TRANSMISSION MODE SELECTION FOR NETWORK-ASSISTED DEVICE TO DEVICE COMMUNICATION: A LEVY-BANDIT APPROACH

Setareh Maghsudi and Sławomir Stańczak

Heinrich-Hertz-Lehrstuhl für Informationstheorie und theoretische Informationsthechnik,
Technische Universität Berlin, Einsteinufer 25, 10587 Berlin, Germany

ABSTRACT

This paper studies device-to-device (D2D) communication underlaying cellular infrastructure, where each device pair is provided with two transmission modes: indirect and direct. Indirect transmission is a two-hop interference-free transmission via a base station. Despite being interference-free, this transmission type might be inefficient in communications scenarios where short-distance connections can be established. Moreover, the need for centralized resource allocation and utilizing extra hardware may lead to excessive complexity and unacceptable costs. In such scenarios, direct transmissions can utilize the proximity- and hop gains to achieve higher rates and lower end-to-end latencies. While having a potential for huge performance gains, direct D2D communications poses some fundamental challenges resulting from the absence of a devoted controller such as uncoordinated interference and unavailability of permanent direct channels. Roughly speaking, in an average sense, while indirect transmission pays safe and steady reward, direct transmission is risky, yielding a stochastic reward which might be lower than the guaranteed reward of indirect transmission, despite the proximity- and hop gains. Transmitters should therefore choose the most efficient transmission mode in the presence of limited information. This paper characterizes the reward process for each transmission mode to model the mode selection problem as a two-armed Levy-bandit game. Accordingly, the reward of the risky arm (direct mode) is considered to be a pure-jump Levy process, following compound Poisson distribution. Mathematical results from bandit and learning theories are used to solve the selection problem. Numerical results complete the paper.

***Index Terms*—** Device-to-device communication, mode selection, compound Poisson process, Levy-bandit game.

1. INTRODUCTION

1.1. Related Works

Despite promising higher flexibility and more reliable services, hybrid D2D-cellular network structures call for the development of novel resource allocation schemes that ensure efficient and robust data transmission under limited radio resources and channel or traffic knowledge. *Transmission mode selection* can be defined as the problem of selecting a single transmission mode given different choices that are enabled by a specific network structure. In [1], for example, a heterogeneous network of device to device (D2D) links and IMT-Advanced[1] in a single cell is considered, where each user is provided with three possible transmission modes, namely cellular mode, overlay relay mode and underlay relay mode. The authors provide a mode selection algorithm which minimizes the power consumption. Reference [2] studies the problem of mode selection for D2D communications in LTE-Advanced[2] networks. A solution based on coalitional game among D2D links is proposed, where coalitions select their communication modes that can be either cellular or direct. In [2], similar to [1], the ultimate goal is to minimize the power consumption. Moreover, in [3], power control through mode selection is studied in a D2D communication network integrated into a traditional cellular system. Further examples include [4], [5] and [6], which study similar networks and performance metrics.

1.2. Our Contribution

Most of the works mentioned above are based on numerical studies and suffer from the lack of analytical models for transmission mode selection [3]. Moreover, transmission mode is either selected by a base station (BS) in a centralized manner [1], or large amount of information (e.g. periodic measurements of signal to noise ratio (SNR)) is required for decision making [4]. In addition, some works completely ignore the interference [5] and almost all of prior works focus on power consumption optimization.

In this paper, we consider a conventional OFDMA-based[3] cellular wireless network in which each device pair is allowed to select one of the following transmission modes:

- indirect transmission via a BS (or a relay) using *permanently* available and *orthogonal* channels,
- direct transmission using a *possibly* available and *interference-corrupted* channel, over which the transmission performance *might be* better than that of the indirect channel.

In contrast to previous works, we consider a dynamic network in which the number of interferers is itself a random variable. Therefore, while indirect mode yields a fixed reward, the reward of direct mode is stochastic, with an expected value higher or lower than the reward of indirect transmission. We model this reward by a pure-jump Levy process that is the compound Poisson process. Moreover, we develop a general framework in which each transmitter aims at maximizing some reward (or minimizing some cost) by selecting the optimal mode as early as possible. This formulation accommodates various performance metrics, including throughput, delay, power consumption, etc. We adapt an analytical model for mode selection, and formulate the problem as a two-armed bandit game with one safe arm and one risky arm, where the reward of the risky arm is a pure-jump Levy process [7], [8]. We suggest some optimal and sub-optimal solutions to this game.

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

[1] International Mobile Telecommunications-Advanced

[2] Long-Term Evolution-Advanced

[3] Orthogonal Frequency Division Multiple Access

2. SYSTEM MODEL AND PROBLEM FORMULATION

We consider an OFDMA-based cellular network, enhanced with the possibility of direct D2D communications. There exist K transmitter-receiver pairs, referred to as users, and denoted either by (k, k') or just by $k \in \{1, ..., K\}$. At time t, the channel coefficient of link $u \to v$ (including Rayleigh fading and path loss) is denoted by $h_{t,uv}$. All nodes transmit at unit power. Orthogonal transmissions are corrupted by zero-mean additive white Gaussian noise (AWGN) with unit variance, which is neglected in case of interference-corrupted transmission for simplicity. At time t, transmitter k selects one of the two possible transmission modes: either cellular (indirect) or D2D (direct), and receives a payoff $u^{(k)}(t)$.

2.1. Cellular (Indirect) Transmission Mode

In order to use the indirect mode, transmitters are required to submit a channel request to a BS. Orthogonal channel allocation is governed by the BS, and data exchange among any transmitter-receiver pair is performed in a two-hop style via the BS that acts as a relay. Since no queuing is allowed, the achievable throughput of each trial is equal to the minimum throughput of the first and second hops. As a result, the throughput for D2D user k yields

$$R_t^{(k)} = \frac{1}{2} \min\{R_{t,1}^{(k)}, R_{t,2}^{(k)}\}, \tag{1}$$

with $R_{t,i}^{(k)}$, $i \in \{1, 2\}$, being the throughput of hop i. $R_{t,1}^{(k)}$ is given by

$$R_{t,1}^{(k)} = \log(1 + |h_{t,km}|^2), \tag{2}$$

where m denotes the BS. $R_{t,2}^{(k)}$ follows similarly. Therefore, $R_t^{(k)}$ is the first order statistics of two random variables. Moreover, the average throughput of indirect transmission link yields

$$\bar{R}^{(k)} = \mathrm{E}\left[R_t^{(k)}\right] = \frac{1}{2}\mathrm{E}\left[\min\left\{R_{t,1}^{(k)}, R_{t,2}^{(k)}\right\}\right]. \tag{3}$$

If selected at time t, the indirect mode yields a reward $u_s^{(k)}(t) = R_t^{(k)}$. Thus, $U_s^{(k)} = \bar{R}^{(k)}$ denotes the expected reward of the indirect mode to user k. Here we assume that upon selection, this expected reward is guaranteed by the indirect mode.

Proposition 1. *Let $\theta_i^{(k)}$, $i \in \{1, 2\}$, denote the average channel gain of hop i for user k. Then we have*

$$U_s^{(k)} = \exp(1/\theta_T^{(k)}) \cdot \mathrm{E}_1(1/\theta_T^{(k)}), \tag{4}$$

where $\theta_T^{(k)} = \frac{\prod_{i=1}^2 \theta_i^{(k)}}{\sum_{i=1}^2 \theta_i^{(k)}}$ and $\mathrm{E}_1(x) = -\mathrm{Ei}(-x) = \int_x^\infty \frac{\exp(-t)}{t} dt$ denotes the exponential integral. Moreover,

$$U_s^{(k)} \le \log\left(1 + \frac{\prod_{i=1}^2 \theta_i^{(k)}}{\sum_{i=1}^2 \theta_i^{(k)}}\right). \tag{5}$$

Therefore, the expected reward of cellular mode depends on average channel gains, which are determined by the nature. Hence, given θ_i, $i \in \{1, 2\}$, indirect transmission type is safe for the user, as it yields, on average, a fixed and highly-predictable reward (or cost). In this paper, we assume that $U_s^{(k)}$ is known by the D2D user k, and that $u^{(k)}(t) = U_s^{(k)}$, if the indirect mode is chosen. This information can be provided by the already-deployed cellular structure with low overhead.

2.2. D2D (Direct) Transmission Mode

As explained before, the two-hop indirect communication might be inefficient. In such situations, it may be preferable for the user to choose a direct channel, which might be available or not.[4] On the other hand, direct channels are not controlled by any central unit, hence interference might occur. We model these two characteristics as two random processes. The availability of direct channels is modelled as a Poisson process with intensity $\lambda \ge 0$, which means that the inter-availability times follow an exponential distribution with mean λ^{-1}. Number of interferers is considered to be a random variable following geometric distribution with parameter (success probability) $q^{(k)} \in [0, 1]$.[5] For each single user, we assume that all interfering signals are independent and identically distributed (i.i.d.);[6] however, interference signals are *not necessarily i.i.d.* for all users. The reward is proportional to the signal to interference ratio (SIR). Therefore, if the direct channel is available, the throughput follows

$$R_t'^{(k)} = \log(1 + \frac{|h_{t,kk'}|^2}{\sum_{l=1}^n |h_{t,lk'}|^2}),\ \sum_{l=1}^n |h_{t,lk'}|^2 > 0, \tag{6}$$

where n is the number of interferers. As a result, if the D2D mode is selected at time t, $u_r^{(k)}(t)$ yields

$$u_r^{(k)}(t) = \begin{cases} 0 & \text{if } x = 0 \\ R_t'^{(k)} & \text{if } x = 1 \end{cases}, \tag{7}$$

where $x = 0$ if the direct channel is *not* available and $x = 1$ otherwise. Thus, for direct mode, $u^{(k)}(t) = u_r^{(k)}(t)$, which follows a compound Poisson process, that is a pure-jump Levy process. Let $U_r^{(k)} = \lambda \bar{R}'^{(k)}$ ($\bar{R}'^{(k)} = \mathrm{E}\left[R_t'^{(k)}\right]$) denote the average reward achievable by user k if the indirect mode is selected.

Proposition 2. *Let $\theta'^{(k)}$ denote the average channel gain of direct link $k \to k'$. Moreover, let $\mu^{(k)}$ denote the average gain of interference links for user k. Then we have*

$$U_r^{(k)} = \frac{\lambda q'^{(k)} \theta'^{(k)}}{q'^{(k)} \theta'^{(k)} - \mu^{(k)}} \log\left(\frac{q'^{(k)} \theta'^{(k)}}{\mu^{(k)}}\right), \tag{8}$$

where $q'^{(k)} = 1 - q^{(k)}$, and $q'^{(k)} \theta'^{(k)} - \mu^{(k)} > 0$ by definition. Moreover,

$$U_r^{(k)} \le \lambda \log\left(1 + \frac{\theta'^{(k)} q'^{(k)}}{\mu^{(k)}}\right). \tag{9}$$

By (8), due to interference, the expected throughput achieved by the D2D mode depends not only one the nature (average channel gains), but also on the actions of other users (inside and outside the cell). While average channel gains are relatively easy to acquire, interference characteristics cannot be estimated easily. As a result, the reward of the D2D mode is considered to be stochastic and risky, unlike the cellular mode which offers guaranteed reward. In order

[4] In the context of cognitive radio networks, this can be seen as primary channels that can be used by secondary users as D2D links, only in the absence of primary users.

[5] Basically, any discrete distribution can be considered here. We use Geometric distribution in order to obtain simple closed-form results.

[6] This assumption is only made to simplify the calculation of probability density function (pdf) of throughput. However, this does *not* restrict the application of the game model and solutions, as the knowledge of pdf is not required in general. In essence, it suffices that the throughput of D2D link is a compound Poisson process, and knowing the distribution of jump sizes (i.e. throughput) is not necessary.

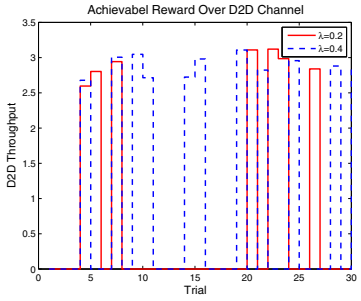

Fig. 1. Sample paths for throughput at risky link ($q = 0.8$).

to simplify the problem, later we assume that the network is aware of *possible* statistical characteristics (types) of the D2D link, and the user exploits this information to make the optimal decision.

Figure 1 illustrates some sample paths of the achievable throughput by using the D2D mode in a simulated network. As it is clear from the figure, the path follows a pure-jump Levy process. Moreover, the availability of the link is higher at larger Poisson rates.

Given the two transmission options, user k aims at maximizing its accumulated discounted reward, $R^{(k)}$ [9]. That is,

$$\text{maximize } R^{(k)} = \int_{t=0}^{\infty} re^{-rt} u^{(k)}(t)dt, \tag{10}$$

where parameter $r > 0$ denotes the discount factor (not to be confused with subscript r for risky channel). Adding a discount factor implies that the application is delay sensitive, and it is important to find the best mode as early as possible.[7]

3. BANDIT-THEORETICAL MODEL OF TRANSMISSION MODE SELECTION

In this paper, we model the problem formulated in Section 2 by a two-armed bandit game. Each user k is modelled as a player that is provided with one safe arm, s, which is considered to be the cellular transmission, and one risky arm, r, which is considered to be the indirect or D2D transmission. While the safe arm yields a known payoff $U_s^{(k)}$, the risky arm results in a stochastic payoff $u_r^{(k)}(t)$. According to the system model, this stochastic payoff is a compound Poisson process where jumps arrive with intensity λ and jump sizes are proportional to SIR values.[8] For this model, a play strategy is defined in the following.

Definition 1. *Assume that player k has one unit of resource at time t. Then, a play strategy, κ, assigns $0 \leq \alpha_{s,t}^{(k)} \leq 1$ part of this unit to the safe arm and $\alpha_{r,t}^{(k)} = 1 - \alpha_{s,t}^{(k)}$ part to the risky arm. Moreover, a strategy in which $\alpha_{s,t}^{(k)} \in \{0, 1\}$ is called a simple strategy.*

Primarily, each player k starts with a belief ϱ_0, which is the prior probability of the direct mode being superior to indirect mode. The player updates its belief after each time playing. Ultimately, it aims at maximizing its accumulated reward, $R^{(k)}$, given by (10). The player however does not know which arm yields larger payoff at each trial. Therefore, it is reasonable to pursue a less ambitious goal, which is to maximize its expected accumulated reward, with respect to both x (availability of D2D channel, see (7)) and $\alpha_{r,t}^{(k)}$, i.e. [9]

$$\text{maximize } \mathrm{E}\left[R^{(k)}\right] = \mathrm{E}\left[\int_{t=0}^{\infty} re^{-rt}\left[(1-\alpha_{r,t}^{(k)})U_s^{(k)} + \alpha_{r,t}^{(k)}U_r^{(k)}x\right]dt|\varrho_0\right]. \tag{11}$$

A strategy that satisfies (11) is said to be optimal (denoted by κ^*).

The problem in (11) can be stated as finding the optimal stopping time for a hypothesis-testing problem, in which user k has to make a decision as fast as possible between $U_r^{(k)} \leq U_s^{(k)}$ and $U_r^{(k)} > U_s^{(k)}$, given prior ϱ_0. To date, the optimal solution of problem (11) for the general case where $u_r^{(k)}(t)$ follows a compound Poisson process is unknown. Sub-optimal solutions *include* [12]

- fit the experimental data into a compound Poisson process in order to estimate its parameters (arrival rate and the distribution of jump sizes) and calculate its statistical characteristics,
- assume a prior distribution for the expected value of reward process and continue with experimentations in order to estimate the mean value.

To solve (11) optimally, we make the following assumption [8].

Assumption A1. *The risky arm belongs to one of the two types, namely high or low, denoted by h and l. Moreover, for each player k*

- $0 < U_s^{(k)} < U_{r,h}^{(k)} < \infty$.
- $\lambda_l = 0$ *and* $0 < \lambda_h < \infty$.

where $U_{r,h}^{(k)}$ is the average reward of the high type. λ_l and λ_h denote the Poisson availability intensity of low and high types, respectively.

In words, Assumption A1 states that while a risky arm of the low type results in zero payoff, the average reward of a risky arm with the high type is larger than that of the safe arm. The reward of the risky arm follows a compound Poisson process, and therefore the first payoff might appear only after some exponentially distributed time. The player does not know the type of the risky arm in advance; however, it is aware of values λ_h and $U_{r,h}^{(k)}$. Note that this information is statistical and if the number of risky arms is small or moderate, it can be provided by the cellular structure at low cost. Moreover, the statistical information about any channel can be obtained by observing the traffic data of that channel for some relatively short time intervals. Thus gathering this data does not cost much to the host network. All in all, even if no information is available, the sub-optimal approaches mentioned above can be still used. As a result, the game model is applicable. We consider two cases of this game model, namely multiple *independent* players and multiple *cooperative* players. We leave the strategic case for future work.

3.1. Multiple Independent Players

All players are independent, in the sense that their actions do not affect the beliefs and rewards of each other. Thus the problem reduces to K independent and identical single-player problems. We describe one of these problems and omit the user index (k) for convenience. Consider a player having access to one direct and one indirect channel. As stated before, the player does not know to which type the direct channel belongs. Let ϱ_t denote the belief of player at time t

[7] Formulation for cost (such as power consumption) minimization follows similarly.

[8] Here we assume that only one risky arm exists per user; Nevertheless, a similar problem with multiple independent risky arms per user can be decomposed into parallel problems with one risky arm and one safe arm (see [10], [7] and [11]). Therefore, the model is not restrictive and results can be easily generalized.

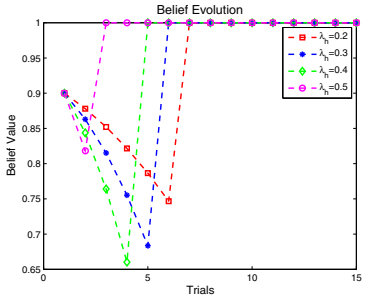

Fig. 2. Belief evolution for different values of λ_h.

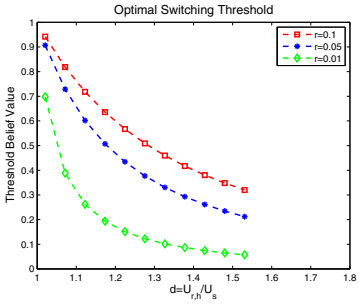

Fig. 3. Optimal switching belief as a function of $d = \frac{U_{r,h}}{U_s}$.

about the risky arm being high. Recall that $\alpha_{s,t}$ denotes the action of player during the time $t + \delta t$. Then $\varrho_{t+\delta t}$ can be calculated as follows. If the risky arm is high, it pays a reward with probability $1-\exp\left(-\alpha_{r,t}\lambda_h\delta t\right)$; otherwise, this probability is zero. Hence, the posterior belief that the risky arm is high yields [8]

$$\varrho_{t+\delta t} = \begin{cases} \frac{\varrho_t \exp\left(-\alpha_{r,t}\lambda_h\delta t\right)}{1-\varrho_t+\varrho_t \exp\left(-\alpha_{r,t}\lambda_h\delta t\right)} & \text{if } x=0 \\ 1 & \text{if } x=1 \end{cases}, \tag{12}$$

where $x = 1$ if a jump occurs in time period $\alpha_{r,t}\delta t$ and $x = 0$ otherwise. The optimal strategy in this case is a simple cut-off strategy.

Theorem 1. *([7], [8]) The optimal strategy for a Levy-bandit game with independent players that satisfies Assumption A*1 *is*

$$\kappa^*: \ \alpha_{r,t} = \begin{cases} 0 & \text{if } \varrho_t < \varrho^* \\ 1 & \text{if } \varrho_t > \varrho^* \end{cases}, \tag{13}$$

where $\omega = \frac{r}{\lambda_h}$ *and*

$$\varrho^* = \frac{\omega U_s}{(\omega+1)(U_{r,h}-U_s)+\omega U_s}. \tag{14}$$

In Figure 2, we illustrate the belief evolution of a user with prior $\varrho_0 = 0.9$ for different values of λ_h, in a communication network with independent users where the D2D channel is set to be high. Other variables include $q = 0.8$, $U_s = 0.57$ and $\frac{U_{r,h}}{\lambda_h} = 2.9$. Clearly, for larger λ_h, the first jump is expected to occur sooner. From the figure, the larger the λ_h, the faster is the reduction of the belief towards the optimal belief threshold in which the player should switch to the safe arm. As we have simulated the risky channel to be high, in all cases a jump occurs before this threshold, and therefore the belief jumps to one.

Figure 3 illustrates the optimal switching threshold as a function of $d = \frac{U_{r,h}}{U_s}$. It can be seen that for a fixed discount factor (r), larger values of d result in smaller values of ϱ^*. This means that the player would be more patient if higher reward is expected from the risky action. Also, for smaller discount factor, the player is more patient. As a result, information (experimenting) becomes more valuable and the user would change to the safe arm later.

3.2. Multiple Cooperative Players

In this case, there exist K players, each provided with one direct and one indirect channel. If some kind of dependency is assumed among risky arms, users would be better off cooperating. Here, we assume the following type of dependency.

Assumption A2. *The D2D channels of all users have the same type.*

As a result of this dependency, users are better off to cooperate with each other. That is, if a user observes a jump, it would announce it to other players. Therefore, users share a unique belief at each time t, namely ϱ_t. Let $A_{r,t} = \sum_{k=1}^{K} \alpha_{r,t}^{(k)}$, $U_S = \frac{1}{K}\sum_{k=1}^{K} U_s^{(k)}$ and $U_{R,h} = \frac{1}{K}\sum_{k=1}^{K} U_{r,h}^{(k)}$. Then, the posterior belief at time $t+\delta t$ is calculated as follows.

$$\varrho_{t+\delta t}^{(k)} = \begin{cases} \frac{\varrho_t^{(k)} \exp\left(-A_{r,t}\lambda_h\delta t\right)}{1-\varrho_t^{(k)}+\varrho_t^{(k)} \exp\left(-A_{r,t}\lambda_h\delta t\right)} & \text{if } x=0 \\ 1 & \text{if } x=1 \end{cases}. \tag{15}$$

In this case, like the case before, the optimal strategy is a simple cut-off strategy for all users [9].

Theorem 2. *([9]) The optimal strategy for a Levy-bandit game with cooperative players that satisfies Assumption A*1 *is*

$$\kappa^*: \ \alpha_{r,t}^{(k)} = \begin{cases} 0 & \text{if } \varrho_t^{(k)} < \varrho^* \\ 1 & \text{if } \varrho_t^{(k)} > \varrho^* \end{cases}, \tag{16}$$

where

$$\varrho^* = \frac{\omega U_S}{(\omega+K)(U_{R,h}-U_S)+\omega U_S}. \tag{17}$$

4. CONCLUSION AND REMARKS

We considered a cellular communications network enhanced with D2D transmission possibility, where each device pair is allowed to select between the noise-limited cellular (indirect) mode and the possibly-available, interference-corrupted D2D (direct) mode. While cellular mode yields guaranteed payoff, D2D mode results in a stochastic reward with an unknown expected value, since the number of interferers is assumed to be a random variable. The stochastic reward is modelled as a compound Poisson process, whose expected value can be higher or lower than that of the cellular mode. The problem is to select the transmission mode so as to maximize (minimize) some reward (cost). This scenario reduces to a two-armed bandit game with one safe arm and one risky arm whose reward series is a pure-jump Levy process. Under the assumption that users are provided with some statistical data by the network (hence the term network-assisted), the mode selection problem is solved by resorting to some recent mathematical results.

5. REFERENCES

[1] S. Xiang, T. Peng, Z. Liu, and W. Wang, "A distance-dependent mode selection algorithm in heterogeneous D2D and IMT-advanced network," in *IEEE Globecom Workshops*, 2012, pp. 416–420.

[2] K. Akkarajitsakul, P. Phunchongharn, E. Hossain, and V.K. Bhargava, "Mode selection for energy-efficient D2D communications in LTE-advanced networks: A coalitional game approach," in *IEEE International Conference on Communication Systems*, 2012, pp. 488–492.

[3] H. Xing and S. Hakola, "The investigation of power control schemes for a device-to-device communication integrated into OFDMA cellular system," in *IEEE International Symposium on Personal Indoor and Mobile Radio Communications*, 2010, pp. 1775–1780.

[4] H. Hu, M. Weckerle, and J. Luo, "Adaptive transmission mode selection scheme for distributed wireless communication systems," *IEEE Communications Letters*, vol. 10, no. 7, pp. 573–575, 2006.

[5] M.H. Han, B.G. Kim, and J.W. Lee, "Subchannel and transmission mode scheduling for D2D communication in OFDMA networks," in *IEEE Vehicular Technology Conference*, 2012, pp. 1–5.

[6] M. Jung, K. Hwang, and S. Choi, "Joint mode selection and power allocation scheme for power-efficient device-to-device (D2D) communication," in *IEEE Vehicular Technology Conference*, 2012, pp. 1–5.

[7] A. Cohen and E. Solan, "Bandit problems with Levy payoff processes," *Mathematics of Operations Research*, vol. 38, no. 1, 2013.

[8] G. Keller and S. Rady, "Strategic experimentation with poisson bandits," *Econometrica*, vol. 5, pp. 275–311, 2010.

[9] G. Keller, S. Rady, and M. Cripps, "Strategic experimentation with exponential bandits," *Econometrica*, vol. 73, pp. 39–68, 2005.

[10] J. C. Gittins and D. M. Jones, "A dynamic allocation index for the discounted multiarmed bandit problem," *Biometrika*, vol. 66, no. 3, pp. 561–565, 1979.

[11] K. Liu and Q. Zhao, "Indexability of restless bandit problems and optimality of Whittle index for dynamic multichannel access," *IEEE Transactions on Information Theory*, vol. 56, no. 11, pp. 5547–5567, 2010.

[12] J.O. Berger, *Statistical decision theory and Bayesian analysis*, Springer, 1985.