Online Rate Selection in Nonstationary Wireless Channels

Haoyang Zhang

An explosive increase of Internet of Things devices motivates the improvement of network throughput. We consider the problem of transmitting at optimal rates for a wireless transmitter-receiver pair in both long- and short-term varying channels with unknown statistics when the feedback is only success/fail transmission signals. Inspired by Multi-armed Bandits and competitive ecosystems, we propose an algorithm to identify the optimal rate by simulating the dynamics of ecosystems, which gives us a larger throughput compared to sticking to the averaged best rates.

Keywords: Rate Selection; Multi-armed Bandits; Lotka-Volterra competition; Bio-inspired computation.

1. Introduction

Future wireless networks such as the fifth-generation and Wi-Fi 6 are expected to support the exponential growth of data traffics in wireless networks. But an explosive increase of smartphones, tablets and the Internet of Things devices will continue increasing wireless data traffics. To accommodate the large scale of connected devices, it is necessary to improve the network throughput, which motivates us in this work to consider a new approach that by utilizing wireless channel conditions, allocating a proper transition rate to achieve a larger throughput in multi-rate wireless channels.

However, traditional communication protocols, which are based on probing and channel estimation techniques, are not necessarily applicable in future wireless networks. Because future networks use an extremely-high-frequency spectrum, which is highly sensitive to mobility and drastically time-varying, we cannot utilize the estimation for a long period. In this sense, the cost of keeping estimating channel state information is too expensive for practical use. Therefore, it is necessary to develop an online rate selection algorithm based on limited feedback, such as success/fail signals, in a wireless channel whose channel statistics are unknown and channel conditions are highly time-varying.

Recently, online wireless rate selection algorithms are widely discussed in the domain of statistical learning [1]. In this paper, we are inspired by amoeba dynamics, which are fluctuating dynamics under a conservative constraint, and developed a low-complexity algorithm for a single user where the only feedback is transmission success/fail signals.

The organization of this report is as follows. In Section 2, we first provide a multi-armed bandit (MAB) model to describe this rate selection problem, followed by an ecosystem metaphor to convert this MAB problem into competitive dynamics. In Section 3, we develop an algorithm based on the Lotka-Volterra (LV) mechanism with non-linear crowding effects, which is used to describe ecosystem dynamics. In Section 4, we conduct an experiment to test our algorithm performance by comparing it with some other policies. In Section 5, we discussed some limitations of our algorithm and provides some possible directions for future research.

2. Model and Problem Statement

2.1. Rate Selection Problems

We consider a wireless transmitter-receiver pair. At each time slot t, the channel can be in one of N states: $s_1, s_2, \dots, s_N$, with corresponding probabilities $\nu_1(t), \nu_2(t), \dots, \nu_N(t)$. Namely, the channel state is influenced by both long- and short-term variations caused by mobility. For each channel state s_i, the maximum possible transmission rate is r_i. The transmitter can choose N possible transmission rates: $r_1, r_2, \dots, r_N$. Let $R = \{r_1, r_2, \dots, r_N\}$. Without loss of generality, we assume that $r_1 > r_2 > \cdots > r_N \geq 0$. Therefore, for each rate r_i, the probability of being successfully transmitted at time t is $\theta_i(t) = \sum_{k=1}^{i} \nu_k(t)$. Without loss of generality, we assume that $\theta_N(t) = 1$ since we can always let $r_N = 0$

Our goal is to maximize the expected throughput. Namely, at each time t, the aim is to identify the optimal rate r^*:

$$r^*(t) = \underset{r_i \in R}{argmax}\, r_i \times \theta_i(t) \tag{1}$$

Since the success rate $\theta_i(t)$ is unknown, we cannot solve problem (1) directly. Instead, we let transmitter transmit at rate r_i to approximate $\theta_i(t)$ based on success/fail signals. The challenge is the exploration-exploitation dilemma. That is, sufficient probing transmission may allow us to identify the temporarily optimal rate $r^*(t)$, but it is at the cost of probing at non-optimal rates, let alone the long-term variation of $\theta_i(t)$. The non-stationary MAB problem models this dilemma perfectly, which we are introducing in the next part.

2.2. Non-stationary MAB Problems

MAB problems are well studied for sequential decision-making problems because of its inherent trade-off between exploration and exploitation [5]. The traditional MAB problem is formulated as follows: given N possible actions or “arms”, each associated with a fixed but unknown reward probability distribution, at each iteration an agent selects an arm to play and receives a reward, sampled from the respective arm’s probability distribution independently from the previous actions [4]. A non-stationary variant is that the reward probability distribution is time-varying. The goal of the agent is to maximize the cumulative rewards over time.

Let $y_i(t) \sim Bernoulli(\theta_i(t))$ be an indicator of successful transmission. We define the reward of each arm $x_i(t) = r_i \times y_i(t)$ at time slot t. Namely, if the rate r_i is less than or equal to the maximum possible rate r_j, which is given by channel state s_j $(j \leq i)$, the throughput is transmission rate r_i. Otherwise, the throughput is 0.

Let $i(t)$ denote the selected transmission rate index at time slot t. For ease of exposition, let $i^*(t)$ denote the index corresponding to the optimal rate, i.e., $r^*(t) = r_{i^*(t)}$. Then, the problem (1) can be rewritten as:

$$i^*(t) = \underset{i(t)}{argmax} \sum_t x_{i(t)}(t) = \underset{i(t)}{argmax} \sum_t r_{i(t)} \times y_i(t) \tag{2}$$

Our aim is to design an algorithm that efficiently identifies the optimal arm in a short period constrained by an exploration scheme so as to adapt the long-term variation.

2.3. LV Competitive Dynamics with Nonlinear Crowding Effects

When multiple species coexist and share the same basic requirements, they usually compete for resources. The Classical LV competition equation is well studied as a typical resource competition model, which is given by:

$$\frac{dz_i}{dt} = z_i\left(a_i - \sum_{j=1}^{N} c_{ij} z_j\right), i = 1,2,\ldots,N, \tag{3}$$

where z_i represents the population size of species i, $\mathrm{a_i}$ represents the growth rate of species i, and $\mathrm{c_{ij}}$ represents the (competitive) effect of species j on i [6, 7].

The analyses of this equation show that if the interspecific competition is stronger than the intraspecific competition, dynamics leads to exclusion of one species among N species [8]. Inspired by this competitive exclusion principle, a decision-making scheme to identify the optimal arm in stationary MAB problems with logarithm complexity is proposed recently [2]. The core idea of this scheme is a metaphor between MAB problems and LV competitive dynamics, which is shown in Table 1. However, this scheme cannot be directly applied to our problem because of the elimination of currently non-optimal arms, which can become the optimal arms in the future due to the long-term variation of channel statics.

Table 1. A metaphor between MAB problems and LV competitive dynamics.

Multi-Armed Bandits	Lotka-Volterra Dynamics
Probability of being selected	Population of a species
Constrain of total probability	Constrain of resource
Expectation of rewards	Natural growth rates
Realization of rewards	Temporary population growth
Identifying the optimal arm	Identifying the dominate (fittest) species
Maximize accumulative rewards	The survival of the fittest
Exploitation and exploration	Dynamics of competitive ecosystem

However, multiple species are generally possible to coexist using LV competitive dynamics with nonlinear crowding effects, which is given by:

$$\frac{dz_i}{dt} = z_i\left(a_i - \sum_{j=1}^{N} c_{ij} z_j\right) - d_i z_i^{1+\delta}, i = 1,2,\ldots,N, \tag{4}$$

where d_i represents the density-dependent factor of species i, and δ represents the nonlinearity of crowding effects [3]. We further extend the identification scheme with nonlinear crowding effects, which leads to the adaptability to nonstationary reward probability distributions.

3. Algorithms and Discussions

3.1. Algorithms Based on Ecosystem Dynamics

Let b be a small step size, $0 < \mathrm{b} \ll 1$. The dynamics (4) can be simulated by:

$$\Delta z_i = b\left(z_i\left(a_i - \sum_{j=1}^{N} c_{ij} z_j\right) - d_i z_i^{1+\delta}\right), i = 1,2,\ldots,N, \tag{5}$$

Therefore, simulation (5) can be rewritten for MAB problems based on the metaphor of Table 1 as:

$$\Delta P_i = b\left(P_i\left(\mu_i - \sum_{j=1}^{N} \mu_j\, P_j\right) - d_i P_i^{1+\delta}\right), i = 1,2,\ldots,N, \tag{6}$$

where P_i represents the probability of choosing $\mathrm{r_i}$, and $\mu_i = r_i \times \theta_i$ represents the expected throughput for $\mathrm{r_i}$. This gives our Algorithm 1.

Algorithm 1

Input hyperparameters b, d, δ, transmission outcomes $x(t)$
Initialize population of each rate $q_i = 1$
For time $t = 1,2,\ldots$
    $P_i(t) = q_i(t)/\sum_\mathrm{i}\left(\mathrm{q_i}(t)\right)$
    Draw an index j with probability $P_j(t)$
    Observe transmission outcome $x_j(t)$ for rate r_j
    $w = bx_j(t)/\left(1 - bx_j(t)\right)$
    $\Delta q_i = -bdq_i^{1+\delta}(t) + \mathbb{I}_{\{i=j\}} w \sum_i\left(q_i(t)\right)$
    $q_i(t+1) = q_i(t) + \Delta q_i$

Note that $\mathrm{P_i}$ is scaled from $\mathrm{q_i}$ with the same ratio for each possible rate $\mathrm{r_i}$, which gives that $w\sum_i\left(q_i(t)\right)$ is equivalent to $bP_i\left(\mu_i - \sum_{j=1}^{N} \mu_j\, P_j\right)$ on expectation [2].

3.2. Discussions About Hyperparameters

For ease of exposition, we choose a cyclically changing environment at a constant interval of $\mathrm{T} = 20000$ steps. We consider $\mathrm{r}_1 = 0.9, \mathrm{r}_2 = 0.5, \mathrm{r}_3 = 0.3, \mathrm{r}_4 = 0.1, \nu_1(0) = 0.2, \nu_2(0) = 0.3, \nu_3(0) = 0.4, \nu_4(0) = 0.1$ (see Fig. 1). To illustrate the function of each hyperparameter, we conduct the following experiment (see Fig. 2).

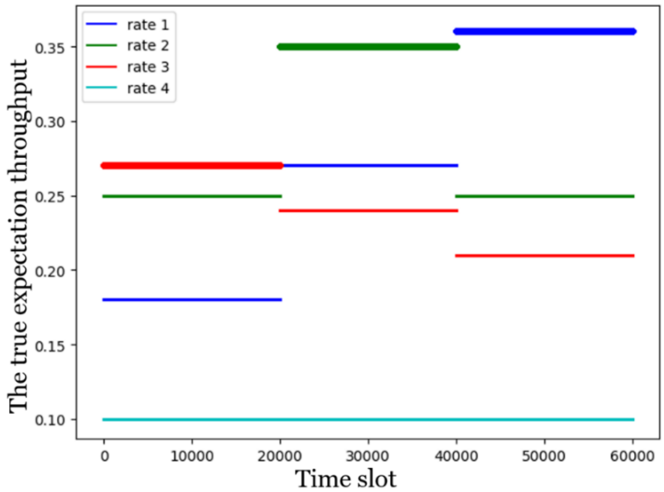

Fig. 1. A cyclically changed environment.

We observed that step size b controls the speed of reaching the equilibrium population. A large b leads to faster identification of the optimal rates, which is helpful for maximizing total throughput, while it also increases the risk of "extinction", i.e., some rates will never be considered again (see Fig. 2c).

The nonlinearity of crowding effects δ controls the strength of exploration. A large δ pushes every P_i to about $1/N$ to keep probing channel statics. By contrast, a small δ puts more importance on exploitation to maximize temporary throughput at risk of "extinction" (see Fig. 2d).

The equilibrium population is influenced by the density-dependent factor d, which causes overshooting before convergence. In Fig. 2f, the equilibrium population is approximately 10^4 at time $t = 19000$, while the initial population is less than 10. This loose crowding effect failed to constrain the population boom of r_3 until $t = 6000$, where its population already overshot to 10^5. Then, the other rates gradually took over r_3, and reached to the equilibrium point.

4. Results

To demonstrate the adaptability of Algorithm 1, we consider $r_1 = 0.9, r_2 = 0.7, r_3 = 0.5, r_4 = 0.1, \nu_\mathrm{i}(\mathrm{t}) = \omega_\mathrm{i}(\mathrm{t}) / \sum_{i=1}^{4} \omega_\mathrm{i}(\mathrm{t}), i = 1,2,3,4$, where

$$\omega_1(t) = \frac{6}{13}\left(2 + cos\left(\frac{\pi}{15000}t\right)\right), \omega_2(t) = \frac{2}{13}\left(2 + cos\left(\frac{\pi}{15000}t + \frac{3}{4}\pi\right)\right),$$

$$\omega_3(t) = \frac{3}{13}\left(2 + cos\left(\frac{\pi}{15000}t + \frac{3}{2}\pi\right)\right), \omega_4(t) = \frac{2}{13}\left(2 + cos\left(\frac{\pi}{15000}t + \pi\right)\right)$$ (see Fig. 3a), and choose $b = 0.01, d = 0.1, \delta = 0.2$ as hyperparameters.

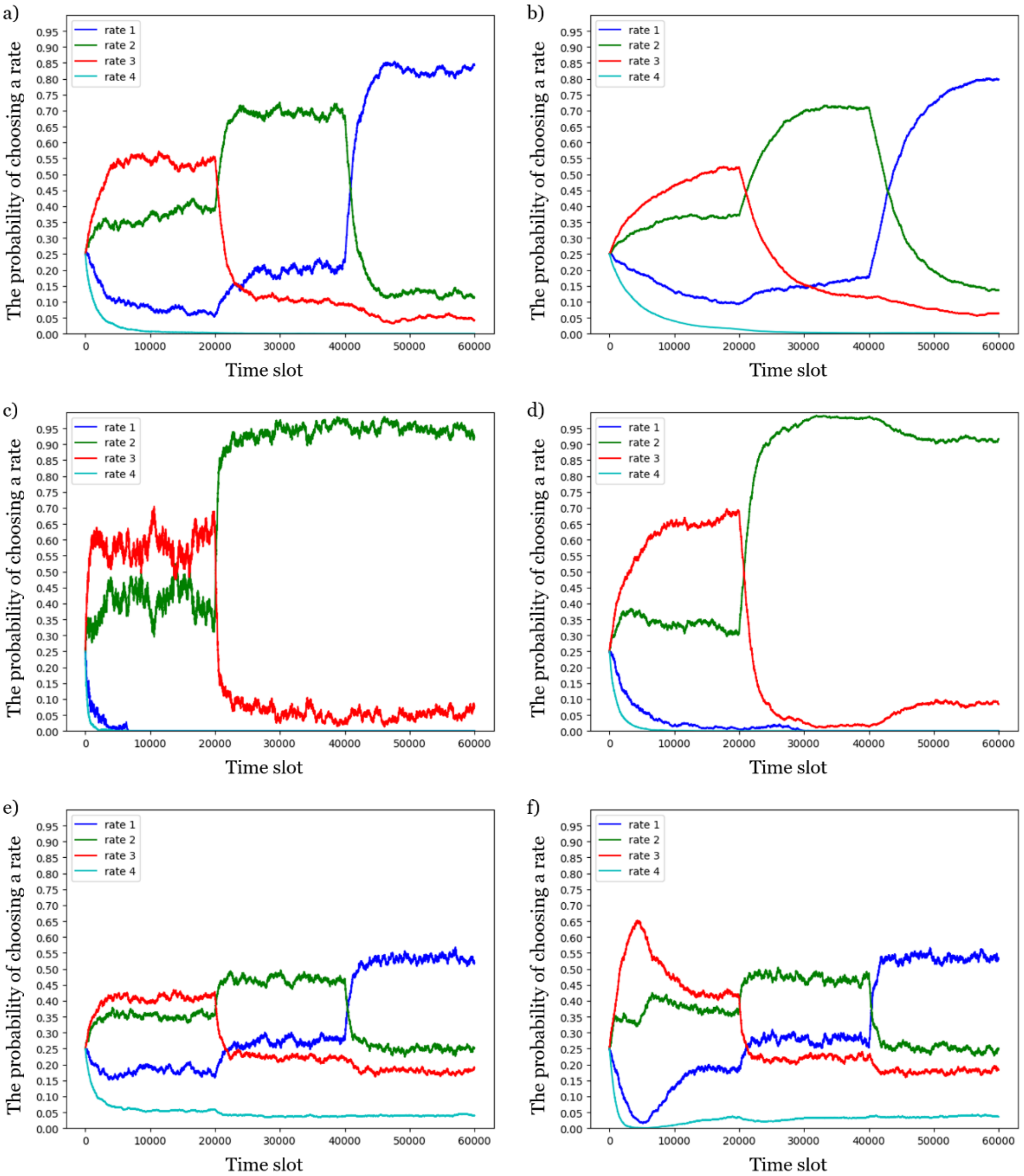

Fig. 2. The evolution of rate selection with given hyperparameters.
a). $b = 0.01, d = 0.9, \delta = 0.2$, b). $b = 0.003, d = 0.9, \delta = 0.2$, c). $b = 0.05, d = 0.9, \delta = 0.2$,
d). $b = 0.01, d = 0.9, \delta = 0.1$, e). $b = 0.01, d = 0.9, \delta = 0.5$, f). $b = 0.01, d = 0.001, \delta = 0.5$

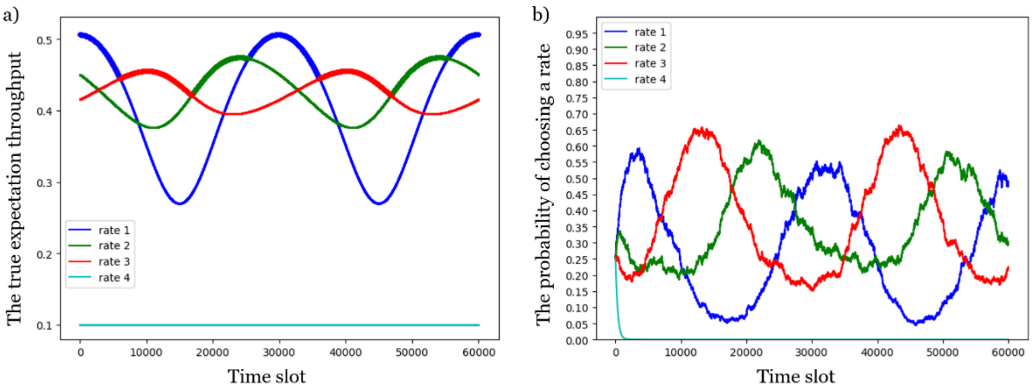

Fig. 3. The evolution of rate selection in the long-term nonstationary environment.

To evaluate the performance of Algorithm 1, define the optimality rate as:

$$opt = \sum_{\tau=0}^{t} r_{\hat{\imath}(t)} \times \theta_{\hat{\imath}}(t) \,/ \sum_{\tau=0}^{t} r_{i^*(t)} \times \theta_{i^*(t)}(t), \tag{7}$$

where $\hat{\imath}(t)$ denotes the rate chosen by a given algorithm at time slot t. We compared our algorithm with random policy and sticking to one rate, the results are shown in Table 2.

Table 2. Optimality rate of different policies.

Policy	Optimality Rate
Randomly selecting 4 rates	72.74%
Randomly selection first 3 rates	89.83%
Stick to Rate 1	86.37%
Stick to Rate 2	91.97%
Stick to Rate 3	91.13%
Stick to Rate 4	21.48%
Algorithm 1	93.78%

Algorithm 1 is about 2% better than sticking to the averaged optimal rate. The intuitive explanation of the performance of Algorithm 1 is the following: In Fig. 3b, we observe that when a rate provides the largest throughput in a short period, for instance, Rate 3 in $t = [7000{,}10000]$, it dominates in a temporary environment. Namely, we have a higher chance to transmit at this rate, which gives us a high throughput on average, while others are limited to a small probability for future exploration. Once its success rate decreases, due to the crowding effect, it can no longer hold such a large population (probability of being selected). In this case, for instance, Rate 1 in $t = [30000{,}35000]$, our algorithm will shift focus to exploring potentially better rates. When multiple rates are approximately optimal like $t = [33000{,}36000]$, our algorithm falls into uncertain zones, where it keeps probing channel states for future decision making. Note that this feature does not hurt much though it seems randomly picking rates, because all rates it explores are nearly optimal, which still gives us a large enough throughput.

5. Conclusion and Future Work

In this report, we consider the optimal rate selection problem in both long- and short- term varying wireless channels with limited feedback. We propose a low-complexity algorithm based on non-stationary MAB models and ecosystem dynamics and show its performance experimentally.

However, the major gap between non-stationary MAB problems and rate selection problems is transmission success rates $\theta_i(t)$ are correlated, while rewards of arms are independent of each other. This raises extra challenges to our algorithm. Yet, a Joint MAB problem takes this extra information into consideration, where arms can inform each other. It will be interesting to further metaphorize this Joint MAB problem to a different ecosystem.

Another key issue is hyperparameter selection, which can be considered as a meta-leveled selection problem. Since we already have our algorithm to deal with selection problems, we can also apply it to hyperparameter selection if we have a good scheme of rewards. But the definition of rewards for this problem requires more research.

6. Acknowledgment

This work is a final project for course CS 530 Principles of Artificial Intelligence, supervised by Prof. Casimir Kulikowski.

References

1. Gupta, H., Eryilmaz, A. and Srikant, R., 2018, April. Low-Complexity, Low-Regret Link Rate Selection in Rapidly-Varying Wireless Channels. In IEEE INFOCOM 2018-IEEE Conference on Computer Communications (pp. 540-548). IEEE.
2. Niiyama, T., Furuhata, G., Uchida, A., Naruse, M. and Sunada, S., 2019. Lotka–Volterra Competition Mechanism Embedded in a Decision-Making Method. Journal of the Physical Society of Japan, 89(1), p.014801.
3. Gavina, M.K.A., Tahara, T., Tainaka, K.I., Ito, H., Morita, S., Ichinose, G., Okabe, T., Togashi, T., Nagatani, T. and Yoshimura, J., 2018. Multi-species coexistence in Lotka-Volterra competitive systems with crowding effects. Scientific reports, 8(1), p.1198.
4. Lai, T.L. and Robbins, H., 1985. Asymptotically efficient adaptive allocation rules. Advances in applied mathematics, 6(1), pp.4-22.
5. Bouneffouf, D. and Rish, I., 2019. A survey on practical applications of multi-armed and contextual bandits. arXiv preprint arXiv:1904.10040.
6. Lotka, A.J., 1926. Elements of physical biology. Science Progress in the Twentieth Century (1919-1933), 21(82), pp.341-343.
7. Volterra, V., 1927. Variazioni e fluttuazioni del numero d'individui in specie animali conviventi (p. 142). C. Ferrari.
8. Hardin, G., 1960. The competitive exclusion principle. science, 131(3409), pp.1292-1297.