Asymptotically Optimal Multi-armed Bandit Algorithms Aimed at Online Contents Selection

論 文 の 内 容 の 要 旨

論文題目　Asymptotically Optimal Multi-armed Bandit Algorithms Aimed at Online Contents Selection（オンラインコンテンツ選択のための漸近最適な多腕バンディットアルゴリズム）

氏　　名　　小宮山　純平

A multi-armed bandit problem is a crystallized instance of a sequential decision-making problem in an uncertain environment. The history of this problem at least goes back to the 1930s, and it has recently attracted attention in the machine learning community. This problem involves conceptual entities called arms, and a forecaster who tries to identify good arms from bad ones. At each round, the forecaster draws one of the K arms and receives a corresponding reward. The aim of the forecaster is to maximize the cumulative reward over rounds, which is achieved by running an algorithm that balances the exploration (acquisition of information) and the exploitation (utilization of information). The notion of strong consistency introduced by Lai and Robbins (1985) defines the optimal balance between the exploration and the exploitation for a certain class of robust algorithms. At the beginning of the 2000s, an algorithm based on the upper confidence bound was established. Around that time, people found that many problems in web systems that involve uncertainty are related to the multi-armed bandit problem.

In each round of the multi-armed bandit problem, the algorithm selects an arm and receives a reward. In other words, the three core notions in the multi-armed bandit problem are (i) the sequential selection of arms, (ii) the criterion of selection (i.e., single arm selecting), and (iii) the reward feedback. However, these three notions are usually violated in practical applications: there are gaps between the multi-armed bandit framework and practical problems in web systems to which we would like to apply it.

In this thesis, we first review the history and the state-of-the-art framework of the multi-armed bandit. Then, we propose three extensions to the multi-armed bandit problem, which are intended to fill these gaps. Namely, (i) we propose the lock-up bandit problem, which models technical restrictions on the sequential selection of arms.

(ii) Motivated by the problem of online advertisement placement, we study a multiple-play version of the multi-armed bandit problem. We propose an extension of the Thompson sampling algorithm and show its effectiveness both theoretically and empirically. Moreover, (iii) motivated by problems arising in the information retrieval domain, we study the dueling bandit problem, a variant of the multi-armed bandit problem in which only the result of a pairwise comparison is available. A family of algorithms based on the likelihood function is proposed, and their effectiveness is verified.

見本 Sample

論 文 の 内 容 の 要 旨

論文題目　□□□□□□□□□□□□□□□□□□□□
（□□□□□□□□□□□□□□□□□□□□）

↑

（※論文目録の記載と同じにしてください。）

（※論文題目が外国語の場合には、和訳を括弧書きで付けてください。）

(* The title typed here must be identical to that shown in the Thesis Table of Contents.)

(* Add a Japanese translation in parentheses if the thesis title is written in a non-Japanese language.)

氏　　名　○○　○○

↑

（※学位記に記載される氏名と同じにしてください。）

（※漢字圏以外の外国人は、カタカナ表記となります。）

(* Type your name in the same manner as you want to have shown on your degree certificate.)

(* Type your name in katakana if you are a non-Japanese without a kanji name.)

　□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

１．**「論文の内容の要旨」は、紙媒体を４部提出してください。**
また、「PDF ファイル」及び「文書ファイル（Word 等で作成したもの）（省略可。）」の電子データも併せて提出してください。

２．**論文博士は日本文で**記入してください。（課程博士は英文でもよい。）
横書き、片面刷りとしてください。

３．大きさはＡ４判とし４ページ以内、１０ポイント程度の活字で印刷したものとしてください。
（日本語の場合は 4, 000 字以内（英語の場合は 2, 000 語以内）とする。）
４．第１ページ上部に、タイトルを「論文の内容の要旨」とした上で、論文題目及び氏名を記入し、その下から内容の要旨を記載してください。

1. **Four copies of your thesis summary must be submitted in paper form.**
   Electronic data of the thesis summary must also be submitted: a PDF file is mandatory, while submission of the original document file (MS Word or other) is optional.

2. **If you are obtaining your Doctorate degree by submitting a thesis (as a Ronpaku),** your thesis summary **must be written in Japanese. (**If you are obtaining your degree by completing the course requirements of a Doctorate program, a thesis summary in English is acceptable.)
   The thesis summary is formatted with **horizontal writing and single-sided print**.

3. The thesis summary is to be printed on **A4-size paper** and digested into **four pages or less** using **approximately a 10 point type**.
   (The restriction is **4,000 characters** for a Japanese summary and **2,000 words** for an English summary.)

4. **In the upper part of the first page, the text “論文の内容の要旨” is typed and the title of the thesis and the name of the applicant are typed on subsequent lines. The main text of the thesis summary begins below the above heading section on the same page.**