Opportunistic Cooperative Channel Access in Distributed Wireless Networks with Decode-and-Forward Relays

Zhou Zhang, Shuai Zhou, Hai Jiang, Senior Member, IEEE, and Lei Dong

Abstract—This letter studies distributed opportunistic channel access in a contention-based wireless network with decode-and-forward relays. If a source wins a contention, the channel state information in the first-hop channel is estimated, and a decision is made for the winner source to either give up the transmission opportunity and let all sources start a new contention, or transmit to the relay. Once the relay gets the traffic, it may have a sequence of probings of the second-hop channel. The optimal decision strategies for the two hops are derived in this letter. Simulation results show that our scheme is beneficial when second-hop channels have larger average signal-to-noise ratio.

Index Terms—Opportunistic transmission, optimal stopping.

I. INTRODUCTION

In a distributed wireless network, normally channel contention is adopted, for example, by using handshakes of request-to-send (RTS) and clear-to-send (CTS) before a data transmission. To efficiently utilize the wireless channel, it may be better if a source that wins the channel contention could give up its transmission opportunity when its channel is not good, i.e., it does not transmit upon reception of CTS, and thus, all sources immediately start a new channel contention. This idea is called distributed opportunistic channel access. The challenge is: how does distributed opportunistic channel access achieve optimality in terms of maximal system throughput? The challenge was addressed in [1]. A source first sends a probing packet (e.g., RTS) to its destination for channel contention. If the contention is successful, the destination estimates the channel signal-to-noise ratio (SNR) and feedbacks (e.g., by using CTS) to the source. If the channel SNR is less than a threshold value, which can be numerically calculated based on the users’ channel statistics, the source gives up its transmission opportunity; otherwise, the source transmits its traffic using the maximal achievable transmission rate of the probed channel. Ref. [2], [3], [4] are follow-ups of [1], considering that channel information is imperfect, multiple transmissions are possible, and there exists a bound on the average interval between two transmissions, respectively. For wireless relaying networks, distributed opportunistic channel access is investigated in [5] and [6] with amplify-and-forward relays and decode-and-forward (DF) relays, respectively. In specific, a distributed DF relaying network is considered in [6], in which each source-destination pair is aided by a DF relay. If a source has traffic to send, it sends a probing packet, and based on reception of the probing packet, the information of the first-hop channel SNR (from the source to its relay and to the destination) is obtained. Then the source decides to 1) give up, or 2) transmit with direct link, or 3) continue to probe the second hop (from the relay to the destination). If the source decides to probe the second hop, the channel SNR of the second hop is estimated, and it is decided either to give up or to transmit (by using the direct link or the relay link, whichever has better utility).

Different from [6], here we propose another method for a distributed DF relaying network. A source also first sends a probing packet, and after the first-hop channel SNR is obtained, we propose that the source either gives up the transmission opportunity or utilizes the first-hop channel first (by transmitting to its relay and letting its relay wait for a good second-hop channel to forward the received traffic, referred to as relay-waiting). The rationale for the relay-waiting is: if the first-hop instantaneous SNR is good but the second-hop instantaneous SNR is bad, the relay-waiting can still exploit the good first-hop channel, while the scheme in [6] is very likely to give up, thus wasting the good first-hop channel. In this letter, we derive an optimal strategy for our proposed scheme, and show that our scheme is beneficial when the second-hop channels have larger average SNR.

II. SYSTEM MODEL

Consider a distributed DF relaying network with $M$ source-destination pairs. Each source-destination pair is assigned a relay. First consider the case with direct links from sources to destinations. Similar to [6], to probe the first-hop channels, a source can send a probing packet. If there is no collision, the probing packet is received by both its relay and its destination. By reception of the probing packet, the relay and the destination can estimate the channel SNRs from the source to themselves. Then the relay reports its channel SNR information to the destination, and the destination makes a decision for the first hop (give up or transmit). For this case, by reception of the reporting message from the relay, the destination can actually estimate the channel SNR from the relay to itself, and thus, the destination has complete channel SNR information for the two hops: from the source to the relay, from the source to itself, and from the relay to itself. Then the destination can calculate the achievable end-to-end transmission rate denoted as $R$ between the source and itself. Therefore, although the communication from the source to the destination is with two hops, it can be treated as a virtual one-hop communication with achievable rate $R$. So the same method as that in [1] (which deals with single-hop networks) can be used to find an optimal opportunistic channel access strategy. Therefore, in this letter, we investigate the case without direct link between any source-destination pair. For source-destination pair $i$, channels in the first hop (from source $i$ to its relay) and the second hop (from the relay to destination $i$) follow Rayleigh fading with average received SNR being $\eta_i$ and $\rho_i$, respectively.

Similar to [1] [5] [6], the $M$ sources use a channel contention procedure as follows. At a minislot with duration $\sigma$, this work was supported in part by the TELUS Communications Company, Canada, and by the Natural Sciences and Engineering Research Council of Canada. Z. Zhang and S. Zhou were with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada (e-mail: zzhang6, szhou4@ualberta.ca). H. Jiang is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada (e-mail: hai1@ualberta.ca). L. Dong is with the State Key Laboratory of Integrated Service Network, Xidian University, Xi’an 710071, China (e-mail: donglei.xian@gmail.com).
each source sends an RTS with probability \( p \) to its relay. So at a
minislot, if no source transmits, i.e., the minislot is idle (the
probability is \( (1 - p)^M \)), then all sources start a new channel
contention in the next minislot; if more than one source sends
RTS (the probability is \( 1 - (1 - p)^M - MP(1 - p)^{M-1} \), a
collision happens, and thus, all sources start a new channel
contention after a time-out duration following the collision; if
only one source sends RTS (with probability \( MP(1 - p)^{M-1} \)),
we call the source a winner source. Define an observation as
the interval from the starting point of the channel contention
until a winner source appears (i.e., its RTS is successfully
received by its relay). The average duration of an observation
is
\[
\tau_1 = \frac{1}{MP(1 - p)^{M-1}} \cdot \sigma + \frac{1}{1 - (1 - p)^M - MP(1 - p)^{M-1}} (\tau_{\text{CTS}} + \tau_{\text{timeout}} + \tau_{\text{RTS}}),
\]
in which \( \tau_{\text{RTS}} \), \( \tau_{\text{timeout}} \), and \( \tau_{\text{CTS}} \) are
RTS and time-out durations, respectively.

At the end of an observation (say, observation \( n \)), the
winner source’s relay can estimate the channel SNR from the
winner source to itself by the RTS reception, and it decides from
two options: 1) option give-up: to give up the transmission
opportunity, and notify the source of the decision by sendin
g a CTS. This CTS is also received by other sources. Thus,
subsequently all sources can start a new contention. 2) option
stop: to stop the process and utilize the first-hop transmission
opportunity, and send back a CTS to notify the decision. In
the CTS, a transmission rate denoted as \( R_n \) is also indicated
for transmission from the winner source to the relay. Then
the winner source transmits for duration of a channel coherence
time denoted as \( \tau_g \) by using transmission rate \( R_n \) (the optimal
value of \( R_n \) is derived in Section IV). The subsequent action
of the relay is detailed in Section III.

For observation \( n \), if the winner source stops, denote reward
\( \lambda T_n \) as the total amount of traffic that is sent by the
winner source and received by its destination, and denote \( T_n \) as
the time duration from observation 1 until observation \( n \) plus
the time used for transmissions in the two hops. Denote \( N \) as
the stopping time, i.e., the winner sources in the first \( N - 1 \)
observations do not stop, and the winner source in the \( N \)th
observation stops. This later targets at an optimal stopping
time denoted as \( N^* \), which makes the system achieve the
maximal system throughput, i.e.,
\[
N^* = \arg \sup_{N \geq 0} \mathbb{E}[Y_N]/\mathbb{E}[T_N]
\]
where \( \mathbb{E}[\cdot] \) means expectation. \( N^* \) is also referred to as optimal
stopping strategy. Based on [7, Chapter 6], we can transform
problem (1) into a problem that maximizes reward \( \lambda T_N \)
with \( \lambda > 0 \). In specific, for \( \lambda > 0 \), an optimal strategy denoted
as \( N^*(\lambda) \) should be found, which maximizes expected reward
of the transformed problem:
\[
U(\lambda) = \sup_{N(\lambda) \geq 0} \mathbb{E}[Y_N(\lambda)] - \lambda \mathbb{E}[T_N(\lambda)].
\]
If we find a \( \lambda^* \) such that \( U(\lambda^*) = 0 \), an optimal strategy
of Problem (1) is given as \( N^*(\lambda) \) with \( \lambda = \lambda^* \) [7].

From [7], an optimal strategy of Problem (2) exists if the
following two conditions are satisfied: \( \mathbb{E}[\sup_n (Y_n - \lambda T_n)] < \infty \); and \( \limsup_{n \to \infty} (Y_n - \lambda T_n) = -\infty \) almost surely. From
Sections III and IV it can be seen that \( R_n \) should be finite
(because if \( R_n = \infty \), the probability for a relay to have a good
enough second-hop channel to forward its received traffic is zero). As \( Y_n \leq R_n \tau_d \), the first condition is apparently satisfied.
Since \( \limsup_{n \to \infty} T_n = \infty \), the second condition is also satisfied.
Therefore, an optimal strategy of Problem (2) exists.

An optimal strategy of Problem (2) has two parts: optimal
first-hop and second-hop strategies, discussed next.

### III. STRATEGY FOR THE SECOND HOP

Consider observation \( n \). Here we first try to find an optimal
strategy for the second hop, i.e., we assume that the winner
source, denoted as \( w(n) \), stops and transmits to its relay with
rate \( R_n \). For the second hop, the relay should find its best
strategy. The relay first sends an RTS to the destination, and
the destination estimates the second-hop channel SNR denoted
as \( r_g \) and feedbacks a CTS that includes the channel SNR
information, referred to as a channel probing. If the achievable
second-hop transmission rate, given as \( \log_2 (1 + r_g) \), is not
less than \( R_n \), the relay transmits to the destination by using
transmission rate \( R_n \) with duration \( \tau_d \); otherwise, the relay
decides to give up or to continue channel probing. If the relay
decides to give up, all sources start a new channel contention.
If the relay decides to continue channel probing, the relay
waits for channel coherence time \( \tau_d \) and has a new RTS-CTS
exchange with the destination (a new channel probing), and
transmits if the achievable second-hop transmission rate is not
less than \( R_n \), or decides to give up or to continue channel
probing otherwise. This procedure is repeated until the relay
transmits or gives up. So there are a sequence of decisions in
the second hop, which is challenging. To address this, we take
a new viewpoint for second-hop strategies, as follows.

Denote \( S_l \) as the second-hop strategy that the relay can have
up to \( l \) channel probeings of its channel to the destination. So
if the relay cannot find a second-hop channel realization with
achievable rate not less than \( R_n \), within \( l \) channel probeings, the
relay is forced to give up. Denote \( V^l(\lambda) \) (which is a function of \( \lambda \)) as the net reward of strategy \( S_l \). Therefore,
an optimal second-hop strategy should achieve net reward
\( \max \{E[V^1(\lambda)], E[V^2(\lambda)], \ldots, E[V^\infty(\lambda)] \} \). We have
\[
E[V^l(\lambda)] = \Pr[r_g \geq r_n](R_n \tau_d - \lambda \tau_d)
+ \Pr[r_g < r_n, r_g^2 \geq r_n](R_n \tau_d - 2 \lambda \tau_d) + \ldots
+ \Pr[r_g^l < r_n, \ldots, r_g^{l-1} < r_n, r_g^l \geq r_n](R_n \tau_d - l \lambda \tau_d)
+ \Pr[r_g < r_n, \ldots, r_g^{l-1} < r_n, r_g^l < r_n]
\times ((-l - 1) \lambda \tau_d - \lambda(\tau_{\text{RTS}} + \tau_{\text{CTS}})) \quad (3)
\]
in which \( \Pr[\cdot] \) means probability, \( \tau_{\text{CTS}} \) is CTS transmission
duration, \( \tau_{\text{RTS}} = \tau_{\text{RTS}} + \tau_{\text{CTS}} + \tau_d \) is the time cost for probing
and waiting (or transmission) in the second hop, \( r_g \neq 2R_n - 1 \)
is the minimum required SNR of the second hop for achievable
transmission rate \( R_n \), and \( r_g^1, r_g^2, \ldots, r_g^l \) are channel SNRs of
1st, 2nd, ..., \( l \)th channel probing of the relay. We have
\[
E[V^\infty(\lambda)] = \Pr[r_g \geq r_n](R_n \tau_d - \lambda \tau_d)
+ \Pr[r_g < r_n, r_g^2 \geq r_n](R_n \tau_d - 2 \lambda \tau_d) + \ldots
+ \Pr[r_g^l < r_n, \ldots, r_g^{l-1} < r_n, r_g^l \geq r_n](R_n \tau_d - l \lambda \tau_d)
+ \Pr[r_g < r_n, \ldots, r_g^{l-1} < r_n, r_g^l < r_n](E[V^\infty(\lambda)] - l \lambda \tau_d). \quad (4)
\]
Denote $F_{w(n)}(\cdot)$ as the cumulative distribution function of the SNR of the second-hop channel of winner source $w(n)$’s relay. From (3) and (4), we have

$$E[V^\infty(\lambda)] - E[V^l(\lambda)] = (F_{w(n)}(r_n))^l (E[V^\infty(\lambda)] - \lambda \tau_d). \quad (5)$$

From (5) with $l$ and (5) with $l + 1$, we have

$$E[V^{l+1}(\lambda)] - E[V^l(\lambda)] = (F_{w(n)}(r_n))^l (1 - F_{w(n)}(r_n)) (E[V^\infty(\lambda)] - \lambda \tau_d). \quad (6)$$

Thus, if $E[V^\infty(\lambda)] \leq \lambda \tau_d$, from (6) we have $E[V^l(\lambda)] \leq \lambda \tau_d$, which means the optimal second-hop strategy should be: the relay keeps probing the second-hop channel until the achievable rate is not less than $R_n$. On the other hand, if $E[V^\infty(\lambda)] > \lambda \tau_d$, from (6) we have $E[V^l(\lambda)] > \lambda \tau_d$, which means the optimal second-hop strategy should be: the relay probes the second-hop channel only once, and transmits if the achievable transmission rate is not less than $R_n$, or gives up otherwise.

IV. Strategy for the First Hop

Based on the derived optimal second-hop strategy, now we derive an optimal strategy for the first hop. In the first hop, at observation $n$, when the RTS of the winner source denoted as $w(n)$ is received by its relay and the first-hop channel SNR denoted as $r_f(n)$ is estimated, the decision is either give-up or stop (i.e., to transmit), whichever has higher reward. If the decision for the first hop is give-up, the net reward is $-\lambda \tau_d$ (since a CTS is needed to notify the decision); if the decision for the first hop is to transmit, the net reward is $\max \{E[V^1(\lambda)], E[V^2(\lambda)], ..., E[V^\infty(\lambda)]\} - \lambda \tau_c + x_n$, in which $\tau_{CTS} + \tau_d$ is cost time in the first hop: the relay uses a CTS to notify the source of the decision and the source transmits with $\tau_d$ duration.

First consider $E[V^\infty(\lambda)] < \lambda \tau_d$ for the second hop. Then from Section III, $\max \{E[V^1(\lambda)], E[V^2(\lambda)], ..., E[V^\infty(\lambda)]\} = \lambda \tau_{CTS} + \tau_d$. So the net reward of transmission in the first hop is $E[V^1(\lambda)] - \lambda \tau_{CTS} + \tau_d$. Since $E[V^\infty(\lambda)] < \lambda \tau_d$, from (5) with $l = 1$ we have

$$E[V^1(\lambda)] = (1 - F_{w(n)}(r_n)) E[V^\infty(\lambda)] + F_{w(n)}(r_n) \lambda \tau_d < \lambda \tau_d$$

which leads to $E[V^1(\lambda)] < \lambda \tau_{CTS} + \tau_d < -\lambda \tau_{CTS}$. In other words, the net reward of transmission in the first hop is less than the net reward of give-up in the first hop, and thus, the winner source will always give up in the first hop. Therefore, when we calculate the net reward of transmission in the first hop, we can ignore $E[V^\infty(\lambda)] < \lambda \tau_d$. Thus, we focus on $E[V^\infty(\lambda)] \geq \lambda \tau_d$, and from Section III, the net reward of transmission (stopping) in the first hop is

$$E[V^\infty(\lambda)] - \lambda (\tau_{CTS} + \tau_d)$$

(a) $= R_n \tau_d - \frac{1}{1 - F_{w(n)}(r_n)} \lambda \tau_2 - \lambda (\tau_{CTS} + \tau_d)$

(b) $\geq \tau_d \log_2 (1 + r_n) - \lambda \tau_{CTS} - \lambda \tau_d - \lambda e^{-\lambda w(n)} \tau_2$

which is from (4) with $l = 1$, and (b) is from $F_{w(n)}(r_n) = 1 - \frac{\lambda w(n)}{w(n)} (\text{Rayleigh fading})$ and $r_n \triangleq 2^{\rho w(n)} - 1$. The net reward (7) is not a monotonically increasing function of $r_n$. Next we set up an optimal $r_n$ that makes the net reward maximal.

For winner source $w(n)$, define function $\phi(x_{w(n)}) = \tau_d \log_2 (1 + x_{w(n)}) - \lambda \tau_{CTS} - \lambda \tau_d - \lambda e^{-\lambda w(n)} \rho w(n) \tau_2$, which is a concave function of $x_{w(n)}$. To find the optimal $x_{w(n)*}$, denoted $x_{w(n)*}$, that maximizes $\phi(x_{w(n)})$, we can solve $\frac{d\phi(x_{w(n)})}{dx_{w(n)}} = 0$, which leads to

$$\frac{\tau_d}{1 + x_{w(n)*}} \ln 2 = \frac{\lambda w(n)}{\rho w(n)} e^{-\lambda w(n)} \tau_2. \quad (8)$$

$x_{w(n)*}$ can be calculated from (8) numerically. So $r_n$ should be set to $x_{w(n)*}$ if feasible. However, it may not be feasible to set $r_n$ to be $x_{w(n)*}$ since $r_n$ should not be more than the first-hop channel SNR $r_f(n)$. Thus, overall we should set $r_n = \min \{r_f(n), x_{w(n)*}\}$ and $R_n = \log_2 (1 + \min \{r_f(n), x_{w(n)*}\})$.

Recall that an optimal stopping strategy of Problem (2) with $\lambda^*$ satisfying $U(\lambda^*) = 0$ is an optimal stopping strategy of Problem (1). So next we focus on optimal stopping strategy of Problem (2) with $\lambda^*$. Maximal expected reward $U(\lambda^*)$ of Problem (2) should satisfy an optimality equation [7]:

$$\sum_{w(n)} \frac{1}{M} \mathbb{E}(w(n)) \left[ \max \{\tau_d \log_2 (1 + \min \{r_f(n), x_{w(n)*}\}) - \lambda (\tau_{CTS} + \tau_d + e^{-\lambda w(n)} \rho w(n) \tau_2), U(\lambda^*) - \lambda^* \tau_{CTS}\} \right] - \lambda^* (\tau_{CTS} + \tau_d + e^{-\lambda w(n)} \rho w(n) \tau_2) \geq 0$$

in which $1/M$ is the probability for a source to be the winner source, and $\mathbb{E}(w(n))$ means expectation when $r_f(n)$ follows Rayleigh fading with mean SNR being $\rho w(n)$ (i.e., source $w(n)$ is the winner source). $\lambda^*$ can be calculated numerically from the optimal equation by setting $U(\lambda^*) = 0$.

Accordingly, an optimal stopping strategy for winner source $w(n)$ in the first hop is given as

$$N^*(\lambda^*) = \min \{n \geq 1: \tau_d \log_2 (1 + \min \{r_f(n), x_{w(n)*}\}) - \lambda^* (\tau_{CTS} + \tau_d + e^{-\lambda w(n)} \rho w(n) \tau_2) \geq 0\}$$

in which $x_{w(n)*}$ can be calculated from (8) with $\lambda = \lambda^*$.

For each winner source $w(n)$, the left-hand side of the inequality in (9) is a non-decreasing function of $r_f(n)$. Denote $\hat{r}_{f,w(n)}$ as the solution of $r_f(n)$ that makes two sides of the inequality in (9) equal. Then the optimal stopping strategy in the first hop is rewritten as $N^*(\lambda^*) = \min \{n \geq 1: \tau_d \geq \hat{r}_{f,w(n)}\}$. Overall, the proposed scheme works as follows. After a successful channel contention, winner source $w(n)$’s relay compares the first-hop channel SNR $r_f(n)$ with the threshold $\hat{r}_{f,w(n)}$: if $r_f(n) < \hat{r}_{f,w(n)}$, the first-hop decision is give-up; otherwise, the first-hop decision is stopping, i.e., source $w(n)$ transmits with rate $R_n = \log_2 (1 + \min \{r_f(n), x_{w(n)*}\})$. Subsequently the relay keeps probing the second-hop channel and comparing the second-hop achievable rate with $R_n$. If the second-hop achievable rate is larger than $R_n$, the relay transmits; otherwise, the relay keeps probing and comparing until it transmits eventually. Note that the values of $\hat{r}_{f,w(n)}$ and $x_{w(n)*}$ for $w(n) = 1, 2, ..., M$ can be calculated
V. PERFORMANCE EVALUATION

We use simulation to evaluate the proposed scheme, the scheme in [6] (with direct links not considered), and a naive scheme that utilizes all transmission opportunities (which is actually the scheme in [6] that never gives up in the first and second level decisions). The simulated network has 18 source-destination pairs, with other parameters as: system bandwidth is 2MHz, \( \sigma = 20\mu s \), \( \tau_{TTS} = 103\mu s \), \( \tau_{CTS} = \tau_{timeout} = 106\mu s \), \( \tau_d = 8ms \), \( p = 0.1 \). \( \tau_i = 1 \) \( \forall i \), \( \rho_i = \rho \) \( \forall i \). We vary the second-hop average SNR \( \rho \) from 1 to 20.

Consider that each source has an infinitely backlogged queue. Fig. 1 shows the system throughput of the three schemes. Our scheme and the scheme in [6] have much better performance than the naive scheme. When \( \rho \) is below 5, the scheme in [6] achieves higher system throughput (for example, when \( \rho = 2 \), throughput of the scheme in [6] is 28\% higher than that of our scheme). When \( \rho > 5 \), our scheme achieves better throughput performance (for example, when \( \rho = 20 \), throughput of our scheme is 14\% higher than that of the scheme in [6]). Fig. 1 also shows the average number of second-hop probings per transmission in our scheme. The number is not large, and decreases when \( \rho \) increases.

Now we evaluate delay performance which is important for delay-sensitive applications. Since delay-sensitive applications often have periodic traffic arrivals, we consider that each source has periodic packet arrivals with packet inter-arrival duration being 10 ms. Each packet has 500 bits. Define packet delay as the duration from a packet arrival until the moment when the packet is transmitted. Fig. 2 shows average packet delay in the three schemes. When \( \rho = 1 \), the traffic load is more than the system capacity, and thus, the packet delay is large in the three schemes. When \( \rho \) increases, the system capacity increases, and the packet delay decreases. For \( \rho \geq 2 \), the packet delay of our scheme and the scheme in [6] are similar and are less than 10\% of that in the naive scheme. This is because, by giving up transmission opportunity (in our scheme and the scheme in [6]) and/or letting relays wait (in our scheme), each transmission can have a higher rate, and thus, queuing delay in the system decreases.

Overall, an adaptive scheme can work as follows. By normalizing the first-hop average SNR to be 1, the curves of system throughput vs. second-hop average SNR \( \rho \) can be numerically plotted off-line for our scheme and the scheme in [6], similar to plotting the two upper curves in Fig. 1. The intersection of the two curves gives a threshold \( \rho^* \) (which is approximately 5 in the example of Fig. 1). Then our scheme should be used when \( \rho > \rho^* \), and the scheme in [6] should be used otherwise. This adaptive scheme can achieve good performance in low to high SNR.

REFERENCES