ON CHANNEL SELECTION FOR ENERGY-CONSTRAINED RATELESS-CODED D2D COMMUNICATIONS

Setareh Maghsudi and Slawomir Stańczak

Technische Universität Berlin, Einsteinufer 25, 10587 Berlin, Germany

ABSTRACT

We consider a distributed channel selection problem for device-to-device (D2D) communications underlaying conventional cellular networks. In our model, underlaying devices exploit the possibly-idle licensed cellular spectrum in order to establish direct communications links, and transmit using rateless coding under energy constraint. While the quality of each channel is assumed to be stochastic, the availability is non-stochastic (adversarial). Moreover, cellular channels are idle only for some finite time. As acquiring prior information about channel quality and/or availability yields excessive cost, we assume that D2D devices do not possess any prior information about channels. Device pairs face the problem of selecting a suitable channel so that a successful data delivery under the energy constraint is guaranteed. We model this problem as a multi-armed bandit game with mortal arms, and provide an algorithmic solution.

1. INTRODUCTION

The basic idea of D2D communications is to reuse cellular spectrum resources by allowing nearby wireless devices with local needs to establish direct transmission links. This approach improves the coverage and capacity performances; nonetheless, it imposes new challenges with respect to radio resource management. In a large body of literature such as [1], the resource management is performed by some central controller, which is in possession of both D2D and cellular channel knowledge. Besides requiring heavy amount of information, centralized approaches yield excessive time and computational complexity. Moreover, the priority of cellular users is not always taken into account. In some research studies, the D2D system is regarded as a distributed secondary network, whose members are allowed to utilize idle cellular channels. In a great majority of such works, however, some information is available to device pairs or a control channel is available [2]. Furthermore, in all previous studies, fixed rate coding is used, which is suboptimal in the sense of outage probability [3]. In addition, channels are available for infinite time, and availability is conventionally a random variable following Bernoulli distribution [4], which is unrealistic.

This research was supported by Grant STA 864/3-3 from German Research Foundation (DFG).

Against this background, we generalize the state-of-the-art along the following directions.

- D2D devices have no prior or side information.
- Rateless coding is used, which not only does not require channel state information, but also guarantees arbitrary small outage probability [3].
- Channels are available only for finite time.
- Channel availability is adversarial.

We consider a D2D energy-constrained application, such as short-range transmission in wireless sensor networks. Based on realistic conditions that primary channels are available only for finite time, and the energy is costly, the problem encountered by D2D devices is to select some suitable channel so that the transmission is successful using some *fixed, predetermined* energy supply. We cast the problem as a mortal multi-armed bandit game [5], where arms are only available for finite time, and the availability is adversarial, i.e., it cannot be attributed to any probability function. We then provide a distributed algorithmic solution to this problem.

2. SYSTEM MODEL

We consider a D2D communications system integrated in a cellular network supplied with M orthogonal frequency channels. By D2D user we refer to any *device pair*, i.e., transmitter-receiver pair, that communicate directly, and is allowed to use primary channels only if they are not occupied by a cellular user. We assume perfect channel sensing. Each channel is idle only for some *finite* number of transmission trials, and we do not make any assumption on the statistics of channel availability. Each D2D user selects its transmission channel on its own, given *no* prior information. The selection is not necessarily performed simultaneously by all D2D users, and each idle channel serves the first D2D user by which it is selected. Ties are broken using any arbitrary rule. As a result, transmissions are only disrupted by the zero-mean additive white Gaussian noise (AWGN) with variance N_0. After selecting a channel, at each time t, any D2D transmitter uses rateless coding in order to transmit its data X_t,[1] which

[1]The data and also the channel distribution are different for different D2D users.

978-0-9928626-3-3/15/$31.00 ©2015 IEEE

is received at its corresponding receiver as Y_t. Formally,

$$Y_t = h_{t,u} X_t + N_t, \tag{1}$$

where $h_{t,u}$ ($0 < |h_{t,u}| < \infty$) is the complex channel coefficient between the transmitter and receiver of the device pair at channel u at time t, and N_t is the AWGN. Let $h_{t,u}$ follow a specific probability function. We gather the channel coefficients from $t = 1$ to $t = T$ in the set $h_u^{(T)} = (h_{u,1}, ..., h_{u,T})$ and refer to $h_u^{(\infty)}$ as a channel realization [3].

3. PROBLEM FORMULATION

3.1. Rateless Coding

Consider an energy-constrained application, and assume that each device is equipped with an infinitely large buffer. L frames are to be transmitted, each consisting of q units of data. We assume that each channel use (transmission trial) costs one unit of energy; That is, the number of possible channel uses is equal to the available energy units. For transmitting L frames, maximum LD channel uses (i.e., LD units of energy) are allowed, where D is some fixed integer. Moreover, each frame $l \in \{1, ..., L\}$ has to be correctly decoded after at most lD channel uses, otherwise an outage occurs. When utilizing rateless codes, receiver accumulates mutual information in order to decode the message [3]. Therefore, the required number of trials for decoding a frame changes with channel conditions. If some frame l is decoded before lD channel uses, extra trials are saved to be used for transmitting the upcoming frames. Theoretically, for a sufficiently large number of frames, a large number of channel uses becomes available, which means that the outage probability approaches zero [3]. Let $n_l(\xi)$ be the required number of trials to decode the l^{th} frame, which is given by

$$n_l(\xi) = \frac{q}{R(\xi)}, \tag{2}$$

with $R(\xi)$ being the realized transmission rate over the channel realization ξ. As L approaches infinity, the sequence $\{n_l(\xi)\}_{l=1}^{L}$ converges to some random variable, $n^*(\xi)$, in distribution [3]. Representing the outage probability by P_L^{out}, the following holds.

Lemma 1 ([3]). *For a rateless coding scheme, if* $\mathrm{E}[n^*(\xi)] > D$*, then for all* $L > 0$, $\mathrm{P}_L^{out} > \alpha$ *for some* $\alpha > 0$. *If* $\mathrm{E}[n^*(\xi)] < D$*, then for any* $\epsilon > 0$, $\mathrm{P}_L^{out} < \epsilon$ *for sufficiently large* L.

Note that Lemma 1 holds for the general case where the coefficients of any given channel are independent but non-identically distributed (i.ni.d.). According to our system model, however, we consider independent and identically distributed channel gains (i.i.d.). Due to the one-by-one mapping given by (2), it can be concluded that the number of channel uses required for transmitting a frame through each channel are also i.i.d.; In other words, the sequence $\{n_l(\xi)\}_{l=1}^{L}$ follows a specific probability distribution. Besides, it should be noted that that the sequence of i.i.d. random variables following some parametric probability distribution function converges in distribution to that function.[2]

3.2. Channel Selection Problem

As mentioned before, each D2D user aims at selecting a channel from the set of primary channels. The selection is performed only once and cannot be changed. According to Section 3.1, the energy supply of each transmitter supports LD channel uses. Therefore, in order to be able to use the entire energy allowance, the selected channel must be available for at least LD trials *following the selection*. On the other hand, as rateless coding is applied, by Lemma 1, the selected channel must demonstrate some minimum quality in order to ensure successful transmission under the energy constraint. In brief, the following two prerequisites are necessary and sufficient for the selected channel to guarantee a successful data delivery.

- $C1:\ \ T_a \geq LD$, where T_a is the number of channel uses (transmission trials) in which the selected channel is idle *following* the selection.
- $C2:\ \ \mathrm{E}[n^*(\xi)] < D$, with ξ being the realization of the selected channel.

Therefore the problem is to select some channel from the set of primary channels so that $C1$ and $C2$ are both satisfied, where we assume that at least one such channel exists.

Definition 1. *A channel is called "suitable" if it satisfies both conditions* $C1$ *and* $C2$.

Definition 2. *A selection is "successful" if the selected channel is suitable. Moreover, for any channel selection algorithm, "success probability" is referred to the probability that a suitable channel is selected by that algorithm.*

4. MORTAL MULTI-ARMED BANDIT MODEL

Multi-armed bandit game is a class of sequential decision making problems with incomplete information. In all bandit settings, there exists a player that is provided with some finite set of arms (actions), each producing some finite reward upon being pulled. The player observes the rewards only partially, for instance only after pulling an arm or for some limited time. Given incomplete observations, the player selects an arm (or a sequence of arms) that satisfies a long-run optimality condition. In mortal bandits, each arm is available only for some finite time. Thus, the aggregate reward that can be achieved by pulling any arm is finite.

[2] Later we see that knowledge of this distribution is *not* required by our channel selection scheme.

In this paper, we model the formulated channel selection problem as a multi-armed bandit game with mortal arms. Any D2D user is a player, and each primary channel is an arm, which is mortal, in the sense that it is occupied by cellular users most of the time and is available to D2D users only for finite number of trials during the game. The availability of arms is adversarial, while the average reward is stochastic. This is justified by the fact that the average reward (e.g. data rate) depends on the environment, which varies much slower compared to the decisions of cellular users (to use the channel or not) that affects channel availabilities. The optimality conditions required to be satisfied by the selected arm are those defined by conditions $C1$ and $C2$. Each player (D2D user) selects its channel only once per game.

We assume that initially, for some short exploration interval T_e, D2D users transmit freely through different channels (upon being idle), in order to gain some perspective of channel qualities. To this end, the transmitter is notified about the transmission rate by the receiver with a feedback. Note that the feedback is short and the process lasts for only T_e trials, hence the imposed overhead and also the wasted energy are negligible. Afterwards the decision interval T_d begins, in which the player is not allowed to explore. Nevertheless, due to the assumption of perfect sensing (see Section 2), it still observes which channels are available. The selection is performed during this interval. We assume that there exists at least one channel that is available for at least T_d transmission trials (channel uses) *following* the exploration interval.

In [6], an algorithm is proposed for solving a problem that is partially similar to the problem considered here. In the multi-armed bandit setting considered in [6], it is assumed that all mortal arms yield one unit of reward upon availability. The problem is then defined as to select an arm that is available for some minimum number of trials following the selection. In other words, it is desired that the selected arm pays at least some minimum amount of reward. By comparison, it can be concluded that this problem is identical to satisfying condition $C1$. However, with respect to our problem formulation, satisfying $C1$ does not suffice and the selected arm must satisfy condition $C2$, as well. Therefore we generalize the algorithm of [6] by adding some steps in order to qualify it for solving the channel selection problem at hand. The result is a procedure with two cascaded steps, described in the following.

During the exploration interval (T_e), at each step, the D2D user (player) selects one of the *available* channels to explore uniformly at random. The D2D user transmits and observes the realized transmission rate. Then, it estimates the required number of channel uses for transmitting a frame through that channel. At the end of this interval, a sequence of experimental data of length $v = \frac{T_e}{M}$ (on average) is available to the player for each arm. Using these data sets, the player performs the following one-sided hypothesis test.

$$H_0 : \mu \leq 1 \ \text{ vs. } \ H_1 : \mu > 1, \tag{3}$$

where $\mu = \frac{\mathrm{E}[n^*(\xi)]}{D}$. In order to perform the test given by (3), the test statistic

$$K(\mathbf{x}) = \Pr\left[\sum_{i=1}^{v} x_i \left(U_{(i+1)} - U_{(i)}\right) \leq 1\right] \tag{4}$$

can be used [7], where $\mathbf{x} = (x_1, ..., x_v)$ and $\mathbf{x} \in [0, \infty]^v$ is the data vector of length v. Moreover, $U_{(1)} \leq ... \leq U_{(v)}$ are the order statistics of v i.i.d. standard uniform random variables $U_1, ..., U_v$ and $U_{v+1} = 1$. Given this statistic, $\{K \leq \alpha\}$ is used as the rejection region for the null hypothesis H_0 of the test in (3). Note that α is a given nominal value that balances the trade-off between the false alarm and misdetection, and indicates that the result of the test is correct with probability $1 - \alpha$ [7]. Through hypothesis testing, the player selects a subset M' from the set of primarily actions that satisfy the null hypothesis H_0.[3] In fact, this subset includes the arms for which condition $C2$ holds with probability $1 - \alpha$. In order to avoid trivial results in the analysis, we make the following assumption.

Assumption A1. *α and T_e are selected so that M' is not empty. That is, at least one arm passes the hypothesis test.*

Using the subset M', the player continues to the next step of the algorithm, as briefly described below [6]. During observation interval, the player observes which arms are available at each step. Denote the set (and the number) of such arms by M''. For any arm $j \in M''$ and starting from the smallest index (j), the player generates a random variable Z_j drawn from the Bernoulli distribution with parameter $M'^{\frac{3a}{T_d}-2}$, where a denotes the number of steps in which arm j has been available so far in the observation interval. If the random experiment results in success ($Z_j = 1$) for any arm j, that arm is selected. Otherwise the process continues. The procedure is summarized in Algorithm 1. The second part of Algorithm 1 (without hypothesis testing) is analyzed in [6]. The result follows.

Theorem 1. *([6]) In Algorithm 1, let $T_o \geq 3LD \log M$ and assume that $M' = M$. Then the strategy picks up an action that is available for at least LD trials following the selection, with probability at least $1 - O(\frac{LD \log(M)}{T_o} + \frac{1}{M})$.*

The expected success probability for Algorithm 1 is stated in the following.[4]

Proposition 1. *Let M_0 be the subset of M that includes all channels with $\mu \leq 1$, and $T_o \geq 3LD \log M'$. Under Assumption A1, the expected success probability of mortal bandit strategy (Algorithm 1) is at least*

$$P_s = \left(1 - \alpha^{M_0}\right) \left(1 - O\left(\frac{LD \log(M')}{T_o} + \frac{1}{M'}\right)\right). \tag{5}$$

[3] For simplicity, a set and its cardinality are referred to by the same letter.
[4] Proof is omitted for the lack of space.

Algorithm 1 Mortal Bandit Channel Selection

Initialize the sample vector $\mathbf{x}_m$ for every channels m.
for $t = -T_e, ..., 0$ **do**
  Observe idle channels, and select one of them, say m, uniformly at random.
  Transmit and observe the realized transmission rate by using the feedback signal from the receiver.
  Estimate the required number of channel uses for transmitting one frame through m, $x_{m,t}$.
  Append $x_{m,t}$ in the sample vector of channel m, i.e. $\mathbf{x}_m$.
end for
Perform the hypothesis test (3) for all $m \in \{1, ..., M\}$ to select the subset $M' \subseteq M$ for which the null hypothesis H_0 holds.
Initialize the availability vector $\mathbf{A} = (a_1, ..., a_{M'})$ with $a_m = 0$.
for $t = 1, ..., T_o$ **do**
  if $m \in \{1, ..., M'\}$ is available, **then**
    Let $a_m = a_m + 1$.
    Draw a Bernoulli random variable Z_m with $\Pr[Z_m = 1] = M'^{\frac{3a_m}{T_o} - 2}$.
  end if
  if $Z_m = 0 \forall m \in \{1, ..., M'\}$, **then**
    continue.
  else
    select the smallest index $j \in M'$ such that $Z_j = 1$ and exit.
  end if
end for

5. NUMERICAL ANALYSIS

We consider a D2D user with fixed average power, willing to select a transmission channel.[5] Signal to noise ratio (SNR) varies by changing the expected value of the channel gains, thereby channel realizations.

First, we investigate the impact of rateless coding on transmission performance. We assume that there exists only one primary channel ($M = 1$) that is *permanently* available, and we simulate the outage probability of the transmission through this channel, as a function of SNR. There exist $L = 10^3$ frames, each consisting of four bits ($q = 4$). Assuming binary phase shift keying (BPSK) modulation, two bits are transmitted per channel use. According to problem formulation, each frame l must be decoded correctly after at most lD channel uses, otherwise an outage occurs. The result is depicted in Figure 1, which shows that there exists a transition phase in the outage performance. The transition is associated with some threshold in the SNR value. While before the threshold the outage is large, it drops abruptly to zero as soon as the SNR value exceeds the threshold, which is smaller for larger D.

[5]According to our system model and problem formulation, each idle primary channel can be used by only one D2D user. Therefore, the number of D2D users does not contribute to algorithm's performance. Hence we avoid unnecessary complexity, and simulate a single-user scenario.

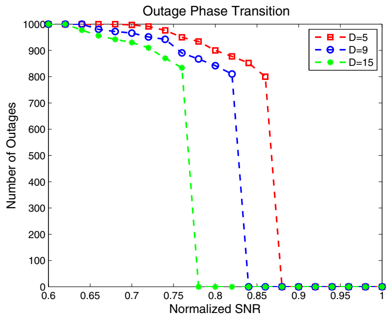

Fig. 1. Outage performance as a function of SNR.

In the next step, we assume that there exist $M = 10$ primary channels, which are characterized by the expected channel gain and the likelihood of availability. The characteristics are summarized in Table 1 for all channels.[6] We evaluate the performance of the following selection methods.

Table 1. Channel Characteristics (AL: Availability Likelihood, EG: Expected Gain)

Channel	1	2	3	4	5
AL	$(0.7, 1]$	$(0.1, 0.6]$	$(0.3, 0.7]$	$(0, 0.4]$	$(0, 0.9]$
EG	$(0, 1]$	$(0, 0.9]$	$(0, 1]$	$(0, 0.8]$	$(0, 0.4]$
Channel	6	7	8	9	10
AL	$(0, 0.3]$	$(0, 0.7]$	$(0.5, 1]$	$(0, 0.4]$	$(0.7, 1]$
EG	$(0, 0.3]$	$(0, 0.5]$	$(0, 0.4]$	$(0, 0.5]$	$(0, 0.4]$

- Random: A channel is selected uniformly at random.
- Informed: The player is provided with full statistical knowledge (Table 1), and the best channel is selected.
- Mortal multi-armed bandit (Algorithm 1).

Number of frames remains as before and we select $D = 5$. Moreover, $T_e = 10$, and $T_d = 3LD \log(M)$. Transmission is performed in 20 independent experiments. The average number of outages (over 20 experiments) is depicted in Figure 2. From this figure, the bandit approach exhibits acceptable performance in comparison to the informed case.

The channels selected by the three approaches are shown in Figure 3, for 10 randomly-selected experiments. Due to its deterministic nature, the informed approach always selects the same channel. Mortal bandit algorithm chooses a channel from the subset of channels with higher possible average gains, due to hypothesis testing. The final output is however randomized, since in the second part of the algorithm randomization is performed.

[6]Availability of each channel is modeled as a Bernoulli process whose parameter is selected adversarially from the given interval, at each *time step*; i.e., the parameter changes during each experiment as well as over experiments. This ensures that some channels are more likely to be available than others. The expected gain of each channel is drawn from the given intervals following uniform distribution for each *experiment*. Hence, it is fixed during each experiment, and is varying *over experiments*.

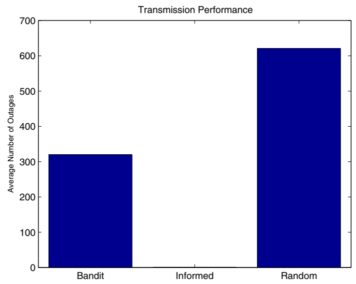

Fig. 2. Average outage performance of three strategies.

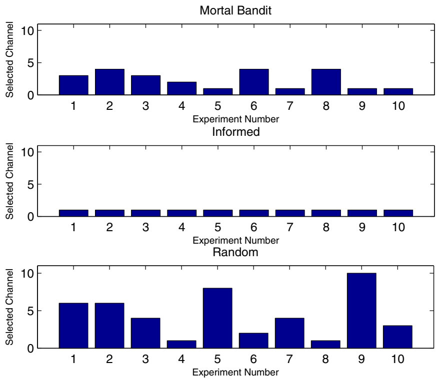

Fig. 3. Channels selected by three strategies.

From Algorithm 1, the performance of mortal bandit approach depends on the length of exploration period, T_e. On the other hand, as rateless coding is used, D plays an important role, as well, by changing the energy constraint. In the following, we investigate the effect of these parameters. To this end, we simulate the outage performance of Algorithm 1 as a function of T_e. For each value of T_e, 20 independent experiments are performed, and the average number of outages is calculated. Results are shown in Figure 4 for different values of D. As expected, the performance improves by increasing T_e. Nevertheless, this improvement has some limit, as it only affects the hypothesis testing part. In the best case, all arms (and only those) that satisfy condition $C2$ pass the test. Beyond that, increasing T_e does not improve the performance. In such case, only increasing D (i.e. relaxing the energy constraint) improves the outage performance. This is due to the fact that by increasing D, a larger number of channels become suitable. For instance, as suggested by the figure, for $D = 15$, all channels that pass the hypothesis test are suitable, thus the error floor of the bandit scheme vanishes.

As the final experiment, we discuss the role of observation period, T_o. Assume that $M' = M$. By Theorem 1, in order to achieve an acceptable error performance, one must have $T_o \geq 3lD\log(M)$. On the other hand, it is assumed that at least one channel is idle for at least T_o channel uses following the selection. Figure 5 shows the success probability as a function of T_o. It is not surprising to see that longer channel availability (T_o) yields better performance. Moreover, as suggested by the figure, success probability remains almost constant with increasing number of channels. It should be however noted that by increasing the number of channels, minimum availability (T_o) must increase logarithmically. As a result, a fixed success probability becomes more difficult to be achieved for larger sets of channels.

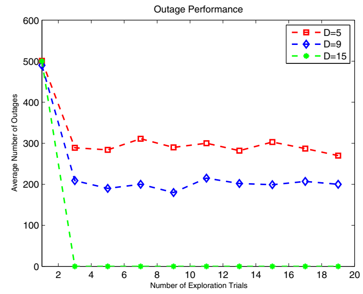

Fig. 4. Average outage performance of Algorithm 1 as a function of exploration interval duration, T_e.

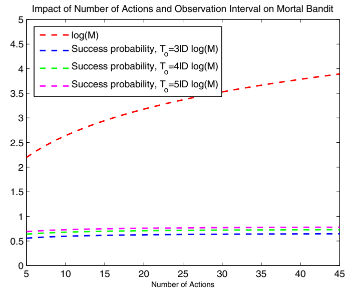

Fig. 5. Impact of the number of actions, M, and the length of observation interval, T_o, on the success probability.

REFERENCES

[1] T. Han et al., "Uplink channel reusing selection optimization for device-to-device communication underlaying cellular networks," in *IEEE PIMRC*, 2012, pp. 559–564.

[2] S. Maghsudi et al., "Dynamic bandit with covariates: Strategic solutions with application to wireless resource allocation," in *IEEE ICC*, June 2013, pp. 5898–5902.

[3] J. Castura et al., "On rateless coding over fading channels with delay constraints," in *IEEE ISIT*, July 2006, pp. 1124–1128.

[4] Y. Xu et al., "Opportunistic spectrum access in unknown dynamic environment: A game-theoretic stochastic learning solution," *IEEE Trans. on Wireless Commun.*, vol. 11, no. 4, pp. 1380–1391, 2012.

[5] D. Chakrabarti et al., "Mortal multi-armed bandits," in *Neural Information Processing Systems Conference*, 2008, pp. 273–280.

[6] B. Awerbuch et al., "Making commitments in the face of uncertainty: How to pick a winner almost every time (extended abstract)," in *ACM Symposium on the Theory of Computing*, 1996, pp. 519–530.

[7] N. Gaffke, "Three test statistics for a nonparametric one-sided hypothesis on the mean of a nonnegative variable," *Mathematical Methods of Statistics*, vol. 14, no. 4, pp. 451–467, 2005.