2022 International Symposium on Nonlinear Theory and Its Applications,
NOLTA2022, Virtual, December 12-15, 2022

Decision making for multi-armed bandit problem using two different dynamics in laser network

Keigo Sasaki†, Takatomo Mihana†, ‡, Kazutaka Kanno†, Makoto Naruse‡, and Atsushi Uchida†

†Department of Information and Computer Sciences, Saitama University, Japan
‡Department of Information Physics and Computing, Graduate School of Information Science and Technology, The University of Tokyo, Japan
Email: k.sasaki.602@ms.saitama-u.ac.jp, auchida@mail.saitama-u.ac.jp

Abstract– We experimentally demonstrate decision making to solve the multi-armed bandit problem using two different dynamics of low frequency fluctuations and chaos in a semiconductor laser network. We make decision by associating slot machines with semiconductor lasers in the laser network, and controlling the coupling strengths among the lasers, based on the result of slot machine selection. We compare the performance of decision making between the low frequency fluctuations and chaos.

1. Introduction

The multi-armed bandit problem [1] is a reinforcement-learning problem whose goal is to maximize the total reward by repeatedly selecting slot machines with unknown hit probabilities. There are two types of actions: exploration to search and find the slot machine with the highest hit probability, and exploitation to maximize the total reward by using the knowledge obtained from exploration. There is a trade-off between exploration and exploitation. Too much exploration ensures that the slot machine with the highest hit probability is found, however, the total reward is reduced. On the contrary, too much exploitation results in incorrect estimation of the slot machine with the highest hit probability. Therefore, the balance between exploration and exploitation is important in the multi-armed bandit problem.

Decision making using light has been studied for solving the multi-armed bandit problem in recent years, and there is several methods using chaotic laser signals and thresholds [2,3], where the chaotic signal is compared with the threshold value to determine the slot machine selection, and the threshold value is adjusted to facilitate the selection of the slot machines according to the results of the selection process. This decision-making method has been applied for channel selection in wireless communications [4,5].

Laser networks provide different types of synchronization [6], and a method for solving the multi-arm bandit problem has been proposed using unidirectionally coupled semiconductor laser networks [7]. In the semiconductor laser networks, synchronization phenomena such as cluster synchronization and group synchronization occur [8,9]. In addition, different dynamics such as chaos and low-frequency fluctuations can be observed [10,11]. Furthermore, lag synchronization of chaos has been observed, and the timing of the dropouts of the low-frequency fluctuations is shifted and synchronized by the coupling delay time. The laser oscillating in advance is called the leader, and the following laser is called the laggard. This relationship is called the leader-laggard relationship [6]. Furthermore, it has been reported that the leader laser is spontaneously exchanged. This switching can be controlled by the coupling strength, which can be applied for decision making. However, it has not been reported whether low-frequency fluctuations are suitable for decision making in laser networks. It is important to investigate which dynamics are suitable for decision making in laser networks.

In this study, we experimentally perform decision making in a unidirectionally coupled semiconductor laser network using two dynamics, low-frequency fluctuations and chaos. We compare the decision-making performance between low-frequency fluctuations and chaos.

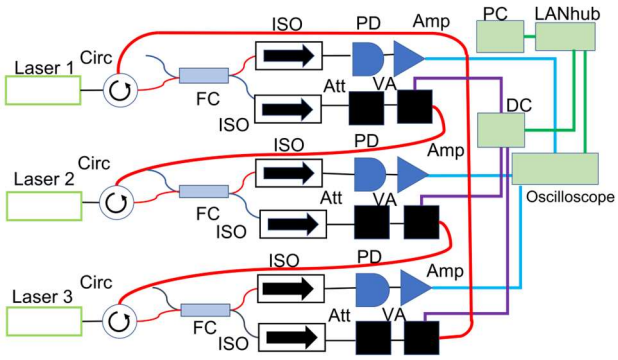

Fig.1 Experimental setup for decision making using semiconductor laser network. Circ: circulator, FC: fiber coupler, ISO: isolator, PD: photodetector, Amp: electric amplifier, VA: variable attenuator, Att: voltage driven variable attenuator, DC: direct-current voltage controller.

2. Methods

Figure 1 shows our experimental setup for decision making used in this study. Three semiconductor lasers are unidirectionally coupled by injecting light from laser 1 to 2, laser 2 to 3, and laser 3 to 1, respectively. The output from each laser propagates through a circulator and is split into two beams by a fiber coupler. One of the beams propagates through an isolator, is attenuated by an attenuator, and is injected into another circulator. The other beam is detected

by a photodetector and converted to an electrical signal by an electrical signal amplifier. Temporal waveforms of the laser output are measured by an digital oscilloscope.

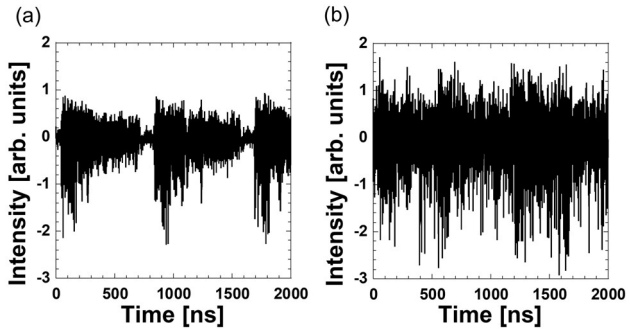

Fig. 2 Temporal waveforms of laser outputs in the case of (a) low-frequency fluctuations and (b) chaotic oscillations.

In this experiment, the injection current was varied to generate temporal waveforms of two different dynamics: low-frequency fluctuations and chaos. The injection currents of the three lasers were set to $(J_1, J_2, J_3) = (1.52J_{th1}, 1.62J_{th2}, 1.72\,J_{th3})$ and $(J_1, J_2, J_3) = (1.1J_{th1}, 1.1J_{th2}, 1.12J_{th3})$ for chaos and low frequency fluctuations, respectively, where J_{th} is the lasing threshold.

Figure 2 shows the comparison of the temporal waveforms between low-frequency fluctuations and chaos. Figure 2(a) shows the temporal waveform of low-frequency fluctuations. Sharp dropouts in the output and their gradual recovery are observed, which is the typical characteristics of low-frequency fluctuations. On the contrary, Figure 2(b) shows the temporal waveform of chaos, in which irregular chaotic oscillations are observed. Thus, different dynamics can be generated by varying the injected currents of the semiconductor lasers. These two dynamics are used for decision making.

We introduce a short-term cross-correlation value to evaluate the degree of chaos synchronization. The short-term cross-correlation value is expressed by the following equation.

$$C_n = \frac{< [I_n(t) - \bar{I_n}][I_{n-1}(t-\tau) - \bar{I}_{n-1}] >_\tau}{\sigma_n \sigma_{n-1}} \qquad (1)$$

This equation represents the short-term cross-correlation value of the temporal waveforms of the laser outputs between laser n and time-delayed laser n-1 (laser 0 is considered as laser 3). The leader laser can be determined by the minimum short-term cross-correlation value [7]. The leader laser can be controlled by changing the coupling strength. The coupling strength $\kappa_{n-1,n}$ from laser n-1 to laser n is expressed as follows.

$$\kappa_{n-1,n} = \frac{P_{inj}}{P_{max}} \qquad (2)$$

P_{inj} represents the light power injected from laser n-1 to laser n, and P_{max} represents the maximum light power injected from laser n-1 to laser n.

The leader probability is investigated by varying the coupling strength. The leader probability represents the probability of being the leader for laser n, expressed as the following equation.

$$L_n = \frac{T_n}{T_{total}} \qquad (3)$$

T_n is the duration that laser n is the leader, and T_{total} is the total duration. By decreasing the light power injected into laser n, the leader probability of laser n converges to 1 and the leader probabilities of the other lasers converge to 0.

Figure 3 shows the leader probabilities of laser n when the coupling strength κ_{31} is changed for the cases of low frequency fluctuations and chaos. By decreasing κ_{31}, the leader probability of laser 1 converges to 1 in both cases of low-frequency fluctuations and chaos, and the leader probabilities of laser 2 and 3 converge to 0. However, in the case of chaos, the convergence of the leader probability of laser 1 is faster than that in the case of low frequency fluctuations. Thus, the characteristics of the leader probability are different between the cases of low frequency fluctuations and chaos. Decision making is performed by controlling the leader by varying the coupling strength.

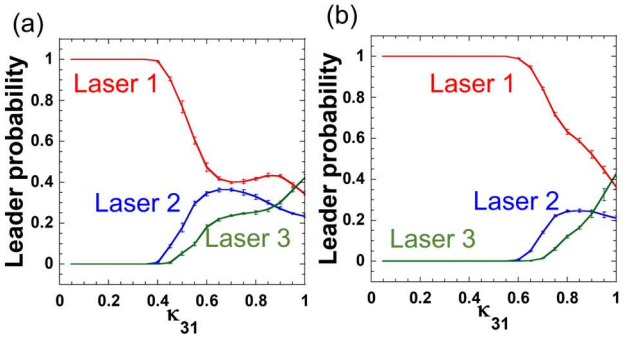

Fig. 3 Leader probabilities for the cases of (a) low-frequency fluctuations and (b) chaotic oscillations when the coupling strength κ_{31} is changed. Leader probabilities of laser 1 (red), laser 2 (blue), and laser 3 (green).

The decision-making process in the experiment is described as follows. Each slot machine is assigned to each laser in the laser network. The temporal waveforms of all laser outputs are acquired, and the short-term cross-correlation values are calculated using these temporal waveforms. The leader laser is identified based on the minimum short-term cross-correlation value. The slot machine corresponding to the leader laser is selected, and the coupling strength is changed based on the result of slot machine selection. For example, if laser 1 is the leader, the player selects slot machine 1 and plays it. If the result is "hit", the coupling strength κ_{31} is decreased, so that the

leader probability of laser 1 increases and slot machine 1 is more likely to be selected. If the result is "miss", κ_{31} is increased, so that the leader probability of laser 1 decreases and slot machine 1 is less likely to be selected. In this manner, the coupling strength is controlled and the decision-making process is repeated.

The evaluation values used in the decision-making process are set as follows.

$$X_n(t) = Q_n(t) - \frac{1}{N-1}\sum_{j\neq n} Q_j(t) \quad (4)$$

$$Q_n(t) = 2H_n(t) - (\bar{P}_{1s} + \bar{P}_{2nd})U_n(t) \quad (5)$$

$$\bar{P}_n(t) = \frac{H_n(t)}{U_n(t)} \quad (6)$$

$X_n(t)$ and $Q_n(t)$ are the relative and absolute evaluation values, $U_n(t)$ is the number of slot machine selection, $H_n(t)$ is the number of hits, and N is the number of slot machines. $\bar{P}_n(t)$ is the estimated probability, where $\bar{P}_{1st}$ and $\bar{P}_{2nd}$ are the maximum and second-maximum estimated probabilities.

The coupling strength is controlled using the relative evaluation value $X_n(t)$ as follows.

$$\kappa_{n-1,n} = \begin{cases} \kappa_{min} (\kappa_{ini} - kX_n(t) < \kappa_{min}) \\ \kappa_{ini} - kX_n(t) (\kappa_{min} \leq \kappa_{ini} - kX_n(t) \leq \kappa_{max}) \\ \kappa_{max} (\kappa_{max} < \kappa_{ini} - kX_n(t)) \end{cases} \quad (7)$$

Where $\kappa_{n-1,n}$ is the coupling strength from laser n-1 to laser n, κ_{ini} is the initial coupling strength, κ_{min} is the minimum coupling strength, and κ_{max} is the maximum coupling strength. k is the step size of $X_n(t)$ for decision making and is set to 0.05.

3. Experimental results

We compare the decision-making performance in the cases of low-frequency fluctuations and chaotic oscillations. We set the hit probability P_n of slot machine n to (P_1, P_2, P_3) = (0.9, 0.1, 0.1), (0.8, 0.2, 0.2), and (0.7, 0.3, 0.3), respectively, by varying the difference in the hit probabilities. The selection of slot machine 1 with the maximum hit probability is the correct decision. Smaller difference in the hit probabilities implies that it is more difficult to identify the best slot machine with the maximum hit probability. In this experiment, 100 plays are considered as one cycle, and 100 cycles are conducted.

Figure 4 shows the change in the selected slot machines in one cycle in the decision-making process for the low-frequency fluctuations and the chaotic oscillations. In the case of low-frequency fluctuations (Fig. 4(a)), all slot machines are selected randomly until 90 plays, which is considered as exploration. Then, slot machine 1 is always selected after 90 plays, which is considered as exploitation. Therefore, we consider that correct decision making (selection of slot machine 1) can be achieved. On the contrary, in the case of chaotic oscillations (Fig. 4(b)), all slot machines are selected until a smaller number of plays (25 plays). After 25 plays, slot machine 3 is always selected, which is wrong decision making, because slot machine 3 is not the slot machine with the maximum hit probability. This result may be owing to the fact that the number of initial exploration is not enough in the case of the chaotic oscillations, compared with that of low-frequency fluctuations.

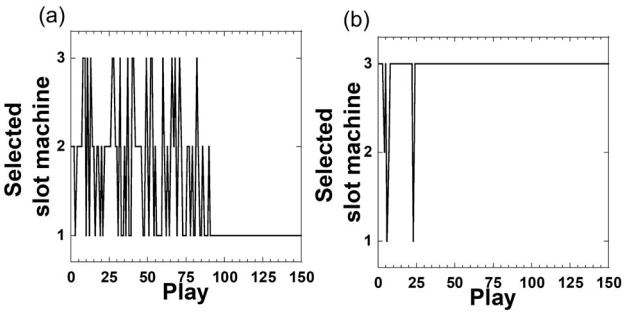

Fig. 4 Examples of selected slot machines at one cycle in the case of (a) low-frequency fluctuations and (b) chaotic oscillations. The hit probabilities are set to (P_1, P_2, P_3) = (0.7, 0.3, 0.3).

Next, the correct decision rate is used to compare the decision-making performance in the case of low-frequency fluctuations and chaotic oscillations. The correct decision rate represents the rate of slot machine selections with the maximum hit probability, expressed by the following equation.

$$CDR(t) = \frac{1}{m}\sum_{i=1}^{m} C(i,t) \quad (8)$$

Where $C(i, t)$ returns 1 if the selected slot machine has the maximum hit probability, and 0 otherwise. m is the number of cycles. An correct decision rate of 1 indicates the best decision-making performance.

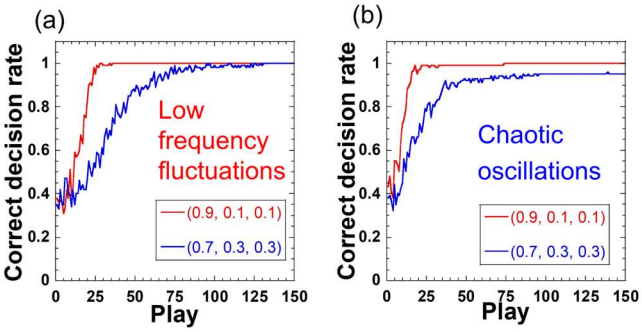

Fig. 5 Correct decision rate for the cases of (a) low frequency fluctuations and (b) chaos oscillations. The hit probabilities are set to (P_1, P_2, P_3) = (0.9, 0.1, 0.1) (red) and (0.7, 0.3, 0.3) (blue).

Figure 5 shows the result of correct decision rate for the low-frequency fluctuations and chaotic oscillations under the hit probabilities of (P_1, P_2, P_3) = (0.9, 0.1, 0.1) and (0.7, 0.3, 0.3). In the case of (P_1, P_2, P_3) = (0.9, 0.1, 0.1) (red), the correct decision rate converges to 1 in both cases. The convergence of the correct decision rate for the chaos oscillations is faster than that for the low-frequency fluctuations. However, when (P_1, P_2, P_3) = (0.7, 0.3, 0.3) is used (green), the convergence of the correct decision rate is slower than the case of (P_1, P_2, P_3) = (0.9, 0.1, 0.1) for both cases, because the difference in hit probabilities is smaller. Compared with the low-frequency fluctuations, correct decision rate for the chaotic oscillations is converged faster. However, the final correct decision rate for the low-frequency fluctuations reaches 1, while that for chaos is only 0.95. This result indicates that the chaotic oscillations provide a faster convergence of the correct decision rate, however, wrong decision is sometimes made when the difference in hit probabilities is small. On the contrary, in the case of low-frequency fluctuations, the convergence of the correct decision rate is slower, however, correct decisions are always made even when the difference in hit probability is small.

From these results, low-frequency fluctuations are more suitable for solving the problem with a small difference in hit probabilities because sufficient exploration is conducted, whereas chaotic oscillations are more suitable when the difference in the hit probabilities is large because the number of exploration is small.

4. Conclusions

We experimentally conducted decision making for solving the multi-armed bandit problem in a unidirectionally coupled semiconductor laser network. The decision-making performance was compared when the dynamics of low-frequency fluctuations and chaotic oscillations were used. In the case of chaotic oscillations, the convergence of the correct decision rate was faster because of less exploration, however, wrong decision was sometimes made when the difference in the hit probabilities was small. On the contrary, in the case of low-frequency fluctuations, the convergence of the correct decision rate was slow because of sufficient exploration, however, the correct decision was always achieved even when the difference in the hit probabilities was small. Therefore, the characteristics of the two different dynamics can be effectively used for different setting of the multi-armed bandit problem.

Acknowledgments

This study was supported in part by JSPS KAKENHI (JP19H00868, JP20K15185, JP20H00233, JP22H05195), JST CREST (JPMJCR17N2), and the Telecommunications Advancement Foundation.

References

[1] H. Robbins, “Some aspects of the sequential design of experiments,” Bulletin of the American Mathematical Society, Vol. 58, No. 5, pp. 527–536 (1952).

[2] Y. Han, S. Xiang, Y. Wang, Y. Ma, B. Wang, A. Wen, and Y. Hao, “Generation of multi-channel chaotic signals with time delay signature concealment and ultrafast photonic decision making based on a globally-coupled semiconductor laser network,” Photonics Research, Vol. 8, No. 11, pp. 1792–1799 (2020).

[3] J. Peng, N. Jiang, A. Zhao, S. Liu, Y. Zhang, K. Qiu, and Q. Zhang “Photonic decision-making for arbitrary-number-armed bandit problem utilizing parallel chaos generation” Optics Express, Vol. 29, No 16, pp.25290-25301 (2021).

[4] S. Takeuchi, M. Hasegawa, K .Kanno, A. Uchida, N. Chauvet and M. Naruse,”Dynamic channel selection in wireless communications via a multi-armed bandit algorithm using laser chaos time series” Scientific Reports, Vol. 10, Article No.1574 (2020)

[5] M. Shimomura, N. Chauvet, M. Hasegawa, and M. Naruse “Experimental demonstration of channel order recognition in wireless communications by laser chaos time series and confidence intervals,” Nonlinear Theory and Its Applications, IEICE, Vol. 13, No. 1, pp. 101–111 (2022).

[6] J. Ohtsubo, R. Ozawa, and M. Nanbu, “Synchrony of small nonlinear networks in chaotic semiconductor lasers” Japanese Journal of Applied physics, Vol. 54, pp. 072702 (2015).

[7] T. Mihana, K. Fujii, K. Kanno, M. Naruse, and A. Uchida, “Laser network decision making by lag synchronization of chaos in a ring configuration,” Optics Express, Vol. 28, No. 26, pp. 40112–40130 (2020).

[8] M. Nixon, M. Fridman, E. Ronen, A. Friesem, N. Davidson and I .Kanter, “Controlling synchronization in large laser networks” Physical Review Letters, Vol. 108, No. 21, pp. 214101 (2012).

[9] T. Dahms, J. Lehnert, and E. Schöll, “Cluster and group synchronization in delay-coupled networks,” Physical Review Letters, Vol. 86, No. 1, pp. 016202 (2012).

[10] T. Heil, I. Fischer, W. Els¨asser, J. Mulet, and C. R. Mirasso, “Chaos synchronization and spontaneous symmetry-breaking in symmetrically delay-coupled semiconductor lasers,” Physical Review Letters, Vol. 86, No. 5, pp. 795–798 (2001).

[11] K. Yamasaki, K. Kanno, A. Matsumoto, K. Akahane, N. Yamamoto, M. Naruse, and A. Uchida, “Fast dynamics of low-frequency fluctuations in a quantum-dot laser with optical feedback” Optics Express, Vol. 29, No. 12, pp. 17962-17975 (2021).