

## Yittoo's Blog

意图博客：关注 ML, NLP, 计算广告, Chatbots

# Contextual Multi-Armed Bandits

## Contextual Multi-Armed Bandits

### Part#1: multi-armed bandit problem (MBP)

[https://en.wikipedia.org/wiki/Multi-armed\\_bandit](https://en.wikipedia.org/wiki/Multi-armed_bandit)

In [probability theory](#), the **multi-armed bandit problem** (sometimes called the  $K$ -[\[1\]](#) or  $N$ -armed bandit problem[\[2\]](#)) is a problem in which a gambler at a row of [slot machines](#) (sometimes known as “one-armed bandits”) (老虎机：投币之后，摇杆，若出现三列图标一致，收益最高。) has to decide which machines to play, how many times to play each machine and in which order to play them.[\[3\]](#)

多臂强盗问题其实是指在多个独立的单臂强盗（老虎机）上的赌博行为，其目的是选择不同的老虎机以达到最大的收益。这样一个问题可以抽象成为多臂强盗问题可以用来建模现实中的很多场景。





## MULTI-ARMED BANDIT PROBLEMS

Multi-armed bandit (MAB) problems are a class of **sequential** resource allocation problems concerned with allocating one or more resources among several alternative (competing) projects. Such problems are paradigms of a fundamental conflict between making decisions (allocating resources) that yield high current rewards (**exploitation**), versus making decisions that sacrifice current gains with the prospect of better future rewards (**exploration**). The MAB formulation models resource allocation problems arising in several technological and scientific disciplines such as sensor management, manufacturing systems, economics, queueing and communication networks, clinical trials, control theory, search theory, etc.

<http://web.eecs.umich.edu/~teneket/pubs/MAB-Survey.pdf>

## Part#2 : Contextual Multi-Armed Bandits

- **Contextual Multi-Armed Bandits**, AISTATS 2010, Google, Inc. <http://www.jmlr.org/proceedings/papers/v9/lu10a/lu10a.pdf>
- 



- **A Contextual-Bandit Approach to Personalized News Article Recommendation**, WWW 2010, Yahoo!  
<http://www.research.rutgers.edu/~lihong/pub/Li10Contextual.pdf>
- 
- Estimation Bias in Multi-Armed Bandit Algorithms for Search Advertising, NIPS 2014  
<http://papers.nips.cc/paper/5146-estimation-bias-in-multi-armed-bandit-algorithms-for-search-advertising.pdf>
- 
- Online Algorithms For Parameter Mean And Variance Estimation In Dynamic Regression Models, 2016 <https://arxiv.org/pdf/1605.05697.pdf>
- 
- ONLINE ADVERTISEMENTS AND MULTI-ARMED BANDITS, PHD thesis 2015, <https://www.ideals.illinois.edu/handle/2142/78369>
- 
- Multi-armed Bandits: Applications to Online Advertising <http://www.lnmb.nl/conferences/2012/programlnmbconference/Zeevi-2.pdf>
- 
- An Empirical Evaluation of **Thompson Sampling**, NIPS 2012 <http://www.research.rutgers.edu/~lihong/pub/Chapelle12Empirical.pdf>
- 
- **Personalized** Recommendation via Parameter-Free Contextual Bandits, SIGIR 2015 <http://users.cis.fiu.edu/~ltang002/papers/sigir2015-tang.pdf>
- 
- todo

打赏

共享此文章：





相关

[Alibaba 2016 双11 之 人工智能篇](#)

2017-05-03

在“Chatbots”中

[Learn Reinforcement Learning](#)

2017-03-20

在“Reinforcement Learning”中

[Resources for Click Through Rate \(CTR\) prediction](#)

2016-03-30

在“Advertising”中



yittoo / 2016-05-28 / CTR、Machine Learning、Papers / CTR、ML

Yittoo's Blog / 自豪地采用WordPress

