase type ON and general OFF periods. Let \(X(t)\) denote the phase of the ON period. The process \(\{X(t)\}\) is a CTMC with a finite state space \(\{0\} \cup X_1\) with 0 \(\notin X_1\). Let us denote by \(S(t)\) the state of the ON-OFF channel at time \(t\). If the channel is ON at time \(t\), then \(S(t) = X(t)\), else \(S(t) = 0\), where 0 is the absorbing state for the CTMC \(\{X(t)\}\). We assume \(X_1\) is irreducible and 0 is reachable from all states. The process \(\{(S_k, D_k, W_k, T_k)\}\) is regenerative with visits to the state \((i_1, R, 1, 1)\), \((i_1 \in X_1)\) (a fixed state), acting as regeneration epochs. We denote by \(N\) the number of transmission attempts made in a regeneration cycle.
Proposition 3. If the OFF periods are i.i.d. with $P(s + M < Y_{off} < s + 3M) \geq \epsilon > 0$ for all $s \leq R$ and

$$P(Y_{off} \leq s + d|Y_{off} \geq d) \geq \epsilon_1 > 0,$$  

(18)

for $d \in \{2M, 4M, 8M, \ldots, T_{max}\}$. Then for all $\alpha > 0$, $E[N^{\alpha}] < \infty$ and $N$ has a finite moment generating function in a neighborhood of 0. Also, $N$ is aperiodic and the stochastic process $\{(S_k, D_k, W_k, T_k)\}$ converges in total variation, exponentially, to its unique stationary distribution (15).

Proof: The proof is somewhat similar to that of Proposition (2) and is omitted due to lack of space.

The conditions in Proposition (3) are satisfied if $Y_{off}$ has a positive density on $[M, 3M]$ and has NBU distribution. It is also satisfied for phase type distributions.

In the proofs for Propositions (2) and (3), we do not assume random packet losses for convenience of notation. The propositions hold even in this case when the packet loss probability, $p < 1$ with a slight modification of proofs. These propositions show stationarity of the regenerative process modelling TCP behaviour in the setup of ON-OFF channels with more general ON and OFF distributions. This also ensures that the time averages for performance metrics such as throughput and probability of RTO converge to the stationary mean values.

C. Simulation Results

We now compare the probability of timeout and throughput obtained via the analytical model with ns2 simulations. The probability of timeout and throughput can be computed using equations (10) and (11) respectively with some modifications. For phase type ON, we replace the term $1_{\{S=1\}}$ by $1_{\{\text{ON}\}}$ and for phase type OFF process we replace $1_{\{S=0\}}$ by $1_{\{\text{OFF}\}}$. The simulation setup is the same as in Section III-B.

We consider two cases (a) ON and OFF periods are both exponentially distributed and (b) ON and OFF are both Erlang-3 distributed. For these experiments, we set RTT to 0.1 sec, $W_{max}$ to 100 and the ON-OFF channel link speed is set to 5 Mbps. The packets undergo Bernoulli random losses with probability 0.01. We vary the average busy duration, $E_{Y_{off}}$ keeping $\alpha$ fixed at 1/3. The results are shown in Figure 7. We see that our theoretical model results match well with simulations with errors less than 5% in most cases. The errors are larger when the average ON and OFF durations are of the order of RTT. However, even for these cases the errors are less than 10%.

In Figure 8, we consider the effect of the cognitive channel link capacity on probability of timeout and throughput. The ON and OFF periods are both Erlang-3 distributed, we set RTT to 0.1 sec and $W_{max}$ to 100. We consider link speeds of 1 Mbps and 5 Mbps. We see that our theoretical model results match well with simulations with errors less than 5% in most cases and always less than 11%.

V. MULTIPLE TCP CONNECTIONS

In this section, we consider the case when multiple secondary TCP connections share a CR channel. Thus the ON and OFF durations for these connections are same. However these connections are subject to different packet error rates as their channel gains may be different. Also these connections can possibly go through different routes; therefore they may have different round trip times. We first consider the case where the queuing delays are negligible. Then, the processes...
phase-type. then \((S_k, P_k, (W_{ij}^k, T_{ij}^k))_{j=1,2,\ldots, N}\) denotes the state of the system and forms a Markov chain, where index \(j\) represents the TCP connection \(j\). The RTT for the different flows at the end of the \(j^{th}\) RTT is given by \(\max\{\Delta, \sum \mu_i W_i\}\) where \(\mu_i\) is the bottleneck link speed (in packets/sec) and \(\Delta\) is the propagation delay (in sec). Our simulation results validate our model assumptions.

In the simulation setup, we have 3 secondary TCP connections with \(\Delta = 0.1\) sec. The maximum window size for all the flows is set to 20 packets. The ON-OFF periods are both exponentially distributed with average ON duration = 20 sec and average OFF duration = 10 sec. The ON-OFF channel has link speed of 2 Mbps and the other links are set to 1 Gbps. We compare the throughput and probability of timeout, \(P_o\) obtained using ns2 simulations with theoretical results in Table II. In this case, the errors are less than 13%.

### VI. Conclusions

We have developed an analytical Markovian model for a TCP flow over an ON-OFF channel with random losses. For the Markovian model, we assume that the ON and OFF periods are both exponential. We then extend our model to include phase-type ON and phase-type OFF periods. We have then considered the case when either the ON or the OFF periods are phase type. We prove the stationarity of the system using regenerative process theory. We compared the results viz., probability of retransmission timeout and secondary TCP throughput obtained using the theoretical models with ns2 simulations and showed that these match quite well. Finally, we have considered the scenario where multiple secondary TCP connections with different RTT and packet error probabilities share the ON-OFF channel. The theoretical results for this scenario also match well with simulations.

### References


