2017 International Symposium on Nonlinear Theory and Its Applications,
NOLTA2017, Cancun, Mexico, December 4-7, 2017

Adaptation of decision making with chaotic semiconductor laser

Takatomo Mihana[1], Yuta Terashima[1], Makoto Naruse[2], Song-Ju Kim[3] and Atsushi Uchida[1]

[1]Department of Information and Computer Sciences, Saitama University,
255 Shimo-Okubo, Sakura-ku, Saitama City, Saitama, 338-8570, Japan
[2]Strategic Planning Department, National Institute of Information and Communications Technology,
4–2–1 Nukui-kita, Koganei, Tokyo 184-8795, Japan
[3]Graduate School of Media and Governance, Keio University,
5322 Endo, Fujisawa, Kanagawa, 252-0882, Japan
Email: t.mihana.229@ms.saitama-u.ac.jp, auchida@mail.saitama-u.ac.jp

Abstract—Decision making using the tug-of-war (TOW) method can be adapted to the change of hit probability in the multi-armed bandit problem. We numerically demonstrate decision making using chaotic temporal waveforms generated from a semiconductor laser. We investigate the adaptation of decision making by changing the memory parameter.

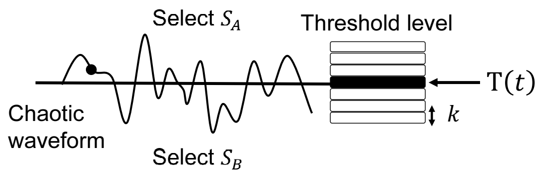

Figure 1: Schematic diagram of decision making using chaotic semiconductor laser.

1. Introduction

Decision making of artificial intelligence has been used in a variety of research fields in information society in recent years. One example of decision making is the Multi-Armed Bandit (MAB) problem [1] aimed at maximizing rewards from multiple slot machines. This is a problem where one can obtain information on hit or miss of the slot machines. However, one cannot directly know hit probabilities of the slot machines, and the hit probability may change over time.

For solving the MAB problem, the following two operations are important. One is the exploration, which is the procedure to select the slot machine with the highest hit probability. The other is the exploitation, which is the usage of the knowledge obtained by the exploration. To increase the total rewards, there is a trade-off between the two operations when the number of trials for selecting the slot machines is fixed. For example, one can correctly estimate the slot machine with the highest hit probability if the number of the exploration is increased. However, the total rewards could be small if the number of the exploration is too large. On the contrary, one cannot make sure whether the selected slot machine has the highest hit probability if the number of the exploration is too small.

In recent years, the tug of war (TOW) method [2, 3] has been proposed to solve the MAB problem. This method imitates the decision making algorithm of primitive organisms and utilizes nonlocality and irregularity. TOW method has been shown to be more adaptable than the conventional methods such as ε-greedy method and soft-max method [2]. As an example of TOW method, physical implementations with near-field photons [4, 5] and the quantum attributes of single-light quanta [6] has been reported. However, the operation speed in these systems is low (in the order of Hz).

Meanwhile, in recent years, physical implementation of high-speed decision making using chaotic waveforms in a semiconductor laser has been reported [7]. Due to its irregularly in GHz order, chaotic laser waveforms are used for physical random number generation [8, 9] and secret key distribution [10, 11]. Chaotic laser waveforms can be used to increase the speed of decision making. However, a limited number of studies have been reported on the decision making using chaotic laser waveform [7]. In addition, the influence of the memory parameter has not been investigated yet.

In this study, we numerically implement high-speed decision making by using TOW method and chaotic temporal waveforms in a semiconductor laser. We investigate the influence of the memory parameter on the decision making performance.

2. TOW method using chaos in semiconductor laser

We consider the two-armed bandit problem that satisfies the following conditions. The total number of the trials to draw slot machines is fixed, and the sum of the hit probabilities of two slot machines is set to 1.0. Furthermore, the sum of the hit probabilities is given to the decision making algorithm as prior knowledge.

In this study, we make decisions using experimentally obtained chaotic temporal waveforms generated from a semiconductor laser with optical feedback. Figure 1 shows the decision-making scheme based on the TOW method using chaotic temporal waveforms of the semiconductor

Table 1: Value of $X(t)$

Selected slot machine	Hit	Miss
S_A	-1	+1
S_B	+1	-1

laser. First, one threshold value is set for the chaotic waveform and the temporal waveform is periodically sampled. If the sampled data is larger than the threshold value, the slot machine S_A is selected. On the other hand, if the sampled data is smaller than the threshold value, the slot machine S_B is selected. Reinforcement learning is performed by changing the threshold value according to the result of the selected slot machine. For example, if the slot machine S_A is selected and the result is "hit", the threshold value is decreased (-1 in Table 1) so that the range of the sampled data for S_A is expanded. On the other hand, if the selected result is "miss", the threshold value is increased (+1 in Table 1) so that the range of the sampled data for S_A becomes narrower. In this way, by moving the threshold according to the result of "hit" and "miss", the range of the sampled data is adjusted and finally all the range becomes either S_A or S_B to make decision.

The threshold value $T(t)$ is changed with respect to the chaotic waveform as follows.

$$T(t) = \begin{cases} k(\mathrm{int})TA(t-1) & (|(\mathrm{int})TA(t-1)| \leq N) \\ kN & (|(\mathrm{int})TA(t-1)| > N) \end{cases} \tag{1}$$

where, k is the width of the threshold step, N is the number of threshold level. The threshold is changed between $-N$ and N (Total number of the threshold level is $2N+1$). These two parameters limit the range of the movement of the threshold. $TA(t)$ is the threshold adjuster variable, which is determined by the results of selecting the slot machines and past decision making. The threshold adjuster variable $TA(t)$ is defined as follows.

$$TA(t) = X(t) + \alpha TA(t-1) \tag{2}$$

where, $X(t)$ is the variable determined by the result of "hit" or "miss" with the selected slot machine, defined in Table 1. α is the memory parameter for weighting past threshold adjuster variables. The memory parameter α ranges in $0 \leq \alpha \leq 1$. For $\alpha = 0$, $TA(t)$ is determined by only the previous value of $X(t)$. On the other hand, for $\alpha = 1$, $TA(t)$ is determined by the results of all the past values of $X(t)$. We investigate the influence of α on the performance of decision making.

3. Experimental setup

We conducted experiments to acquire chaotic temporal waveforms of a semiconductor laser for decision making. Our experimental setup for the semiconductor laser is shown in Fig. 2. The output light of the semiconductor laser was reflected by a external mirror (Reflector), and

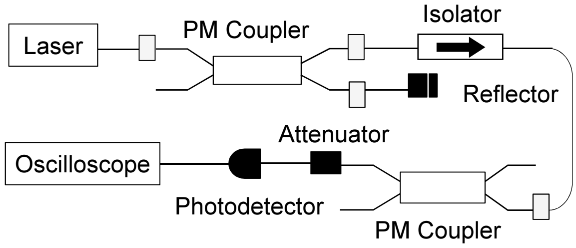

Figure 2: Experimental setup for generation of chaotic temporal waveforms in semiconductor laser.

feed back to the laser to generate chaos. The feed back light power was 210 μW. Chaotic output of the laser was injected into a photodetector and converted into an electric signal, and temporal waveforms were acquired by an digital oscilloscope with 8-bit vertical resolution.

4. Performance evaluation of decision making

4.1. Correct decision rate (CDR)

We used the parameter values in Table 2 for our numerical simulations. The number of trials for selecting slot machine is set to 5000. The hit probabilities of two slot machine S_A and S_B (defined as P_A and P_B) are set to $P_A = 0.4$ and $P_B = 0.6$, and switched every 1000 times. 5000 trials of selected the slot machines are defined as one cycle, and 1000 cycles are used for the evaluation of the performance of decision making. We evaluate the correct decision rate (CDR) at which the slot machine with the highest hit probability is selected. The CDR is defined as follows.

$$CDR(t) = \frac{1}{n}\sum_{i=1}^{n} C(i,t) \tag{3}$$

where, $C(i,t) = 1$ if the slot machine with the highest hit probability is selected, and $C(i,t) = 0$ otherwise, for the t-th decision making at the i-th cycle. In this evaluation, we can evaluate the number of trials for selecting the slot machine with the highest hit probability.

Figures 3(a) and 3(b) show the temporal evolutions of CDR when the memory parameter is small ($\alpha = 0.990$),

Table 2: Parameter values of decision making

Parameters	Symbols	Values
Memory parameter	α	0.990 or 0.999
Width of threshold step	k	64
Number of threshold level	N	2
Width of threshold steps	R	256
Hit probability of S_A	P_A	0.4 or 0.6
Hit probability of S_B	P_B	0.6 or 0.4
Flip interval	FI	1000
Number of trials	m	5000
Number of cycles	n	1000

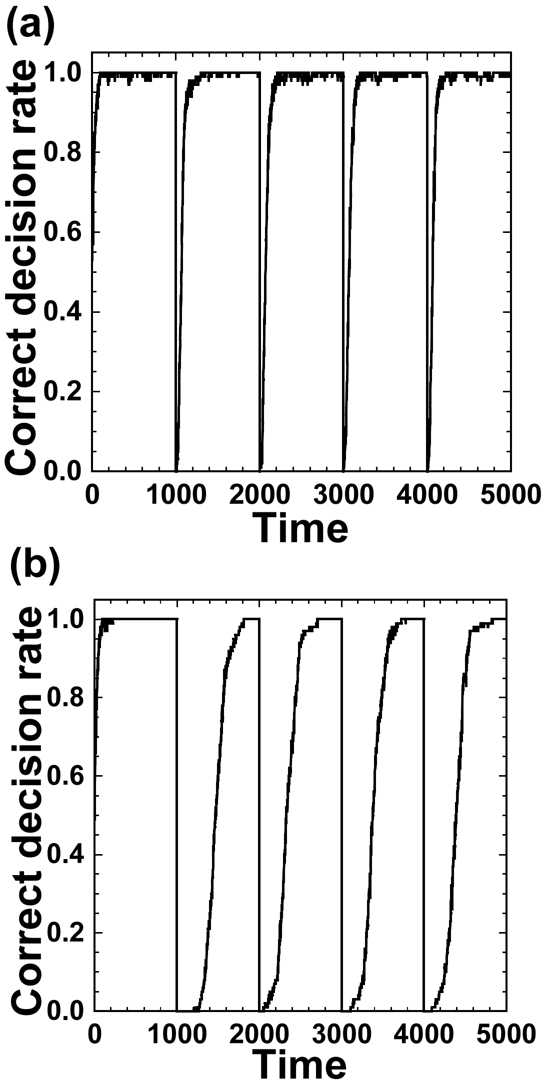

Figure 3: Temporal evolution of correct decision rate (CDR) for the memory parameter of (a) $\alpha = 0.990$ and (b) $\alpha = 0.999$ when the hit probabilities are changed for every 1000 times.

and large ($\alpha = 0.999$), respectively. As shown in Fig. 3(a), when the memory parameter is 0.990, CDR approaches 1.0 quickly even though the hit probabilities are switched, and the decision making can be adapted to environmental change (i.e., switching of the hit probabilities). However, CDR near 1.0 slightly fluctuates and does not converge to 1.0. On the other hand, when the memory parameter is 0.999 in Figure 3(b), it takes time until CDR approaches 1.0 after the hit probabilities are switched, and the adaptation speed is slow. However, CDR converges to nearly 1.0.

From this result, it is found that when the memory parameter is small, fast adaptation to environmental change can be obtained, however, the convergence to the correct decision is not completed. On the other hand, when the

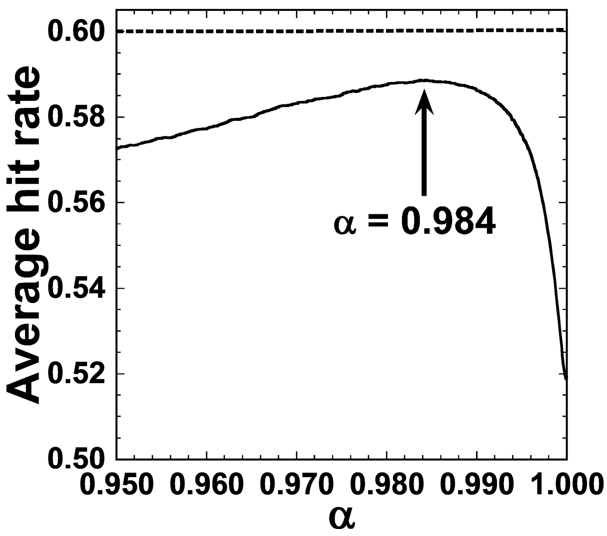

Figure 4: Average hit rate as a function of the memory parameter. The dotted line indicates the maximum value of average hit rate ($P_A = 0.4$, $P_B = 0.6$).

memory parameter is large, slow adaptation to the environmental change is obtained, however, good convergence to the correct decision is achieved.

4.2. Average hit rate (AHR)

Next, we evaluate the average hit rate (AHR), corresponding to the total reward of the decision making. AHR is defined as follows.

$$AHR = \frac{1}{mn}\sum_{i=1}^{n}\sum_{j=1}^{m} H(i, j) \tag{4}$$

where, $H(i, j) = 1$ if the selected slot machine is "hit", and $H(i, j) = 0$ if the selected slot machine is "miss" for the j-th decision making at the i-th cycle. AHR represents the total reward acquisition rate.

Figure 4 shows AHR when the memory parameter α is continuously changed. The upper limit value of AHR is 0.6 because it is the highest hit probability when $P_A = 0.4$ and $P_B = 0.6$ are switched. From Fig. 4, the maximum value of the AHR is 0.589. In addition, the memory parameter at which AHR becomes the largest is $\alpha = 0.984$. We found that it is possible to maximize the reward by choosing an optimal memory parameter, due to the trade-off between the adaptation and the convergence.

5. Conclusions

We numerically investigated high-speed decision making based on TOW method by using temporal waveforms of a chaotic semiconductor laser. We performed decision making in two slot machines in which hit probabilities are switched, and evaluated the role of memory parameter. When the memory parameter is small, fast adaptation to environmental change is obtained, however, the convergence of the correct decision is not completed. On the other hand, when the memory parameter is large, slow adaptation is obtained, however, good convergence to the correct decision is achieved. Therefore, there is a trade-off between the adaptation and the convergence. Finally, we investigated the memory parameter, and the reward of decision making can be maximized by choosing an appropriate memory parameter.

The decision making method using chaotic temporal waveforms in a semiconductor laser could be used for applications in high-speed trading.

Acknowledgments

We acknowledge support from Grants-in-Aid for Scientific Research from Japan Society for the Promotion of Science (JSPS KAKENHI Grant Numbers JP16H03878 and JP17H01277). This work was supported in part by the Core-to-Core Program A. Advanced Research Networks of JSPS.

References

[1] H. Robbins, "Some aspects of the sequential design of experiments," *Bulletin of the American Mathematical Society*, Vol. 58, No. 5, pp. 527–535 (1952).

[2] S. -J. Kim, M. Aono, and M. Hara, "Tug-of-war model for the two-bandit problem: Nonlocally-correlated parallel exploration via resource conservation," *BioSystems*, Vol. 101, No. 1, pp. 29–36 (2010).

[3] S. -J. Kim, M. Aono, and E. Nameda, "Efficient decision-making by volume-conserving physical object," *New Journal of Physics*, Vol. 17, 083023 (2015).

[4] S. -J. Kim, M. Naruse, M. Aono, M. Ohtsu, and M. Hara, "Decision maker based on nanoscale photo-excitation transfer," *Scientific Reports* Vol. 3, 02370 (2013).

[5] M. Naruse, W. Nomura, M. Aono, M. Ohtsu, Y. Sonnefraud, A. Drezet, S. Huant, and S. -J. Kim, "Decision making based on optical excitation transfer via near-field interactions between quantum dots," *Journal of Applied Physics*, Vol. 116, No. 15, 154303 (2014).

[6] M. Naruse, M. Berthel, A. Drezet, S. Huant, M. Aono, H. Hori, and S. -J. Kim, "Single-photon decision maker," *Scientific Reports*, Vol. 5, 13253 (2015).

[7] M. Naruse, Y. Terashima, A. Uchida, and S. -J. Kim, "Ultrafast photonic reinforcement learning based on laser chaos," *Scientific Reports*, Vol. 7, 8772 (2017).

[8] A. Uchida, K. Amano, M. Inoue, K. Hirano, S. Naito , H. Someya , I. Oowada , T. Kurashige , M. Shiki , S. Yoshimori , K. Yoshimura and P. Davis, "Fast physical random bit generation with chaotic semiconductor lasers," *Nature Photonics*, Vol .2, No. 12, pp. 728–732 (2008).

[9] R. Sakuraba, K. Iwakawa, K. Kanno, and A. Uchida, "Tb/s physical random bit generation with bandwidth-enhanced chaos in three-cascaded semiconductor lasers," *Optics Express*, Vol. 23, No. 2, pp. 1470–1490 (2015).

[10] K. Yoshimura, J. Muramatsu, P. Davis, T. Harayama, H. Okumura, S. Morikatsu, H. Aida, and A. Uchida, "Secure key distribution using correlated randomness in lasers driven by common random light," *Physical Review Letters*, Vol. 108, No. 5, 051102 (2012).

[11] H. Koizumi, S. Morikatsu, H. Aida, T. Nozawa, I. Kakesu, A. Uchida, K. Yoshimura, J. Muramatsu, and P. Davis, "Information-theoretic secure key distribution based on common random-signal induced synchronization in unidirectionally-coupled cascades of semiconductor lasers," *Optics Express*, Vol. 21, No. 15, pp. 17869–17893 (2013).