Optimal Advertisement Strategies for Small and Big Companies

Tossou Aristide(✉) and Christos Dimitrakakis

Chalmers University of Technology, Gothenburg, Sweden
aristide@chalmers.se

Abstract. Many small and big companies in developing countries struggle to make their products or services known to the public. This is especially the case when there are new or have a new product. Most of them use publicity through radio, tv, social networks, billboard, SMS... Moreover, they also need to decide at what time to display their publicity for maximal effects. The companies which have more money typically used a simple strategy which consists in doing the publicity at many sources at different time or at a time such as to maximize the number of viewers. The smaller ones typically target the best popular programs.

However, this strategy is not the best as many users listening to your publicity might not be interested in it. So, you are more likely to miss the interested readers. Moreover, there will be many other competing publicities.

We propose a strategy by using the Multi-Armed bandit problem to optimally solve this problem under realistic assumptions. We further extend the model to deal with many competing companies by proposing the use of a time-division sharing algorithm.

Keywords: Bandit algorithm · Advertisement · Developing countries

1 Introduction

Advertisement is one of the most important component of the development of any company. This is especially true in developing countries where people tend to buy what their friends are buying. By choosing correctly when and where to make publicity, a company can then boost its profit significantly.

However, many companies in developing countries do not put too much effort in designing a strategy for where and when to make advertisements. They simply target the program which they believe is the most popular hoping that the more people see their publicity the more clients they will get.

This is not always the case, as the most potential interested users might not be available when a publicity is delivered. Or their publicity will be hidden by many other similar ones.

This problem is related to the budget allocation in marketing which has received a lot of attention. In [1], they modeled it as a bipartite graph where

R. Glitho et al. (Eds.): AFRICOMM 2015, LNICST 171, pp. 94–98, 2016.
DOI: 10.1007/978-3-319-43696-8_10

on one side are the customers and the other the medias. They then formulate it as a knapsack problem and solve the resulting maximization. This approach is further extended in [8] who proposes more efficient algorithms. However, their model only deal with a single advertiser. In [7], the problem is modeled in a game theoretic setting where advertisers are the players wishing to maximize their own utility. They solved it using a best response strategy. In our case, we solved this problem using the more natural multi-armed bandit (MAB) setting. All of the previous approach, only care about the budget to allocate to each media and not when is the appropriate time to advertise.

The MAB setting has previously being used for advertisement in [9]. But they solved the problem of which advertisement a search engine or media should display.

2 Modelisation

2.1 Problem Description

We assume that a company has at its disposal K spaces (combination of medias and time-slot) for publicity. For every new product, the company wants to make T publicities. The company incurs a cost of c_i for each publicity and expects in return to have many new clients giving him a profit of b_i. The overall goal of the company is to maximize its expected earnings after making T publicities.

We mapped this problem to the classical issue of exploration/exploitation which arises in many domains. Indeed, if a given publicity space leads to an immediate high reward, the company needs to keep choosing it to maximize its short term revenue. However, the company needs also to choose other spaces to check if its long-term revenue will not improve. This issue has been formally studied as the multi-armed bandit problem.

We first give the formal definition of multi-armed bandit and then detailed its mapping to our setting.

2.2 Multi-armed Bandit Problem

The K-armed bandit problem [5] involves an agent sequentially choosing among a set of K arms $\mathcal{A}$. At each time step t, the agent selects an action $a_t = i \in \mathcal{A}$ and obtains a reward r_t. The goal of the agent is to draw arms so as to maximize the total reward obtained after T interactions. An equivalent notion is to minimize the total regret against an agent who knew the best sequence of arms to play before the game starts. This is defined by:

$$\mathcal{R} \triangleq \mathbb{E}^{\pi_*} \sum_{t=1}^{T} r_t - \mathbb{E}^{\pi} \sum_{t=1}^{T} r_t. \tag{1}$$

where π_* is the optimal policy and π the one of the learning agent.

When the reward is drawn from some fixed unknown distribution, it leads to a stochastic MAB, otherwise it leads to an adversarial MAB. The stochastic MAB can be efficiently solved using the *UCB* algorithm [2] whereas the adversarial can be efficiently solved using the *EXP3* algorithm [3].

We now described the mapping from our setting to MAB. The set of K arms $\mathcal{A}$ is the set of spaces available for publicity. A reward is received for all previously played spaces at each *aggregation period* which can be taken to be a day for simplicity. We defined the reward as $r_t = \frac{p_s}{p_m \cdot c_s}$ with p_m the maximum profit possible for a *period*, c_s the percentage of costs incurred when choosing space s and p_s the profit due to space s for the *period*. We played a space at the end of each *period* when the budget allocated to publicity allowed it.

2.3 Dealing with Competition

In practice, however, the company is not the only one looking for publicity space. And they might be many other companies willing to advertise similar products. To deal with it, we consider the distributed multiple-player multi-armed bandit setting [6].

The multiple-player multi-armed bandit involves M players who have to choose between the same set of arms. They play simultaneously and whenever any two of them pick the same arm there is a collision. Two collision models are considered. When a collision happens, only one of the player receives the full reward and the other receive 0. The second collision model assumes that each player equally share the reward. The collision are observed by all players when it happens and the goal of each agent is to strategically maximize its own expected utility.

The mapping to the multiple-player multi-armed bandit to our setting is direct. And we can note that each player/company can observe collision by monitoring other publicities presented in the same slot as theirs or simply by asking the publicity provider.

To solve the distributed multi-armed bandits problem, we used the time division sharing technique [6]. In the first time step, the first player will target the best arm, and the k-th player target the k-th best arm. In the second time step, the first player will target the second best arm, etc. This idea can be modeled by associating an offset o to each player p. At each time step t, player p will target the $(t - 1 + o - 1) \mod M + 1$ best action. The offset used by each player is generated randomly. When a collision happens after a number of times, each player involved in the collision change its offset with probability 0.5 and keeps it old offset with the same probability (See [6]).

It is shown in [6] that if all players used this strategy they will get the same expected utility.

2.4 Determining Which Publicity Space Generate Which Customer

Our mapping to MAB requires that we are able to determine the exact publicity that a customer has followed.

This information can be obtained by asking customers to fill a survey where they are simply ask to tell which publicity push them to the company. However, customers might not be willing to tell us this information and when they do they could lie.

To solve this issue, we proposed to use the robust Bayesian truth serum (RBTS) [10]. We first described what is RBTS and then we described how it can be used in our setup.

The description of this RBTS is as follows:

- There are $n \geq 2$ agents who observe some signal about the same phenomenon and report it. Depending on the quality of the report, the agents receive a score or money. The goal of the agents is to maximize their expected score after multiple interaction with the system.
- It is assumed that all agents have the same prior belief about the outcome of the phenomenon and after observing the signal they all update their posterior similarly.

In [10], the score received by each agent is described by the following two steps:

1. Each agent i is asked to provide two reports:
   - Information report x_i which represents agent i's reported signal.
   - Prediction report y_i which represents agent i's prediction about the frequencies of signal values in the overall population.
2. Each agent i is linked with her peer agent $j = i + 1 \pmod{n}$ and is rewarded with a score:
$$\frac{1}{y_j(x_i)} \cdot \mathbb{1}_{x_j = x_i} + R(y_i, x_j)$$
   with the first term the information score, the second the prediction score, where R is a strictly proper scoring rule and $\mathbb{1}_{x_j = x_i}$ is the indicator variable.

The mapping of this setting in our system is as follows:

- The agents here are our customers. We assume that the client who viewed the same publicity share the same common belief a about the quality of our services. Then, after becoming our customers, they update their prior belief accordingly.
- We map the phenomenon to observe as the quality of our product/service.
- The signal is the publicity about the product/service.
- We will ask each customer three questions: The first is where/when he has heard about us, the second is the rating about the quality of our service/product (QOS) he guessed after viewing the publicity (x_i) and the last is the rating about our QOS when he interacts with us (y_i).
- We then assign a score to each of our client by choosing R as the Brier Score [4].

If each agent goal is to maximize its score, then the previous mechanism is Bayes-Nash incentive-compatible [10]. That is, all agents will report their true signals. If they don't, then their expected score will be lower.

To make the customers willing to increase their score, we can give them a monetary reward proportional to their score. For a supermarket, it could mean giving reduction on overall purchase which is proportional to the score of the customer. This could also be a discount in other stores or in the form of Internet, mobile credits bonus.

3 Conclusion and Outlook

We have presented an algorithm which will allow companies to optimally advertise their product. For that we observe that this problem is a standard exploration/exploitation issue which allows us to convert it to a Multi-Armed bandit problem. We also defined the reward sequence to be used. Furthermore, we extend the model to deal with competing companies by using the distributed multiple players multi-armed bandit which is solved using a time fair division sharing algorithm.

Finally, we present an incentivizing mechanism that will help companies to know where their customer heard about them. The Mechanism is a truthful one and encourage customers not only to participate but not to lie.

As future work, we would like to experiment this model with a company in Benin and check if the theory guarantees are observed in practice.

References

1. Alon, N., Gamzu, I., Tennenholtz, M.: Optimizing budget allocation among channels and influencers. In: Proceedings of the 21st International Conference on World Wide Web, WWW 2012, pp. 381–388 (2012)
2. Auer, P., Cesa-Bianchi, N., Fischer, P.: Finite time analysis of the multiarmed bandit problem. Mach. Learn. **47**(2/3), 235–256 (2002)
3. Auer, P., Cesa-Bianchi, N., Freund, Y., Schapire, R.E.: The nonstochastic multiarmed bandit problem. SIAM J. Comput. **32**(1), 48–77 (2003)
4. Brier, G.W.: Verification of forecasts expressed in terms of probability. Mon. Weather Rev. **78**(1), 1–3 (1950). 30 Sept 2015
5. Lai, T.L., Robbins, H.: Asymptotically efficient adaptive allocation rules. Adv. Appl. Math. **6**(1), 4–22 (1985)
6. Liu, K., Zhao, Q.: Distributed learning in multi-armed bandit with multiple players. IEEE Trans. Signal Process. **58**(11), 5667–5681 (2010)
7. Maehara, T., Yabe, A., Kawarabayashi, K.: Budget allocation problem with multiple advertisers: a game theoretic view. In: ICML 32, JMLR Proceedings, vol. 37, pp. 428–437. JMLR.org (2015)
8. Miyauchi, A., Iwamasa, Y., Fukunaga, T., Kakimura, N.: Threshold influence model for allocating advertising budgets. In: ICML 32, JMLR Proceedings, vol. 37, pp. 1395–1404. JMLR.org (2015)
9. Pandey, S., Olston, C.: Handling advertisements of unknown quality in search advertising. In: NIPS 20, pp. 1065–1072. MIT Press (2006)
10. Radanovic, G., Faltings, B.: A robust Bayesian truth serum for non-binary signals. In: AAAI 27, pp. 833–839 (2013)