POLITECNICO DI MILANO
Master of Science Degree in Engineering of Computing Systems
School of Industrial and Information Engineering

A comparative study of mechanisms for Sponsored Search Auctions

Advisor: Prof. Nicola Gatti

Master's Thesis by
Lorenzo Gentile 800721
Marianna Gentile 799054

Academic Year 2014-2015

Contents

List of Figures

List of Tables

Abstract

Sponsored Search Auctions are the workhorse auction mechanism for web-advertising, producing a revenue of about $19 billions in the U.S. alone in 2014. When a user submits a query on a search engine, the search engine returns a list of results including sponsored links. The latter represent sponsored ads: when a user clicks them, he is automatically directed to the advertiser's web page. Generally, search engines decide which ads to display and how to allocate them by using a pay-per-click auction mechanism. The purpose of this thesis is to study mechanisms for Sponsored Search Auctions in terms of theoretical properties and effectiveness through experimental analysis. After a study of the limits of VCG mechanism, we present and analyze a mechanism, RVCG, that goes beyond the limits of VCG and leads to ex-post Incentive Compatibility in multi-slot environments.

Sommario

I meccanismi d'asta per la visualizzazione di link sponsorizzati costituiscono la componenente principale della pubblicità attraverso Internet. Nel 2014 hanno generato negli Stati Uniti un profitto di 19 miliardi di dollari.
Quando un utente utilizza un motore di ricerca, oltre ai risultati effettivi della ricerca viene visualizzata una lista di link sponsorizzati. Tali link rappresentano le pubblicità sponsorizzate. Quando un utente vi clicca sopra, la pagina dell'inserzionista associato viene visualizzata. Generalmente i motori di ricerca decidono quali pubblicità visualizzare e in quale ordine utilizzando le cosiddette "Sponsored Search Auctions". Tali meccanismi d'asta impongono un pagamento agli inserzionisti solo nel caso in cui le loro pubblicità vengano cliccate. Lo scopo di questo lavoro è studiare questi meccanismi, sia in termini di proprietà teoriche che efficacia pratica. A tale proposito sono state effetuate delle analisi sperimentali. Dopo aver studiato i limiti del meccanismo VCG, presentiamo un meccanismo chiamato RVCG. Tale meccanismo permette di superare i limiti di VCG ed è applicabile al caso generale in cui più spazi pubblicitari sono disponibili all'asta.

Chapter 1

Introduction

1.1 Research area

In recent years, there has been a surge of interaction between computer scientists and economists. Computational microeconomics, including algorithmic game theory and algorithmic mechanism design, is receiving a lot of attention in computer science and artificial intelligence communities.
The expansion of the Internet has led to numerous examples of interaction of economic agents (e.g. web auctions). Game theory and economics have been rich in providing models and solution concepts as well as in prescribing strategies for rational agents. However, the outcomes proposed by the economic theory often involve problems with unknown efficient solutions. Resolving such complexity requires a combination of methodologies from computer science and economics.
Computational microeconomics provides the most elegant formal and algorithmic tools to deal with strategic interaction situations with multiple rational agents. Its main purpose is to improve economical transactions and automating them by using computational models.

Sponsored search auctions constitute one of the most successful applications of microeconomic mechanisms. It is now common to sell online ads using an auction. Auctions are used for search ads by Google, Microsoft and Bing/Yahoo!, for display ads by DoubleClick, for social network ads by Facebook.

In every year since 2005, the annual growth rates of Internet advertising have exceeded those of other advertising media. According to the latest IAB internet advertising revenue report [1], Internet advertising revenues in the United States totaled 49.5 billion dollars for the full year of 2014, in which search-related revenues contribute for 19.0 billion dollars alone.
In microeconomic literature, SSAs have been formalized as a mechanism design problem, where the objective is to design an auction mechanism that incentivizes advertisers to bid their truthful valuations (needed for economic stability) and that assures both the advertisers and the auctioneer to have a nonnegative utility. A crucial issue is estimating which advertisers tend to get clicked on more often. Naturally, whenever the mechanism display an ad which is not clicked, the mechanism receives no profit. Moreover, it is essential to study accurate models of the user behaviour and their effective exploitation in the economic mechanisms.

1.2 Original contributions

Our work is focused on the study of mechanisms for Sponsored Search Auctions in terms of theoretical properties and effectiveness through experimental analysis. First we show that, by using the VCG mechanism, it is possible to achieve both truthfulness in expectation and ex-post truthfulness in the single-slot setting.
Then we consider multi-slot settings affected by only position-dependent externalities. We show that, generally, by using VCG it is not possible to obtain neither truthfulness in expectation nor ex-post truthfulness. Under the restriction of click precedence property it is possible to obtain truthfulness in expectation when VCG is adopted, ex-post truthfulness is still not achievable.
For this reason we exploit the generic transformation presented in [2] in order to obtain a new mechanism (RVCG) for the multi-slot setting that, under the restriction of click precedence property, is ex-post truthful.
Finally, we simulate a real auction and we perform an experimental analysis in order to verify its practical usability.

1.3 Thesis structure

The thesis is structured in the following way.
Chapter 2 is devoted to the presentation of the state of the art. We introduce some concepts of mechanism design and the theory of Sponsored Search

Auctions, reporting the most important available results on which our work is based.
Chapter 3 is the heart of this thesis. It is devoted to the study of mechanisms for Sponsored Search Auctions under the assumption that all the parameters are known. We provide some theoretical results and we present an experimental analysis of an Incentive Compatible mechanism (RVCG) obtained by exploiting the generic transformation presented in chapter 2 in a multi-slot environment.
Finally, in Chapter 4, we summarize our work and we discuss about future possible developments.

Chapter 2

State of the art

In this chapter we introduce the basis and the most relevant works in literature we had to deal with during our research.
In Section 2.1 we introduce the main concepts of mechanism design, considering the mechanism properties we are interested in and presenting some of the main implementation results in the field. In Section 2.2 a description of Sponsored Search Auctions is presented, accompanied by the classification of the possible externalities and the presentation of known solution models. In the last section (Section 2.3.1) we report some of the main results about MAB mechanisms that we will exploit in our work.

2.1 Mechanism Design

Mechanism design has recently emerged as an important tool to model, analyze and solve *decentralized* design engineering problems involving multiple agents who interact *strategically* in a *rational* and *intelligent* way [3] [4]. A *rational* agent is an agent who makes decisions in order to maximize his own profit, that is measured through an utility scale. An *intelligent* agent is an agent that uses his knowledge about the underlying game in order to make inferences about it. In particular, *strategic* agents are taken into account, which means agents who take in consideration their expectation about the behavior of other agents and decide how to act basing on it.

The theory of mechanism design deals with settings in which a policy maker (or social planner) faces the problem of aggregating the *announced preferences* of multiple agents into a collective (or social) decision when the *actual preferences* are not publicly known.
Mechanism design theory uses the framework of *non-cooperative games with*

incomplete information and attempts to study how the private preference information can be elicited.
The main goal is to design mechanisms which satisfy some desirable properties, assuming that the individual agents can interact between them and will act in a strategic way.
Mechanism design can be viewed as the *art of designing the rules of a game to achieve a specific desired outcome.*

2.1.1 The Mechanism Design Problem

We now introduce the formal description of the mechanism design problem:

- There are n agents, indexed by $i = 1, 2, ...n$, who must make a collective choice from some set X, called the *outcome set.*

- Prior to the choice, each agent i privately observes his preferences over X. This is modeled by supposing that agent i observes a parameter θ_i (that is observable only by agent i) that determines his preferences. The parameter θ_i is said agent i's type. The set of possible types of agent i is denoted by Θ_i.

- The agents' types, denoted by $\theta = (\theta_1, ..., \theta_n)$ are drawn according to a probability distribution function $\Phi \in \Delta\Theta$, where $\Theta = \Theta_1 \times ... \times \Theta_n$, and $\Delta\Theta$ is the set of all the probability distribution functions over the set Θ. Let ϕ the corresponding probability density function.

- Each agent i is rational and intelligent. This is modeled by assuming that agents try to maximize a utility function $u_i : X \times \Theta_i \to \mathbb{R}$. The utility function $u(\cdot)$ of agent i depends on the outcome x and his type θ_i.

- The type sets $\Theta_1, ..., \Theta_n$, the probability density $\phi(\cdot)$, and the utility function u_i are common knowledge among the agents (it means that, even though the type θ_i is not common knowledge, for any type θ_i, every other agent can evaluate the utility function of agent i).

Given the above situation, we can now present the two main problems the social planner has to deal with.

Preference Aggregation Problem

The first problem is the following: *"For a given type profile $\theta = (\theta_1, ..., \theta_n)$ of the agents, which outcome $x \in X$ should be chosen?"*

In order to solve it, the social planner has to choose a *Social Choice Function.*

Definition 1 *A Social Choice Function (SCF) is a function $f : \Theta \to X$, used by the social planner to assign a collective choice $f(\theta_1, ..., \theta_n)$ to each possible profile of agents' types $\theta = (\theta_1, ..., \theta_n) \in \Theta$.*

Information Elicitation Problem

The problem of mechanism design does not end just by choosing a SCF.
The second problem is the following: *"How to extract the true type θ_i of each agent i, which is private information of agent i?".*
A trivial solution is to request the agents to reveal their types θ_i and use these information to compute the social outcome $x = f(\theta)$.
However, since the agents try to maximize their utility, it could be that agent i reveals an untruthful type, say $\hat{\theta}_i$, so as to drive the social outcome towards his most favorable choice.
A way in which the social planner can tackle this problem is the use of an appropriate *mechanism.*

Definition 2 *A Mechanism $M = ((S_i)_{i \in N}, g(\cdot))$ is a collection of action sets $(S_1, ..., S_n)$ and an outcome function $g : S_1 \times ... \times S_n \to X$.*

The set S_i describes the set of available actions for agent i. Based on his type θ_i, each agent i will choose some actions, say $s_i \epsilon S_i$.
Once all the agents have chosen their actions, the social planner uses this actions profile to select a social outcome $x = g(s_1, ..., s_n)$.

Definition 3 *Given a SCF $f : \Theta \to X$, a mechanism $D = ((\Theta_i)_{i \epsilon N}, f(\cdot))$ is said direct revelation mechanism corresponding to $f(\cdot)$.*

Given a SCF $f(\cdot)$, a direct revelation mechanism is a special case of a mechanism $M = ((S_i)_{i \in N}, g(\cdot))$, with $S_i = \Theta_i \ \forall i \in N$. Any other mechanism is said *indirect mechanism.*

Bayesian Game Induced by a Mechanism

A social planner can use an indirect mechanism M or a direct mechanism D to elicit information about agents' preferences in an indirect or a direct manner, respectively.

Recall that we assumed that agents are rational and intelligent. After knowing about the mechanism $M = ((S_i)_{i \in N}, g(\cdot))$ chosen by the social planner, each agent i selects an action $s_i : \Theta_i \rightarrow S_i$ in order to maximize his utility. This phenomenon leads to a game among the agents, that is said *Bayesian game induced by the mechanism* M and is denoted by Γ^b. It can be defined as follows:

$$\Gamma^b = (N, (S_i)_{i \in N}, (\Theta_i)_{i \in N}, \phi(\cdot), (u_i)_{i \in N})$$

The social planner now worries if the outcome of the game induced by the mechanism matches with the outcome of the SCF (if all the agents reveal their true types). This notion is captured in the definition that follows.

Definition 4 *The mechanism* $M = ((S_i)_{i \in N}, g(\cdot))$ *implements the SCF* $f(\cdot)$ *if there is a pure strategy equilibrium* $s^*(\cdot) = (s_1^*(\cdot), ..., s_n^*(\cdot))$ *of the Bayesian game* Γ^b *induced by* M *such that* $g(s_1^*(\theta_1), ..., s_n^*(\theta_n)) = f(\theta_1, ..., \theta n)\ \forall(\theta_1, ..., \theta n) \in \Theta$.

Depending on the underlying equilibrium concept, it is possible to implement a SCF in Dominant Strategy Equilibrium, Ex-Post Nash Equilibrium and Bayesian Nash Equilibrium.
We now define the notion of Dominant Strategy Equilibrium because it will be used extensively in the rest of this work.

Definition 5 *A pure strategy profile* $s^d(\cdot) = (s_1^d(\cdot), ..., s_n^d(\cdot))$ *of the game* Γ^b *induced by the mechanism* M*, is said to be a Dominant Strategy Equilibrium if and only if it satisfies the following condition.*

$$u_i(g(s_i^d(\theta_i), s_{-i}(\theta_{-i})), \theta_i) \geq u_i(g(s_i^{'}(\theta_i), s_{-i}(\theta_{-i})), \theta_i)$$

$$\forall i \in N, \forall \theta_i \in \Theta_i, \forall \theta_{-i} \in \Theta_{-i}, \forall s_i^{'}(\cdot) \in S_i, \forall s_{-i}(\cdot) \in S_{-i}$$

where S_i is the set of pure strategies of the agent i in the induced Bayesian game Γ^b, ans S_{-i} is the set of pure strategy profiles of all the agents except agent i.

2.1.2 Incentive Compatibility (IC)

Definition 6 *The SCF* $f(\cdot)$ *is said to be incentive compatible (or truthfully implementable) if the direct revelation mechanism* $D = ((\Theta_i)_{i \epsilon N}, f(\cdot))$ *has a pure strategy equilibrium* $s^*(\cdot) = (s_1^*(\cdot), ..., s_n^*(\cdot))$ *in which* $s_i^*(\theta_i) = \theta_i, \forall \theta_i \in \Theta_i, \forall i \in N$.

That is, truth telling by each agent constitutes an equilibrium of the game induced by D.
Therefore if the SCF $f(\cdot)$ is incentive compatible then the direct revelation mechanism D can implement it.

Based on the underlying equilibrium concept, three types of incentive compatibility are given below. Note that we will use the notation $u_i(\theta_i, \theta_{-i}, \hat{\theta}_i, \hat{\theta}_{-i})$ for the utility.

Definition 7 *A SCF $f(\cdot)$ is said to be Dominant Strategy Incentive Compatible (DSIC) (or Dominant Strategy Truthful or Truthfully Implementable in Dominant Strategy) if the direct revelation mechanism $D = ((\Theta_i)_{i\epsilon N}, f(\cdot))$ has a dominant strategy equilibrium $s^d(\cdot) = (s_1^d(\cdot), ..., s_n^d(\cdot))$ in which $s_i^d(\theta_i) = \theta_i, \forall \theta_i \in \Theta_i, \forall i \in N$.*

Such a condition is true iff:

$$u_i(\theta_i, \theta_{-i}, \theta_i, \hat{\theta}_{-i}) \geq u_i(\theta_i, \theta_{-i}, \hat{\theta}_i, \hat{\theta}_{-i}) \ \ \forall \theta_i, \ \forall \theta_{-i}, \ \forall \hat{\theta}_i, \ \forall \hat{\theta}_{-i}.$$

If the SCF $f(\cdot)$ is DSIC, whatever the reported types of other agents are, to report the true type θ_i is always in the best interest of agent i.

Definition 8 *A SCF $f(\cdot)$ is said to be Ex-post Nash Truthful (or Truthfully Implementable in Nash Equilibrium) if the direct revelation mechanism $D = ((\Theta_i)_{i\epsilon N}, f(\cdot))$ has a Nash equilibrium $s^n(\cdot) = (s_1^n(.), ..., s_n^n(\cdot))$ in which $s_i^d(\theta_i) = \theta_i, \forall \theta_i \in \Theta_i, \forall i \in N$.*

Such a condition is true iff:

$$u_i(\theta_i, \theta_{-i}, \theta_i, \theta_{-i}) \geq u_i(\theta_i, \theta_{-i}, \hat{\theta}_i, \hat{\theta}_{-i}) \ \forall \theta_i, \ \forall \theta_{-i}, \ \forall \hat{\theta}_i, \ \hat{\theta}_{-i} = \theta_{-i}.$$

Note that this condition is weaker than the one required for DSIC since, for each agent i, reporting his true type θ_i is guaranteed to be best strategy only when all the other agents behave truthfully.
The set of constraints required by ex-post Nash Truthfulness is generally strictly contained in the set of constraints required by Dominant Strategy Truthfulness. DSIC implies ex-post Nash Truthfulness but not vice versa. Therefore, when the evaluation of an agent over an allocation depends on the true types of other agents, it is not possible to obtain DSIC although ex-post Nash Truthfulness is guaranteed.
On the other hand, the utilities of the agents do not depend on the true types of other agents if the evaluations do not; in this case the constraints required for ex-post Nash Truthfulness are exactly the same required for DSIC.

Definition 9 *A mechanism is said to be Bayesian Incentive Compatible (BIC) (or Bayes Nash Truthful or Truthfully Implementable in Bayesian Nash equilibrium) if the direct revelation mechanism* $D = ((\Theta_i)_{i\epsilon N}, f(\cdot))$ *has a Bayesian Nash equilibrium* $s^b(\cdot) = (s_1^b(\cdot), ..., s_n^b(\cdot))$ *in which* $s_i^b(\theta_i) = \theta_i, \forall \theta_i \in \Theta_i, \forall i \in N$.

Such a condition is true iff:

$$E_{\theta_{-i}}[u_i(\theta_i, \theta_{-i}, \theta_i, \theta_{-i})] \geq E_{\theta_{-i}}[u_i(\theta_i, \theta_{-i}, \hat{\theta}_i, \hat{\theta}_{-i})] \ \forall \theta_i, \ \forall \hat{\theta}_i$$

This kind of truthfulness is weaker than ex-post Nash Truthfulness since it states that the best response of any agent i is reporting the truth only in expectation over the types of the other agents. It could be the case that, for some particular types of other agents, agent i would prefer to lie.
Ex-post Nash truthfulness implies Bayes Nash truthfulness but not vice versa.

The **revelation principle**, that is one of the fundamental results in the theory of mechanism design, states that if a SCF $f(\cdot)$ is implementable by an indirect mechanism $M = ((S_i)_{i\in N}, g(\cdot))$, then it is truthfully implementable by the direct mechanism $D = ((\Theta_i)_{i\epsilon N}, f(\cdot))$.
This result enables to restrict the inquiry about truthful implementation of a SCF to the class of direct revelation mechanism only.

2.1.3 Other desirable properties

Allocative Efficiency

The Allocative Efficiency (AE) property guarantees that, for any profile of agent's type, the social choice function provides a collective choice which is not Pareto dominated.Therefore, it is not possible to make any one individual better off without making at least one individual worse off.
Formally:

Definition 10 *A SCF* $f : \Theta \to X$ *is said to be Allocative Efficient if, for no profile of agents' type* $\theta = (\theta_1, ..., \theta_n)$*, does exist an* $x \in X$ *such that* $u_i(x, \theta_i) \geq u_i(f(\theta), \theta_i)$ $\forall i$ *and* $u_i(x, \theta_i) > u_i(f(\theta), \theta_i)$ *for some* i.

Individual Rationality

The *Individual Rationality* (IR) property ensures that each agent never receives negative utility by participating in the mechanism and reporting his

true type.
There are three stages at which an agent may decide to participate the mechanism:

- *Ex-post stage:* it arises after all the agents have announced their types and an outcome x has been chosen;
- *Interim stage:* it arises after all the agents have learned their type but before they have chosen their actions in the mechanism;
- *Ex-ante stage:* it arises before the agents have learned their type.

Definition 11 *In order to ensure agent i's participation at the ex-post stage, so as to obtain ex-post Individual Rationality, the following constraint must be satisfied: if agent i is truthful then* $u_i \geq 0 \ \forall\theta_i, \forall\theta_{-i}$.

Definition 12 *In order to ensure agent i's participation at the ex-interim stage, so as to obtain ex-interim individual rationality, the following constraint must be satisfied: if agent i is truthful then* $E_{\theta_{-i}}[u_i] \geq 0 \quad \forall\theta_i$.

Agent i has a probability associated for each agent and for each type.
The ex-interim individual rationality implies that if agent i is truthful then, in expectation over the types of the other agents, he gets a positive utility. It could be the case that for some particular types of other agents, utility of agent i is negative.
It is evident that ex-post individual rationality implies ex-interim individual rationality but the opposite is not true.

Definition 13 *In order to ensure agent i's participation at the ex-ante stage, so as to obtain ex-ante Individual Rationality, the following constraint must be satisfied: if agent i is truthful then* $E_{\theta_i}[u_i] \geq 0$.

2.1.4 Implementability results

Ideally a social planner would prefer to implement a SCF $f(\cdot)$ which is AE, DSIC and non-dictatorial.

Definition 14 *A SCF* $f : \Theta \rightarrow X$ *is said to be non-dictatorial if does not exist an agent d such that, for every profile of agents' type* $\theta = (\theta_1, ..., \theta_n)$, *holds* $f(\theta_1, ..., \theta_n) \in \{x \in X | u_d(x, \theta_d) \geq u_d(y, \theta_d \forall y \in X\}$.

The *Gibbard-Satterthwaite* impossibility theorem [5] [6] states that for a very general class of problems there is no SCF that satisfies the above properties simultaneously. In particular, a SCF $f(\cdot)$ is truthfully implementable in dominant strategy if and only if it is dictatorial.
Given this impossibility result, the social planner should attempt to find a satisfying SCF in an environment in which at least one of the hypothesis of the *Gibbard-Satterthwaite* theorem does not hold.
Quasi-linear environment is one of these environments in which all the SCFs are non-dictatorial.

Quasi-linear Environments

In quasi-linear environments the SCF gives as outcome the vector

$$x = (y, p_1, ..., p_n),$$

where y is the *allocation* and p_i are the *payments*.

The *utility* of agent i for a specific outcome x takes the form:

$$u_i(x, \theta_i) = v_i(y, \theta_i) - p_i,$$

where the term $v_i(y, \theta_i)$ is the *evaluation* of agent i over the allocation y.
The allocation y depends on the *reported types* $\hat{\theta}_i$ of the agents.
If y is fixed then $v_i(y, \theta_i)$ depends only on the true type of agent i while agent i's payment p_i depends on the reported types $\hat{\theta}_{-i}$ of all the other agents.

The amount $\sum_{i=0}^{n} p_i$ represents the *revenue* of the mechanism.

Generally, the utilities of the agents and the revenue of the mechanism can be negative, positive or equal to zero.

Budget Balanced

Definition 15 *A SCF is ex-post Weak (Strict) Budget Balanced (BB) if the following condition is satisfied:*

$$\sum_{i=0}^{n} p_i(\theta) \geq 0 \ (= 0) \ \forall \theta.$$

A SCF is ex-ante Weak (Strict) Budget Balanced if:

$$E_\theta[\sum_{i=0}^{n} p_i(\theta)] \geq 0 \ (= 0).$$

Intuitively, a mechanism is Weakly Budget Balanced when the auctioneer is guaranteed to have no loss. Note that ex-post BB implies ex-ante BB but not vice versa.

Allocative Efficiency

In this contest, the concept of *Allocative Efficiency* can be further characterized.

Definition 16 *A SCF $f : \Theta \to X$ is said to be Allocative Efficient (AE) if, for each $\theta \in \Theta$, the output $x(\theta) \in X$ satisfies the following condition:*

$$x(\theta) = \arg\max_{x \in X} \{\sum_{i=1}^{n} v_i(x, \hat{\theta}_i)\}$$

More intuitively, a mechanism is AE when the outcome x, chosen by the allocation function f, is the one that maximizes the Social Welfare (SW).

Groves mechanisms

Definition 17 *A mechanism that implements the SCF $f(\hat{\theta}) = (y, p_1, ..., p_n)$ which satisfies the allocative efficiency condition*

$$y = \arg\max_{x \in X} \{\sum_{i=1}^{n} v_i(x, \hat{\theta}_i)\}$$

and that uses the payment scheme

$$p_i(\hat{\theta}) = h_i(\hat{\theta}_{-i}) - \sum_{j=1, j \neq i}^{n} v_j(y, \hat{\theta}_j)$$

is called Groves mechanism.

It is possible to prove that Groves mechanisms are DSIC ($\hat{\theta}_i = \theta_i \ \forall\theta, \ \forall\hat{\theta}$). *Green-Laffont-Holmstrom* theorem [7][8] states that Groves mechanisms are the only DSIC mechanisms that can be obtained when the allocation rule $f(\hat{\theta})$ is AE.

Vickrey-Clarke-Groves mechanism

Definition 18 *A Vickrey-Clarke-Groves (VCG) mechanism is a particular case of Groves mechanism in which*

$$h_i(\theta_{-i}) = \max_{x} \{ \sum_{j=1, j \neq i}^{n} v_j(x_{-i}(\theta_{-i}), \theta_j)\}.$$

Linear Environments

Linear environments are a subclass of quasi-linear environments in which the following constraints hold:

- each agent i's type lies in the interval $\Theta_i = [\underline{\theta_i}, \overline{\theta_i}] \subset \mathbb{R}$ with $\underline{\theta_i} < \overline{\theta_i}$;
- agents' types are statistically independent, that is, the density $\phi(\cdot)$ has the form $\phi_1(\cdot) \times ... \times \phi_n(\cdot)$;
- $\phi_i(\theta_i) > 0 \ \forall \theta_i \in [\underline{\theta_i}, \overline{\theta_i}] \ \forall i = 1, ..., n$;
- each agent i's utily function takes the form $u_i(x, \theta_i) = \theta_i v_i(y) - p_i$.

Myerson's Characterization Theorem

In linear environment, in the single-parameter case (i.e. agents' types are scalars), the following theorem holds.

Theorem 1 *A mechanism that uses an allocation function f and a payment function p is truthful if and only if $\rho_i(f(\hat{\theta}_i, \hat{\theta}_{-i}))$ is non-decreasing (w.r.t. $\hat{\theta}_i$) and the payment p_i for each agent i satisfies*

$$p_i(\hat{\theta}_i) = h_i(\hat{\theta}_{-i}) + \hat{\theta}_i \rho_i(f(\hat{\theta}_i, \hat{\theta}_{-i})) - \int_0^{\hat{\theta}_i} \rho_i(z) dz$$

where $\rho_i(f(\hat{\theta}_i, \hat{\theta}_{-i}))$ is the reward of agent i when the allocation is $f(\hat{\theta}_i, \hat{\theta}_{-i})$ and $h_i(\hat{\theta}_{-i})$ is a constant [9], [10].

2.2 Sponsored Search Auctions

When a user submits a query on a search engine, the search engine returns a list of results. In this list, in addition to the algorithmic results (the real outcome of the research), some sponsored results appear (from here on *ads*). The ads are displayed in a way similar to the algorithmic results: they are usually located above or on the side of them and organized in a list. If the user clicks one of these ads (they are represented by links), he is automatically directed to the advertiser web page.
The advertiser pays a given price each time the link is clicked.
Search engines generally make use of an auction mechanism to decide how to allocate the different ads in the available positions (from here on *slots*) of the resulting list in the web pages.
This kind of mechanisms is called *Sponsored Search Auction* (SSA).

2.2.1 Preliminaries

There are some parameters that are shared by almost all the different models:

- $N = \{1, 2, \ldots, n\}$ is the set of indices of the ads a_i ($i \in N$). We assume that each advertiser has a single ad, so each advertiser i can be identified by ad a_i.
- $K = \{1, 2, \ldots, k\}$ is the set of indices of the slots s_j ($j \in K$). The indices are assigned to the slots in increasing order from the top to the bottom of the list. E.g. the first slot of the list has index 1, the last one has index k.
  We assume that the number of advertisers is higher or equal to the number of available slots: $|N| \geq |K|$.
- $v_i \in V_i \subseteq \mathbb{R}^+$ is the *value* obtained by advertiser i when ad a_i is clicked by an user. We denote as $v = \{v_1, v_2, \ldots, v_n\}$ the *value profile.* Each advertiser exactly knows his own value but does not know the values of other advertisers.
- $b_i \in V_i$ is the reported value (the bid) of advertiser i. It represents how much the advertiser is willing to pay for each click on his ad. We denote as $\mathbf{b} = \{b_1, b_2, ..., b_n\}$ the reported value profile.
- $q_i \in [0, 1]$ is the *quality* of ad a_i and it represents the probability that a user clicks ad a_i when visualized. The values of the qualities are not known a priori and need to be estimated.
- $CTR_i \in [0, 1]$ is the *Click Through Rate* and represents the probability that ad a_i is clicked by a user when displayed. It is obtained by the combination of quality q_i and the probability that the user visualizes it.
- $u_i \in \mathbb{R}$ is the *utility* of advertiser i;
- $\rho_{i,j}(t) \in \{0, 1\}$ is the *click realization* and indicates whether ad a_i gets a click if displayed at round t in slot s_j or not.

It is now possible to describe how a SSA works.
At the beginning each advertiser specifies a bid b_i and the auctioneer (in this case the search engine) has to select the best way to allocate the ads in the slots devoted to sponsored links. The auctioneer's goal is to solve the optimization problem (e.g. maximize the social welfare or the revenue)

designing mechanisms which are truthful in dominant strategies: every agent maximizes his utility u_i by bidding truthfully ($v_i = b_i$), for any bid of the other and for any click that would have been received.
In each round each advertiser i derives value v_i from clicks and pays the auctioneer a sum defined by a specific payment scheme.
Generally SSA adopts a *pay-per-click* payment scheme that considers positive payments to an advertiser only if his ad has been clicked.

2.2.2 User model and externalities

In literature there are several works that focus on modeling the user behavior. We will focus on the most common one, which assumes that the user scans the links from the top to the bottom. Slot number 1 is the top slot, and receives the maximum attention, while higher-numbered slots are situated in lower position in the page and naturally receive less attention.
By intuition, an high-quality ad can divert from another one, while a low quality ad can cause the user to abandon the page. For this reason it's reasonable to consider the externality effect among ads. The value of an ad impression on a page is affected not just by the slot that the ad is placed in, but also by the set of the other ads displayed on the page.
In [11][12] externalities are classified in three different categories:

1. *Position-dependent externalities*: in this case the probability that a user visualizes an ad is given only by the position of the slot in which the ad is displayed, not by the influence of the other ads. We define the parameter $\lambda_j \in [0, 1]$ for each slot j and we call it *prominence*. It is commonly assumed that λ_j is monotonically non-increasing in k.

2. *Ad-dependent externalities*: the probability that a user visualizes an ad depends solely on the ads that are placed in the slots, but not the slots themselves. We define the parameter $c_j(i) \in [0, 1]$ as the probability that a user observing ad i in the slot $j-1$ will observe the ad in the next slot (c_1 is set to 1 by definition). We call it *continuation probability*.
This models both high-quality ads, which satisfy the need of the user, and low-quality ads, which might frustate the user (i.e. if an ad a with very small continuation probability precedes another slot a', then a' is unlikely to be ever seen by the user).

3. *Position/ad-dependent externalities*: this is a combination of the previos two externalities. In this case the probability that a user observes an ad is influenced by both the position of the slot the ad is displayed in, and the ads that precede it in the list.

2.2.3 Model and allocation

The auctioneer needs to choose k ads to display between the n available ones. The goal is to maximize either the social welfare or the revenue of the auctioneer. In both cases it is necessary to solve an optimization problem in order to find the optimal allocation rule. If the objective is to maximize the social welfare we should select $l \leq k$ ads $a_1, a_2, ..., a_l$ and allocate them in slots $s_1, s_2, ..., s_l$ in order to maximize the value given by $\sum_{i=1}^{l} b_i CTR_i$.
Therefore it is necessary to predict the CTR_i for each ad a_i(i.e. the probability that an ad is clicked) and take this prediction into account during the allocation phase (i.e. when allocating the slots to ads).
Depending on the externalities taken into account, the computations of the CTR and the allocation are based on different parameters and therefore they give rise to three different models:

1. The simplest model is the one in which only position-dependent externalities are considered [13]. We call it *Separable Model* because it completely ignores externalities between ads and it is based on separable click through rates. This model is the basis for most works in the area. It assumes that the CTR_i of an ad a_i is the product of its quality q_i and the prominence λ_j of the slot s_j in which the ad a_i is displayed.

   $$CTR_i = q_i \lambda_j$$

   The main advantage of this model is its simplicity: an optimal allocation of the ad space can be obtained sorting the advertisers by decreasing $b_i q_i$.
   The maximization problem can be solved in polynomial time.

2. We refer to the second model as the *Cascade Model*. In this case only ad-dependent externalities are taken into account. We assume that all the users are identical and scan the ads in the same order, from top to bottom.
   Under this model, the CTR_i of ad a_i is obtained with the product of the quality q_i of the ad, and the probability of reaching the slot s_j in which ad a_i is displayed.
   Assuming ads a_1, ..., a_k are placed in slots s_1, ...,s_k, the *click through rate* of ad a_i is:

   $$CTR_i = q_i \prod_{j=1}^{i-1} c_j$$

The maximization problem can be solved in polynomial time by using dynamic programming [12].

3. The third model combines the two models previously described and it is the most realistic one since it considers both ad-dependent and position-dependent externalities.
   Assuming ads a_1, ..., a_k are placed in slots $s_1, \ldots, s_k$ and $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k$, the click through rate of ad a_i displayed in slot s_j is given by:
   $$CTR_i = \lambda_j q_i \prod_{l=1}^{i-1} c_l$$
   Note that the CTR of an ad depends not only on the preceding ads but also on the order in which they are allocated in the previous slots. So far there are no algorithms that can solve the maximization problem in polynomial time [12][14][15] .

2.3 MAB mechanisms for SSA

SSA provides an environment where a mechanism design problem is inherently coupled with a learning problem. The solution of the optimization problem requires knowledge of the parameters q_i, $c_j(i)$, λ_k for each ad a_i and each slot s_j, according to the model of externalities that is used.
In this section we are interested in the process in which the $CTRs$ are learned. Thus we will consider the problem of designing truthful Sponsored Search Auctions in which the process of learning the CTR is part of the game. Supposing that the advertisers bid truthfully, our problem can be reduced to the Multi Armed Bandit (MAB) one.

2.3.1 The Multi Armed Bandit problem

The Multi Armed Bandit (MAB) problem is a classical paradigm in machine learning in which an online algorithm chooses from a set of strategies in a sequence of trials so as to maximize the total payoff of the chosen strategies.

The name "multi-armed bandits" comes from the scenario in which a gambler faces several slot machines (sometimes known as "one-armed bandits") that look identical at first but produce different expected winnings. The gambler has to decide which machines to play at each time step, how many times to play each machine and in which order to play them. The objective of the gambler is to maximize the sum of rewards earned, receiving feedback only for the chosen decision.

The crucial aspect in MAB problems is the tradeoff between acquiring more information (*exploration*) and using the current information to choose a good machine (*exploitation*).
A *policy*, or *allocation strategy*, is an algorithm that chooses the next machine to play based on the sequence of past plays and obtained rewards.

The performance of a MAB algorithm is usually evaluated through the *regret*, originally introduced by Lai and Robbins in [16]. It represents the loss in reward: the difference between the optimal reward, obtained by choosing always the best machine, and the total reward that could be achieved through the strategy adopted.
Lai and Robbins showed that the regret for MAB problems has to grow at least logarithmically in the number of plays and they devised policies which asymptotically achieve it. In [17], Auer, Cesa-Bianchi and Fischer showed that the optimal logarithmic regret is also achievable uniformly over time, rather than only asymptotically, with simple and efficient policies like *UCB1*.
In *UCB1* each machine is played once during the initialization phase. Then, at each round, the machine j that maximizes $\overline{x}_j + \sqrt{\frac{2 \ln n}{n_j}}$ is played, where $\overline{x_j}$ is the average reward obtained from machine j, n_j is the number of times machine j has been played so far, and n is the overall number of plays done so far. There is not a distinct separation between exploration and exploitation, since the two steps are carried on at the same time.
This algorithm can be applied to our problem supposing that each advertiser bids truthfully. However, the advertisers are strategic and the *UCB1* algorithm cannot be applied directly to design a regret minimizing mechanism.

2.3.2 Social Welfare maximization and truthful MAB mechanisms in single-slot environment

Two main contributions are presented by Babaioff, Sharma and Slivkins in [18]: a structural characterization of dominant strategy deterministic truthful mechanisms, and a bound on the regret that such mechanisms must suffer due to truthfulness.
They found that truthful mechanisms must strictly separate exploration from exploitation and that they incur much higher regret than the optimal MAB algorithms.

The setting

In [18] it is considered a multi-round auction setting.
There is a set N of n agents, numbered from 1 to n, and a single slot.
Each agent i has a single ad a_i to which is associated a private value v_i for every click he gets and a reported value b_i, possibly different from v_i, that is the bid initially submitted by the advertiser(b_i is public information).

For given click realization ρ and bid profile b, the total number of clicks received by agent i in the time horizon T is denoted $C_i(b;\rho)$. Call $C = (C_1, ..., C_n)$ the *click allocation* for the allocation rule A.

A quality q_i is associated to each ad. The qualities are not known at the beginning and need to be estimated during the exploration phase.

The mechanism

The goal is to design a dominant strategy truthful mechanism that maximizes the social welfare.
The mechanism is an online algorithm which first solicits bids from the agents and then runs for T rounds, where T is the given *time horizon*.
At each round the mechanism selects an agent, displays his ad and receives feedback (if there was a click or not). Initially no information is known about the probability of each agent to be clicked.

The mechanism differentiates two distinct phases:

- *Exploration*
  Each ad is displayed T_0 times, where
  $$T_0 = n^{-\frac{2}{3}} T^{\frac{2}{3}} (\log T)^{\frac{1}{3}}.$$
  Since the ads are n, the exploration phase will last for $\tau = nT_0$ rounds.
  During this stage the allocation is not dependent from the bid profile b and no payments are charged to the agents.
  Let c_i be the number of clicks on ad a_i in the exploration phase. The qualities can be estimated as $\hat{q}_i = \frac{C_i}{T_0}$.
- *Exploitation*
  It lasts for $T - \tau$ rounds.
  Basing on the reported values b and the qualities $\hat{q}$ estimated in the previous phase, an agent i^* such that
  $$i^* \in \arg\max_{j \in N} \hat{q}_j b_j = \arg\max_{j \in N} c_j b_j$$

is chosen and selected to be displayed for all the rounds.

The pay-per-click payment scheme is used and the amount charged at each step to each agent is:

$$p_i = \begin{cases} \frac{\max\limits_{j \in N/\{i\}} (c_j b_j)}{c_i}, & \text{if click} \\ 0, & \text{if no click} \end{cases}$$

Note that the agent displayed will be the only one who always pays.

Upper bound regret

In the computation of the regret it is taken into account the difference in Social Welfare between the algorithm and the benchmark which always selects the best ad.
If the advertisers bid their true private values our problem would be equivalent to the classical MAB problem, whose achievable regret is $O(T^{1/2})$. Deterministic truthful mechanisms incur much higher regret than the optimal MAB algorithms since they need to separate exploration from exploitation.

The *worst-case regret* is

$$R(T, v_{max}) = O(v_{max} n^{1/3} T^{2/3} (\log T)^{1/3}).$$

Below the proof.

Proof. The mechanism is truthful and therefore $b = v$.
Since the exploration lasts for $\tau = nT_0$ rounds and no probability of click is taken in consideration during this phase, the exploration regret can not be greater than the one obtained by always selecting the ad with higher value (v_{max}):

$$R_{exploration} \leq nT_0 v_{max}.$$

By Chernoff bounds, for each agent i we have $P(|\hat{q}_i - q_i| > \epsilon) < \delta$, for $\epsilon = \sqrt{8 \log T / T_0}$ and $\delta = T^{-4}$.

For the exploitation phase, with probability $1 - \delta$ (when for each agent i all the estimates $\hat{q}_i$ lie in the intervals specified above), let $j = \arg\max\limits_{i \in N} q_i v_i$ and $i^* = \arg\max\limits_{i \in N} \hat{q}_i v_i = \arg\max\limits_{i \in N} c_i v_i$.
Then we have:

$$(q_{i^*} + \epsilon) v_{i^*} \geq \hat{q}_{i^*} v_{i^*} \geq \hat{q}_j v_j \geq (q_j - \epsilon) v_j$$

which implies

$$q_{i^*}v_{i^*} + \epsilon v_{i^*} \geq q_j v_j - \epsilon v_j$$

$$q_j v_j - q_{i^*}v_{1^*} \leq \epsilon(v_{i^*} + v_j) \leq 2\epsilon v_{max}.$$

With probability δ the expected regret is at most v_{max}.
Therefore, the regret in the exploitation phase is at most:

$$R_{exploitation} \leq \delta(T-\tau)v_{max} + (1-\delta)(T-\tau)2\epsilon v_{max} \leq \delta T v_{max} + (1-\delta)T2\epsilon v_{max}.$$

As a consequence, the total regret is as claimed:

$$R(T, v_{max}) = O(v_{max} n^{1/3} T^{2/3} (\log T)^{1/3}).$$

Truthfulness characterization

As stated in Theorem 1, a mechanism (A, P) for the MAB mechanism problem is truthful with unrestricted payment computation if and only if, for any given *click realization* ρ, the corresponding *click allocation* C is non-decreasing:

$$C_i(b_i^+, b_{-i}; \rho) \geq C_i(b_i, b_{-i}; \rho) \text{ for } b_i^+ \geq b_i$$

and the payment rule is given by

$$P_i(b_i, b_{-i}; \rho) = b_i C_i(b_i, b_{-i}; \rho) - \int_0^{b_i} C_i(x, b_{-i}; \rho) dx.$$

However, in Myerson payment scheme, it is supposed the full knowledge of $\rho_i(t)$ for each ad a_i at each round t. In our case we can observe $\rho_i(t)$ only if ad a_i is displayed at round t and we can not know the values of the click realization otherwise.

Definition 19 *An allocation rule A is pointwise monotone if increasing the bid cannot cause a loss of impression: for each click realization ρ, bid profile b, round t and agent i, if $A_i(b_i, b_{-1}; \rho; t) = 1$ then $A_i(b_{i^+}, b_{-1}; \rho; t) = 1$ for any $b_{i^+} \geq b_i$.*

Lemma 1 *Consider the MAB mechanism design problem. Let (A, P) be a normalized truthful mechanism such that A is a non-degenerate deterministic allocation rule. Then A is pointwise monotone.*

Definition 20 *A round t is called influential for a given click realization if, for some bid profile, changing the click realization at round t can affect the allocation in some future round t'.*

Definition 21 *An allocation rule A is called exploration-separated if for any given click realization ρ and round t that is influential for ρ it holds that $A(b, \rho, t) = A(b', \rho, t)$ for any two bid vectors b, b'.*

Proposition 1 *Consider the MAB mechanism design problem with two agents. Let A be a non-degenerate scale-free deterministic allocation rule. If (A, P) is a normalized truthful mechanism for some P, then it is exploration separated.*

In other words: *truthful implies exploration-separated*. In order to obtain a truthful mechanism it is necessary to strictly separate the phases of exploration and exploitation.

2.3.3 Revenue maximization and truthful MAB mechanisms in single-slot environment

In [19], Devanur and Kakade considered the problem of designing a truthful pay-per-click auction in which the click-through-rates ($CTRs$) of the ads are unknown to the auctioneer and only a single slot is available.
Such an auction faces the classic explore/exploit dilemma typical of MAB mechanisms. They show that the achievable regret in the expected revenue, under truthful restrictions, is $\Theta(T^{2/3})$ (where T is the number of rounds), while for bandit algorithm without truthful restrictions it is $\Theta(T^{1/2})$. The extra $T^{1/6}$ factor is called the *price of truthfulness*. With analogous considerations to those of [18] they show also that truthful mechanisms must separate exploration from exploitation.

The model

The model proposed in [19] considers a repeated auction, where the single slot is auctioned in each of T rounds.
There are n advertisers. At each round t, each advertiser bids a value $b_i(t)$, which is the reported value of i per click at round t. Then the auctioneer decides which ad will be displayed.
Let $x(t)$ be the allocation vector, and say $x_i(t) = 1$ iff ad i is displayed (since only one advertiser is allocated, $x_j(t) = 0$ for all $j \neq i$). The auction then observes the event $\rho_i(t)$, which is equal to 1 if the ad is clicked and 0 otherwise. The auction observes $\rho_i(t)$ if and only if $x_i(t) = 1$.

At the end of each round, the auction charges advertiser i the amount $p_i(t)$ only if i is clicked. The revenue of the auction is $A = \Sigma_{i,t} p_i(t)$.
It is assumed that advertiser i's true value for a click at round t is $v_i(t)$. Hence, the utility of i is $\Sigma_t(v_i(t)\rho_i(t)x_i(t) - p_i(t))$.
Using this notation, such an auction is truthful for a given sequence $\rho_{seq}\epsilon\{0,1\}^{n\times t}$, if bidding $v_i(t) = b_i(t)$ is a dominant strategy for all bidders. As the auction depends on the advertisers previous bids, an advertiser could try to manipulate their current bid in order to improve their future utility. If an auction is truthful for all $\rho_{seq}\epsilon\{0,1\}^{n\times t}$ then it is said to be *always truthful* and such manipulation is not possible. Subject to the constraint of being always truthful, the goal of the auction is to maximize its revenue. Let $T - Regret = \Sigma_{t=1}^{T} smax_i\{CTR_i b_i(t)\} - E_C[A]$ be the expected truthful regret of the auction. The first term is the expected revenue of the Vickrey auction that knows the true CTR_i's ($smax_i\{u_i\}$ is the second largest element of a set of numbers $\{u_i\}$). The second term is the expected revenue of the auction. Below it is presented an upper bound for $T - Regret$.

Upper bound analysis

The algorithm works as follows: for the first τ steps, the auction explores. This means that each ad is displayed for $\lfloor \tau/n \rfloor$ rounds. All payments are 0 during the exploration phase. Let $\widehat{CTR}_i$ be the empirical estimate of the CTR_i obtained during the exploration phase. With probability greater than 1-δ, the following upper bound holds for all i:

$$CTR_i \leq \widehat{CTR}_i + \sqrt{2\lfloor \tfrac{n}{\tau} \rfloor \log \tfrac{n}{\delta}} := \widehat{CTR}_i^{+}$$

For $t > \tau$ (which is the exploitation phase), the auction allocate the slot to the bidder i^* at time t which maximizes $\widehat{CTR}_i^{+} b_i(t)$, i.e. $i^* = argmax_i \widehat{CTR}_i^{+} b_i(t)$. The payment of i^* at time t is:

$$p_i(t) = \frac{smax_i \widehat{CTR}_i^{+} b_i(t)}{\widehat{CTR}_{i^*}^{+}}$$

The auction is truthful. It can be shown as follows: consider a set of positive weights ω_i. It is possible to costruct a truthful auction in the following manner: let the winner at time t be $i^* = argmax_i \omega_i b_i(t)$ and charge i^* the amount $\frac{smax_i \omega_i b_i(t)}{\omega_{i^*}}$ this time. This auction is truthful for any click sequence and for any T. The weights used by the auctions are $\omega_i = CTR_i^{+}$ which are not functions of the bids. Hence the auction is truthful since during the exploration phase the auction is truthful (for any set of weights).
For all t during the exploitation phase (all $t > \tau$) $E[\rho_{i^*}(t)] = CTR_{i^*}$. Hence,

the expected revenue of the auction at time t is $\frac{smax_i \widehat{CTR}_i^+ b_i(t)}{\widehat{CTR}_{i^*}^+} CTR_{i^*}$.
By construction holds that:

$$\frac{smax_i \widehat{CTR}_i^+ b_i(t)}{\widehat{CTR}_{i^*}^+} \leq b_{i^*}(t) \leq b_{max}$$

and with probability greater than $1 - \delta$:

$$\frac{smax_i CTR_i b_i(t)}{smax_i \widehat{CTR}_i^+ b_i(t)} \leq 1$$

since $CTR_i \leq \widehat{CTR}_i^+$ (with probability greater than $1-\delta$). Using these facts, the istantaneous regret is bounded as follows:

$$\begin{aligned}
& smax_i CTR_i b_i(t) - \frac{smax_i \widehat{CTR}_i^+ b_i(t)}{\widehat{CTR}_{i^*}^+} CTR_{i^*} \\
&= \frac{smax_i \widehat{CTR}_i^+ b_i(t)}{\widehat{CTR}_{i^*}^+} \left(\frac{smax_i CTR_i b_i(t)}{smax_i \widehat{CTR}_i^+ b_i(t)} \widehat{CTR}_{i^*}^+ - CTR_{i^*} \right) \\
&\leq b_{max} \left(\frac{smax_i CTR_i b_i(t)}{smax_i \widehat{CTR}_i^+ b_i(t)} \widehat{CTR}_{i^*}^+ - CTR_{i^*} \right) \\
&\leq b_{max} (\widehat{CTR}_{i^*}^+ - CTR_{i^*}) \\
&\leq b_{max} \sqrt{2 \frac{n}{\tau} \log \frac{n}{\delta}}
\end{aligned}$$

Hence, since there are $T - \tau$ exploitation rounds and τ exploration rounds (with no revenue), the expected regret is:

$$T - Regret \leq b_{max}((T - \tau)\sqrt{2 \frac{n}{\tau} \log \frac{n}{\delta}} + \tau + \delta T)$$

where the δT term comes from the failure probability. Choosing $\delta = 1/T$ and $\tau = n^{1/3} T^{2/3} \sqrt{log(nT)}$ it is obtained that $T - Regret = O(T^{2/3})$.

2.3.4 Truthful mechanisms with implicit payment computation in single-slot environment

Before the results of Babaioff, Kleinberg and Slivkins presented in [2], it was believed that computing payments needed to induce truthful bidding was somehow harder than simply computing the allocation. They showed the opposite: to create a randomized truthful mechanism is as easy as a *single* call to a monotone allocation rule. Their main result is a general procedure (referred as the *generic transformation*) that takes a monotone in-expectation allocation rule A for a single-parameter domain (auctions in which the private information of each agent is a single parameter: his value per item) and transform it into a randomized mechanism that is truthful in expectation and individually rational for every realization. The mechanism

obtained implements the same outcome as the original allocation rule with probability arbitrary close to 1, and requires evaluating that allocation rule only once.
Moreover its regret is $O(T^{1/2})$. This upper bound matches the information-theoretic lower bound in the same setting. This stands in contrast to the lower bound of [18] and [19], where it was shown that deterministic ex-post truthful MAB mechanisms have a regret of $\Omega(T^{2/3})$.
In [2] is also designed a new MAB allocation rule (UCB1) that is ex-post monotone and has a regret $O(T^{1/2})$ in order to obtain a randomized ex-post truthful MAB mechanism with the same regret.

The model

Let n be the number of agents and N be the set of agents. Each agent $i \in N$ has some private type $v_i \in V_i$. It is assumed that V_i is an open subset of $\mathbb{R}$. Let $V = V_1 \times ... \times V_n$ denote the domain of types and let $v \in V$ denote the vector of true types. There is a set of *outcomes* X. For each agent $i \in N$ there is a function $a_i : X \to \mathbb{R}_+$ specifying the *allocation* of agent i. The *value* of an outcome $x \in X$ for an agent $i \in N$ is $v_i \cdot a_i(x)$. The *utility* that agents $i \in N$ derives from outcome $x \in X$ when he is charged p_i is quasi-linear: $u_i = v_i \cdot a_i(x) - p_i$. A direct revelation mechanism M consists of the pair (A, P), there $A : V \to X$ is the *allocation rule* and $P : V \to \mathbb{R}^n$ is the *payment rule*, i.e. the vector of payment functions $P_i : V \to \mathbb{R}$ for each agent i. The vector of bids is denoted by $b \in V$. The mechanism picks an outcome $A(b)$ and charges agent i payment of $P_i(b)$. The allocation for agent i is $A_i(b) = a_i(A(b))$. Agent's i utility is

$$u_i(x_i, b) = v_i \cdot A_i(b) - P_i(b)$$

For a randomized allocation rule $A_i(b)$ and $P_i(b)$ denote the expected allocation and payment charged to agent i. The expectation is taken over the randomness of the mechanism. Sometimes is is considered explicitly the deterministic allocation and payment that is generated by a specific random seed of the mechanism w or by a specific random seed of the nature r. In these cases $A_i(b, w, r)$ and $P_i(b, w, r)$ are used.
The notion of truthfulness considered in the model is DSIC. It is useful to estabilish terminology to indicate when truthfulness holds not only in expectation but also for specific realizations.
The term *universally truthful* is used when the mechanism is truthful not only in expectation over the mechanism's randomness, but for every realization of that randomness. Similarly, the term *ex-post truthful* is used when the

mechanism is truthful not only in expectation over the nature's randomness, but for every realization of that randomness. These terms can be refered also to other properties. In order to emphasize that a property holds only in expectation over the nature's randomness the term *stochastically* is used. If all types are positive, in addition to individual rationality it is desiderable that all agents are charged a non-negative amount; this is said *no-positive-transfers* property. Finally, the *welfare* of a truthful mechanism is defined to be the total utility $\Sigma_i v_i \cdot A_i(t)$.

The generic transformation

The main one result of [2], that is the existence of the generic transformation with the desired properties, can be stated as follows.

Theorem 2 *Consider an arbitrary single-parameter domain with n agents. Let A be a monotone allocation rule for this domain. Then for each $\mu \in [0, 1]$ there exist a truthful mechanism $M = (\tilde{A}, \tilde{P})$ with the following properties:*

- *M executes a single call to $A(\tilde{b})$, with a pre-processing step to compute the modified bid vector $\tilde{b}$ using a* ***self-resampling procedure*** *(a procedure that samples a smaller bid value w.r.t. b), and a post-processing step to compute the payments. Both pre- and post-processing step do not execute a call to A.*

- *For any bid vector b and for any random seed of nature allocations $\tilde{A}(b)$ and $A(b)$ are identical with probability at least $1 - n\mu$.*

- *M is truthful, universally ex-post individually rational.*

- *If $V = \mathbb{R}^n_+$ (all types are positive), then M is ex-post no-positive-transfers, and never pays any agent i more than $b_i \cdot A_i(\tilde{b}) \cdot (\frac{1}{\mu} - 1)$.*

For [Myerson 1981] a mechanism (A, P) is truthful if and only if A is monotone and morover for the payment $P_i(b)$ for each agents must satisfy

$$P_i(b) = P_i^0(b_{-i}) + b_i A_i(b_{-i}, b_i) - \int_{-\infty}^{b_i} A_i(b_{-i}, u) du$$

where $P_i^0(b_{-i})$ does not depend on b_i.

The payment imposed by the generic transformation is equal, in expectation, to the payment defined by Myerson. In such a way truthfulness is obtained. An agent is charged for his reported type and she receives a random rebate that is an estimation of the integral that appears in the Myerson payment. This estimation is obtained by using a **method for estimating integrals** by evaluationing the integrand at a randomly sampled points. As said before, the computation of a payment doesn't require to execute a call to the allocation rule.

2.3.5 An ex-post truthful mechanism in multi-slot environment

Narahari and Mandal in [20] introduced two different contributions. First, they proved that when the user clicks are governed by ad-dependent externalities and position-dependent externalities, it is impossible to design an ex-post monotone allocation rule with sublinear regret. This impossibility result motivates their second contribution: they designed an ex-post truthful mechanism for multi-slot SSAs with sublinear regret for the case in which the $CTRs$ are affected by only position-dependent externality and follow *click precedence* property.

The setting

It is a multi-slot SSA. There is a set N of n agents, numbered from 1 to n, and a set K of k slots, numbered from 1 to k.
Each agent i has a private value v_i and submits a bid b_i. A quality q_i is associated to each ad a_i.
The *cascade model* is adopted to predict the user behavior and both position and ad-dependent externalities are taken into account.

Impossibility of Sublinear Regret for Ex-Post Monotone Allocations

It can be proved that, when the user clicks are governed by ad-dependent or position-dependent externalities on the ads, it is impossible to design an allocation rule with sublinear regret. The regret of any ex-post monotone allocation rule is $\Omega(T)$.
This also means that it is impossible to design an ex-post truthful mechanism in the presence of ad-dependent externalities.

Ex-Post Truthful Mechanism for Multi-Slot SSA

Motivated by the above impossibility result, it is considered the case in which only position-dependent externalities are taken into account and the clicks follow a special property called *click precedence*.

Definition 22 *The click precedence property is verified if:*

$$\rho_{i,j}(t) = 1 \implies \rho_{i,j'}(t) = 1 \forall j' \leq j.$$

In other words: if ad a_i is clicked at time t when displayed at slot s_j than ad a_i will also be clicked at time t if displayed at slots higher that s_j.

If the clicks follow this property, then an allocation rule is ex-post monotone if and only if:

$$\exists j' \leq j : A_i(b_i, b_{-i}, t) = j \implies A_i(b_i^+, b_{-i}, t) = j' \;\; \forall t, \; \forall b_i^+ > b_i, \; \forall i.$$

The proposed allocation rule maintains k lower confidence bounds $\{L_i^j\}_{j=1}^k$ and k upper confidence bounds $\{U_i^j\}_{j=1}^k$ for each agent i. For all rounds, $L_i^j \leq b_i q_i \leq U_i^j$ with high probability.

It maintains also k different *activation sets* S_{act}^j for $j = 1, \ldots, k$. S_{act}^j is the set of agents that can be assigned to slot j. Let $L_*^j = \max\{L_i^j\}_{i=1}^n$. Then S_{act}^j contains agent i if $U_i^j \geq L_*^j$.

The algorithm runs for T rounds. In every round the allocation rule decides which agents to allocate in the k slots in two steps:

1. *Designation procedure*: the allocation rule *designates* k different agents to allocate in the k available slots.

2. *Allocation:* the allocation rule verifies the presence of the designated agents in the corresponding activation sets: if j-th designated agent is present in S_{act}^j, then that agent is allocated to slot j. Otherwise, another agent is allocated to slot j: the agents are sorted in S_{act}^k in non-decreasing order of their k-th lower confidence bounds. Then all the designated agents are deleted from the sorted list and the top m agents (m is the number of free slot) are allocated to the m empty slots starting from top.

The resulting allocation rule is ex-post monotone if the click events follow click precedence property and the ads are affected only by position-dependent externalities.
Since the proposed allocation rule is ex-post monotone, using the procedure described in [2] , an ex-post truthful mechanism can be derived for the multi-slot SSA problem.

Regret Analysis

It is possible to prove that the total regret lower bound of the algorithm is:

$$R_{tot} = \tau b_{max} + (T - \tau) r_t + \delta T b_{max}$$

Chapter 3

SSA - Known parameters

This chapter is devoted to the study of truthful mechanisms in environments in which all the parameters are known to the auctioneer. The objective is to find a mechanism that is ex-post truthful without considering the problem of learning and estimating the unknown parameters.

In next sections we present the results obtained by studying the works reported in the previous chapter. In Section 3.1 it is showed how, by using the VCG mechanism, it is possible to achieve truthfulness in expectation (Section 3.1.1) and ex-post truthfulness (Section 3.1.2) in the single-slot setting. In Section 3.2 we extend the single-slot solution to the multi-slot setting and we consider the two different cases in which click precedence property does not hold and holds.
Due to the impossibility of obtaining ex-post truthfulness by using the VCG mechanism, in Section 3.3 we extend the method reported in Section 2.3.4 to the multi-slot setting. In Section 3.3.1 an apposite algorithm (RVCG) is introduced. Then, the setting and the results of an experimental analysis are reported in Section 3.3.2 and Section 3.3.3. It follows a comparison with theoretical results in Section 3.3.4.

3.1 VCG mechanisms and single-slot SSA

We consider a single-slot setting in which all the qualities q_i of the agents $i \in N$ are known to the auctioneer. In the single-slot case the VCG mechanism turns out to be truthful in expectation (3.1.1) and ex-post truthful (3.1.2).

3.1.1 Truthfulness in expectation

Statement: considering a single-slot SSA in which all the parameters are known, the mechanism obtained by the use of VCG is truthful in expectation.

Proof: given the allocation rule $A : \arg\max\{SW(\mathbf{b})\}$, where $SW(\mathbf{b}) = \sum_{i=1}^{n} b_i CTR_i$ and $\mathbf{b}$ is the vector of the values reported by the agents, it is easy to verify that it is AE and ex-post monotone. Therefore the resulting mechanism will be truthful.
Using the Groves payment scheme, to each advertiser i it is charged an amount equal to:

$$p_i = SW(\mathbf{b}_{-i}) - SW_{-i}(\mathbf{b})$$

where $SW(\mathbf{b}_{-i})$ represents the social welfare computed not considering the presence of agent i and $SW_{-i}(\mathbf{b})$ is the social welfare computed considering agent i but removing its contribute from the total sum.

We want to demonstrate that, for each agent, reporting the truth is always the best strategy.
We will do it presenting three different possible situations and considering only 2 agents for simplicity. Suppose that $q_1 v_1 \geq q_2 v_2$:

- *Scenario 1:* agent 1 bid truthfully ($b_1 = v_1$ and $q_1 v_1 \geq q_2 v_2$).
  This means that a_1 is allocated and the payment of agent 1 will be:

  $$p_1 = q_2 b_2.$$

  The expected utility of agent 1 is:

  $$u_1 = q_1 v_1 - p_1 = q_1 v_1 - q_2 b_2.$$

- *Scenario 2:* agent 1 does not bid truthfully ($b_1 \neq v_1$ and $q_1 b_1 < q_2 b_2$).
  This means that ad a_1 is not allocated. The payment of agent 1 will be null and the utility will be $u_1' = 0 \leq u_1$.

- *Scenario 3:* agent 1 does not bid truthfully ($b_1 \neq v_1$ and $q_1 b_1 \geq q_2 b_2$).
  This means that ad a_1 is still allocated and the payment of agent 1 will be:

  $$p_1'' = q_2 b_2.$$

  The expected utility of agent 1 is:

  $$u_1'' = q_1 v_1 - p_1'' = q_1 v_1 - q_2 b_2 = u_1.$$

In all the cases agent 1 can not receive an utility higher then the one obtained by bidding truthfully. Reporting the true value is the best strategy. The resulting mechanism is truthful in expectation.

3.1.2 Ex-post truthfulness

In order to obtain ex-post truthfulness, we suppose that the *click realization* $\rho_i(t)$ for each round t is known by all the agents $i \in N$ but not by the auctioneer.

Statement: considering a single-slot SSA in which all the parameters are known, the mechanism obtained by using the VCG is ex-post truthful.

Proof: using a VCG and pay-per-click mechanism in which the allocated agent i^* is given by $i^* = \arg\max_{i \in N} q_i v_i$ it is easy to observe that the allocation rule is ex-post monotone and therefore the resulting mechanism will be ex-post truthful. Moreover, since the payment is given by

$$p_i = \begin{cases} \dfrac{\max\limits_{j \in N/\{i\}} (q_j v_j)}{q_i}, & \text{if click} \\ 0, & \text{if no click} \end{cases}$$

it is possible to verify that the mechanism is also *Individual Rational* since $u_i \geq 0$ for each agent $i \in N$.
For simplicity we suppose to have only 2 agents ($N = \{a_1, a_2\}$) and that $q_1 v_1 \geq q_2 v_2$. Because of the previous hypothesis, the mechanism will select agent 1 ($i^* = 1$) and allocates ad a_1.
We compute the payment p_1 of agent 1 (note that $p_2 = 0$ since the ad a_2 is not displayed):

$$p_1 = \begin{cases} \frac{q_2 v_2}{q_1}, & \text{if } \rho_1(t) = 1 \\ 0, & \text{if } \rho_1(t) = 0 \end{cases}$$

and the respective utility

$$u_1 = \begin{cases} v_1 - \frac{q_2 v_2}{q_1} \geq 0, & \text{if } \rho_1(t) = 1 \\ 0, & \text{if } \rho_1(t) = 0 \end{cases}$$

The utility u_1 of agent 1 will always be positive or at least equal to 0, while $u_2 = 0$. Bidding a value b_1 higher or lower than the real value v_1 cannot generate an utility greater than the one obtained by a truthful bid.
Moreover agent 2 is not induced to lie bidding a value $b_2 \geq v_2$ in order to be allocated: in this case his utility u_2 would be negative since $v_2 \leq \frac{q_1 v_1}{q_2}$.

3.2 VCG mechanisms and multi-slot SSA

We now consider a multi-slot setting affected by only position-dependent externalities.

First we show that generally it is not possible to obtain ex-post monotone allocation rules (3.2.1) but it is in the particular case in which click precedence property holds (3.2.2). For this second case we demonstrate the truthfulness in expectation of the mechanism induced (3.2.3) and the impossibility of ex-post truthfulness (3.2.4).

3.2.1 Impossibility of ex-post monotonicity, truthfulness in expectation, ex-post truthfulness in absence of click precedence property

Statement: considering a multi-slot SSA, affected by only position-dependent externality in which all the parameters are known, the allocation rule A : $\arg\max\{\mathbb{E}[SW]\}$ (where $\mathbb{E}[SW]$ is the expected value with respect to clicks) is not ex-post monotone. The VCG mechanism obtained is not truthful in expectation nor ex-post truthful.

Proof: in order to have an ex-post monotone allocation rule, for each round t, agent $i \in N$ and $b_i \leq b'_i$, calling θ the *allocation vector*, if

$$\theta = \arg\max\{\mathbb{E}[SW(b_i, b_{-i})]\}$$

$$\theta' = \arg\max\{\mathbb{E}[SW(b'_i, b_{-i})]\}$$

it must hold that:

$$\rho_i(\theta, t) = 1 \implies \rho_i(\theta', t) = 1.$$

In the current case there are no limitations on the values the click realization can assume and it could be that $\rho_i(\theta, t) = 1$ and $\rho_i(\theta', t) = 0$.
Since the allocation rule is not ex-post monotone the resulting mechanism can not be truthful (Definition 1).
Intuitively, suppose that the mechanism does not know the *click realization* ρ but ρ_i is known by agent i. Agent i can manipulate the allocation vector by modifying his bid and be allocated in a different slot (higher or lower) basing on the knowledge he has about the future clicks.

The mechanism is not truthful in expectation.
To demonstrate it, consider a setting with 3 agents ($N = \{a_1, a_2, a_3\}$) and 2 slots ($K = \{s_1, s_2\}$). We assume that $q_1v_1 \geq q_2v_2 \geq q_3v_3$ and $\Lambda_1 \geq \Lambda_2 \geq \Lambda_3$. Supposing that agent 1 knows that, if displayed at slot s_1 his ad will not be clicked while if displayed at slot s_2 it will be:

Λ_1	a_1
Λ_2	a_2

- *Case 1:* agent 1 bids his true value ($b_1 = v_1$).

  Agent 1 is allocated in the first slot and will not be clicked. In this case the payment p_1 and the utility u_1 of agent 1 will be null.

- *Case 2:* agent 1 does not bid his true value ($b_1 \neq v_1$).

$$q_2 v_2 \geq q_1 b_1 \geq q_3 v_3$$

Λ_1	a_2
Λ_2	a_1

Agent 1 is displayed at slot s_2 and will be clicked. In this case the payment p_1' and the utility u_1' of agent 1 will be:

$$\mathbb{E}[p_1'] = \Lambda_1 q_2 v_2 + \Lambda_2 q_3 v_3 - \Lambda_1 q_2 v_2 = \Lambda_2 q_3 v_3$$

$$p_1' = \frac{\Lambda_2 q_3 v_3}{\Lambda_2 q_1}$$

$$u_1' = \Lambda_2 q_1 v_1 - \Lambda_2 q_3 v_3 = \Lambda_2 (q_1 v_1 - q_3 v_3) \geq 0$$

It is easy to observe that $u_1' > u_1$. Therefore the utility of agent 1 is greater when he bids not truthfully.

3.2.2 Click precedence and ex-post monotonicity

Statement: considering a multi-slot SSA, affected by only position-dependent externalities in which all the parameters are known, the allocation rule $A : \arg\max\{\mathbb{E}[SW]\}$ is ex-post monotone if the mechanism satisfies the *click precedence* property.

Proof: in order to have an ex-post monotone allocation rule, for each round t, agent $i \in N$ and $b_i \leq b_i'$, calling θ the *allocation vector*, if

$$\theta = \arg\max\{\mathbb{E}[SW(b_i, b_{-i})]\}$$

$$\theta' = \arg\max\{\mathbb{E}[SW(b_i', b_{-i})]\}$$

it must hold that:

$$\rho_i(\theta, t) = 1 \implies \rho_i(\theta', t) = 1.$$

If agent i reports a bid $b'_i \geq b_i$, his ad a_i will be displayed in a slot $s'_i \leq s_i$. Since the click precedence property is satisfied, $\rho_i(\theta) = 1 \implies \rho_i(\theta') = 1$ ad the utility of agent i can not decrease.

From Theorem 1, a necessary condition to obtain ex-post truthfulness is the ex-post monotonicity of the allocation rule and the use of Mayerson payments.
Note that, when the click precedence property holds, the above condition is satisfied (the allocation rule A is ex-post monotone and the payments used in VCG mechanism are a particular case of Mayerson payments). Nevertheless the ex-post truthfulness of the mechanism induced is not guaranteed. In the next two sections we will show that this mechanism results truthful in expectation but not ex-post truthful.

3.2.3 Click precedence and truthfulness in expectation

Statement: considering a multi-slot SSA, affected by only position-dependent externalities in which all the parameters are known, the VCG mechanism is truthful in expectation under the hypothesis of click precedence property.

Proof: Without loss of generality, we consider only 3 agents ($N = \{a_1, a_2, a_3\}$) and 2 slots ($K = \{s_1, s_2\}$) and we suppose that $q_1v_1 \geq q_2v_2 \geq q_3v_3$ and $\Lambda_1 \geq \Lambda_2$.
In order to show that the mechanism is truthful in expectation, we consider two different cases:

- *Case 1:* agent 1 bids his true value ($b_1 = v_1$).
  Since $q_1b_1 = q_1v_1 \geq q_2v_2 \geq q_3v_3$, the mechanism allocates ad a_1 in slot s_1 and ad a_2 in slot s_2.

Λ_1	a_1
Λ_2	a_2

  The payment p_1 which is charged to agent 1 is:

  $$p_1 = \Lambda_1 q_2 v_2 + \Lambda_2 q_3 v_3 - \Lambda_2 q_2 v_2$$

  and the expected utility u_1 is:

  $$u_1 = \Lambda_1 q_1 v_1 - \Lambda_1 q_2 v_2 - \Lambda_2 q_3 v_3 + \Lambda_2 q_2 v_2.$$

Note that the same utility is obtained if agent 1 bids a value $b_1 \neq v_1$ such that $q_1 v_1 \geq q_1 b_1 \geq q_2 v_2 \geq q_3 v_3$.

- *Case 2:* agent 1 bids a value $b_1 \neq v_1$ such that $q_2 v_2 \geq q_1 b_1 \geq q_3 v_3$. The mechanism allocates ad a_2 in slot s_1 and ad a_1 in slot s_2.

Λ_1	a_2
Λ_2	a_1

In this case the payment p_1' which is charged to agent 1 is:

$$p_1' = \Lambda_1 q_2 v_2 + \Lambda_2 q_3 v_3 - \Lambda_1 q_2 v_2 = \Lambda_2 q_3 v_3$$

and the expected utility u_1' is:

$$u_1' = \Lambda_2 q_1 v_1 - \Lambda_2 q_3 v_3.$$

It is now possible to demonstrate that the utility when agent 1 bids the true value, u_1, is not lower than the utility u_1' obtained when agent 1 bids untruthfully.

$$u_1 \geq u_1'$$
$$\Lambda_1 q_1 v_1 - \Lambda_1 q_2 v_2 - \Lambda_2 q_3 v_3 + \Lambda_2 q_2 v_2 \geq \Lambda_2 q_1 v_1 - \Lambda_2 q_3 v_3$$
$$\Lambda_1 (q_1 v_1 - q_2 v_2) \geq \Lambda_2 (q_1 v_1 - q_2 v_2)$$
$$\Lambda_1 \geq \Lambda_2 \leftarrow true\ for\ hypothesis.$$

3.2.4 Click precedence and ex-post truthfulness

Statement: considering a multi-slot SSA, affected by only position-dependent externality in which all the parameters are known, the VCG mechanism is not ex-post truthful in the hypothesis of pay-per-click payments and click precedence property.

Proof: consider 3 agents ($N = \{a_1, a_2, a_3\}$) and 2 slots ($K = \{s_1, s_2\}$) and suppose that $q_1 v_1 \geq q_2 v_2 \geq q_3 v_3$ and $\Lambda_1 \geq \Lambda_2$. The click realization ρ is known by all the agents, not by the auctioneer. Supposing that agent 1 knows that he will be clicked both in slot s_1 and s_2 at time t ($\rho_1(1, t) = \rho_1(2, t) = 1$):

- *Case 1:* agent 1 bids his true value ($b_1 = v_1$). Using the pay-per-click payment scheme:

$$p_1'^c = \begin{cases} \frac{(\Lambda_1 - \Lambda_2) q_2 v_2 + \Lambda_2 q_3 v_3}{\Lambda_1 q_1}, & \text{if click} \\ 0, & \text{if no click} \end{cases}$$

Λ_1	a_1
Λ_2	a_2

and, since $\rho_1(1,t) = 1$:

$$u_1 = v_1 - p_1^c = v_1 - \frac{\Lambda_2 q_3 v_3}{\Lambda_2 q_1}.$$

Note that the same utility is obtained if agent 1 bids a value $b_1 \neq v_1$ such that $q_1 v_1 \geq q_1 b_1 \geq q_2 v_2 \geq q_3 v_3$.

- *Case 2:* agent 1 bids a value $b_1 \neq v_1$ such that $q_2 v_2 \geq q_1 b_1 \geq q_3 v_3$. The mechanism allocates ad a_2 in slot s_1 and ad a_1 in slot s_2. In this

Λ_1	a_2
Λ_2	a_1

case the pay-per-click payment $p_1'^c$ is given by

$$p_1'^c = \begin{cases} \frac{\Lambda_2 q_3 v_3}{\Lambda_2 q_1}, & \text{if click} \\ 0, & \text{if no click} \end{cases}$$

and, since $\rho_1(1,t) = 1$:

$$u_1' = v_1 - p_1'^c = v_1 - \frac{\Lambda_2 q_3 v_3}{\Lambda_2 q_1}.$$

It is now possible to prove that the mechanism is not ex-post truthful by showing that $u_1 \leq u_1'$:

$$v_1 - \frac{(\Lambda_1 - \Lambda_2) q_2 v_2 + \Lambda_2 q_3 v_3)}{\Lambda_1 q_1} \leq v_1 - \frac{\Lambda_2 q_3 v_3}{\Lambda_2 q_1}$$

$$\Lambda_1(q_2 v_2 - q_3 v_3) - \Lambda_2(q_2 v_2 - q_3 v_3) \geq 0$$

$$\Lambda_1(q_2 v_2 - q_3 v_3) \geq \Lambda_2(q_2 v_2 - q_3 v_3)$$

$$\Lambda_1 \geq \Lambda_2 \rightarrow true\ for\ hypotesis.$$

Note that if ad a_i is not clicked, no amount is charged to agent i. Therefore $u_i \geq 0$ at each round and the resulting mechanism is ex-post Individual Rational.

3.3 Randomized mechanisms and multi-slot SSA

In the previous section we have demonstrated that, by using the VCG mechanism, it is not possible to obtain ex-post truthfulness in a multi-slot environment.
In this section we introduce a new mechanism in order to obtain ex-post truthfulness in the multi-slot setting when click precedence property holds. We called this mechanism Randomized VCG (RVCG).

3.3.1 RVCG mechanism

As suggested in [20], we tried to extend the method discussed in Section 2.3.4 to the multi-slot environment, giving as input of the generic transformation the AE allocation rule $A : \arg\max\{\mathbb{E}[SW]\}$.

We built a Matlab algorithm and tried to simulate a real auction in order to verify the truthfulness and the practical usability of the randomized mechanism in the multi-slot environment under the restriction of click precedence. The algorithm that we produced is reported in Algorithm 1.

Our procedure can be divided in five different phases:

1. *Values generation:* generate the vectors of the bids, the qualities and the prominences supported by an external generator.

2. *Self-resampling procedure:* execute each agent's self-resampling procedure to obtain two vectors of modified bids: x and y (Algorithm 2). We implemented the same procedure presented in Section 2.3.4.

3. *Allocation:* allocate according to the AE allocation rule $A : \arg\max\{\mathbb{E}[SW]\}$ but using the modified bid vector x.

4. *Click realization simulation:* the click realization is simulated by an apposite generator taking in consideration the ads $CTRs$.

5. *Payments computation:* the rebate R_i is generated for each agent i using the modified bid vector y. To each agent i is charged the amount $\rho\ (b_i A_i(x) - R_i)$.

Once these steps are accomplished, it turns out to be easy to compute all the desired parameters as the utility or the revenue of the auctioneer.

```
Data: μ,K,N
/* Values generation                                    */
[b, q, λ] = adSlotGenerator(N);
/* Self-resampling procedure                            */
[x, y] = selfResampling(b, μ);
/* Allocation                                           */
A = allocationRuleVCG(q, x);
/* Click realization simulation                         */
ρ = clickRealizationGenerator(q, λ, A);
/* Payments computation                                 */
for i=1:N do
    if y(i) < b(i) then
        R(i) = b(i)/μ;
    else
        R(i) = 0;
    end
    P(i) = ρ(i) * (b(i) − R(i));
end
```

Algorithm 1: Generic transformation with self-resampling procedure.

```
Data: b,μ
Result: [x,y]
for i=1:length(b) do
    if randomUniform(0, 1) < 1 − μ then
        x(i) = b(i);
        y(i) = b(i);
    else
        b(i)′ = randomUniform(0, b(i));
        x(i) = recursive(b(i)′);
        y(i) = b(i)′;
    end
end
```

Algorithm 2: Self-resampling procedure.

3.3.2 Experimental setting

We tested the algorithm on different settings in order to study the results in more situations.

Each experimental setting is defined by:

Data: b
Result: $bNew$
if $randomUniform(0,1) < 1 - \mu$ **then**
  $bNew = b$;
else
  $b' = randomUniform(0, b)$;
  $bNew = recursive(b')$;
end

Algorithm 3: Recursive.

- n: the number of agents, $n \in N = \{10, 100\}$;
- k: the number of slots, $k \in K = \{1, 2, 4, 6, 8, 10\}$;
- μ: the resampling probability, $\mu \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$;
- $M = 10$: the number of experiments done for each experimental setting.

We keep fixed the following parameters for each single experiment:

- $T = 100000$: the time horizon (i.e. the number of rounds of an experiment);
- (v_i, q_i): the values and the qualities of the ads;
- λ_k: the prominences of the slots.

Note that v_i, q_i and λ_k are generated from real data from *Yahoo! Webscope A3* dataset [14].

3.3.3 Experimental results

In this section we present the results of our experimental analysis, focusing on the desirable properties that a mechanism should own.

Allocative Efficiency

The randomized mechanism is Allocative Efficient with probability $1 - \mu$ since the Social Welfare is maximized only when the self resampling procedure causes no changes in the allocation.

In Figure 3.1 it is possible to observe how, varying the resampling probability μ, the curve of the mean Social Welfare (with respect to the time passed from the beginning of the auction) changes. The higher is μ and the higher is the difference with the optimal value obtained by the VCG. In particular, after a sufficient amount of time, the curve of the mean Social Welfare converges to the optimal value obtained by the VCG multiplied by μ. For this reason we say that RVCG loses a factor of efficiency equal to μ.

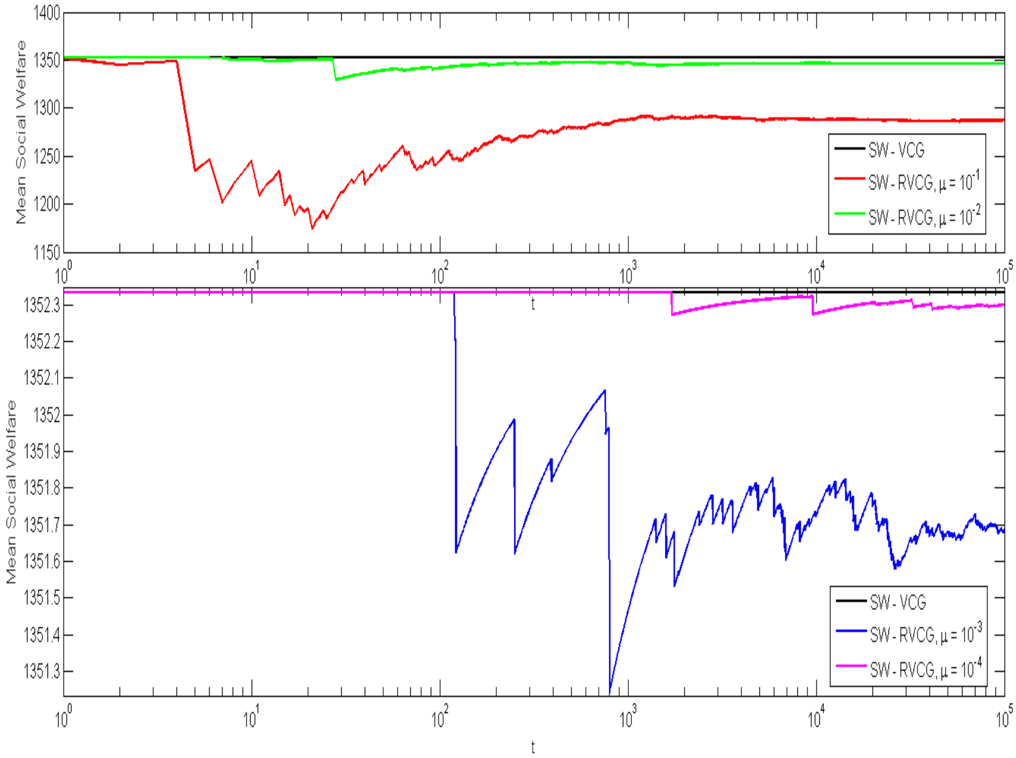

Figure 3.1: Mean Social Welfare for different values of μ.

The graphics in Figure 3.1 have been obtained by using an auction simulation with $n = 10$ agents and $k = 4$ slots.

Truthfulness

A combination of the allocation rule A with the generic transformation gives rise to an ex-post truthful mechanism with respect to clicks.

We experimentally verify it by showing that, for each agent, it is not possible to receive an utility higher than the one obtained by bidding truthfully.

For simplicity, we will report the results of a single run of the algorithm in which only 2 agents participate ($n = 2$) and only 2 slots are available ($k = 2$). The qualities of the ads are $q = \{1, 1\}$ and the real values of the advertisers are $v = \{1,000, 4\}$. The prominence vector is $\Lambda = \{1, 0.25\}$. Fixing the time horizon to $T = 100,000$ and changing the reported values b and the click realization ρ, we distinguished 3 possible scenarios.

1. *Scenario 0*: both agents bid truthfully, $v = b$. We suppose that the ads are clicked independently from the slots in which they are displayed: $\rho_{i,j}(t) = 1 \ \forall i \in N, \ \forall j \in K$.

2. *Scenario 1*: agent 1 reports a value lower than the real one but greater than the value reported by the other agent, $b = (20, 4)$. In this case the pure VCG mechanism will allocate ad a_1 in slot s_1. We still suppose that $\rho_{i,j} = 1 \ \forall i \in N, \ \forall j \in K$.

3. *Scenario 2*: agent 1 reports a value lower than the real one and lower than the value reported by the other agent, $b = (1, 4)$. Note that the pure VCG mechanism will allocate ad a_1 in second position. We still suppose that $\rho_{i,j}(t) = 1 \ \forall i \in N, \ \forall j \in K$.

In Figures 3.2, 3.3, 3.4 and 3.5 we report the functions of the agents' mean utility with respect to the time passed from the beginning of the auction (in ex-post) for the different scenarios listed above and for different values of the resampling probability μ.
It is possible to note that, in all the cases, agents can never achieve an utility higher than the one obtained by bidding truthfully. The above algorithm is truthful with respect to click realization.
It is also interesting to notice that, after 100,000 rounds, the utility obtained by agent 1 when when he is truthful tends to the one obtained by using VCG.

Afterward we compared the agents' utility reached by our mechanism with the values obtained by the pure VCG mechanism in case of truthfulness.
In Figure 3.6 it is interesting to notice how, varying the resampling probability μ, the value of agents' utility changes.

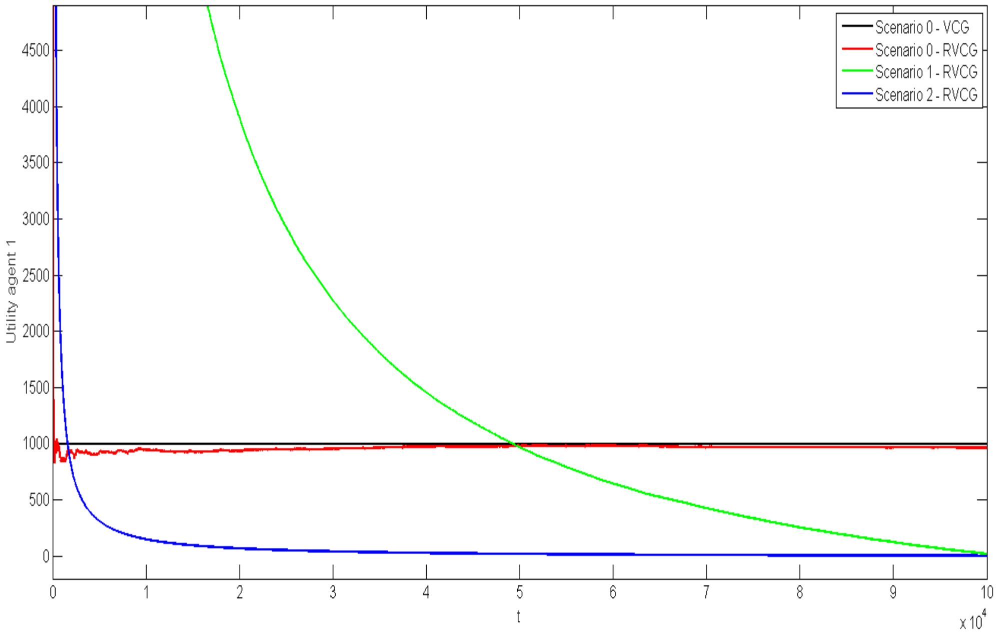

Figure 3.2: Utility of agent 1, $\mu = 10^{-1}$.

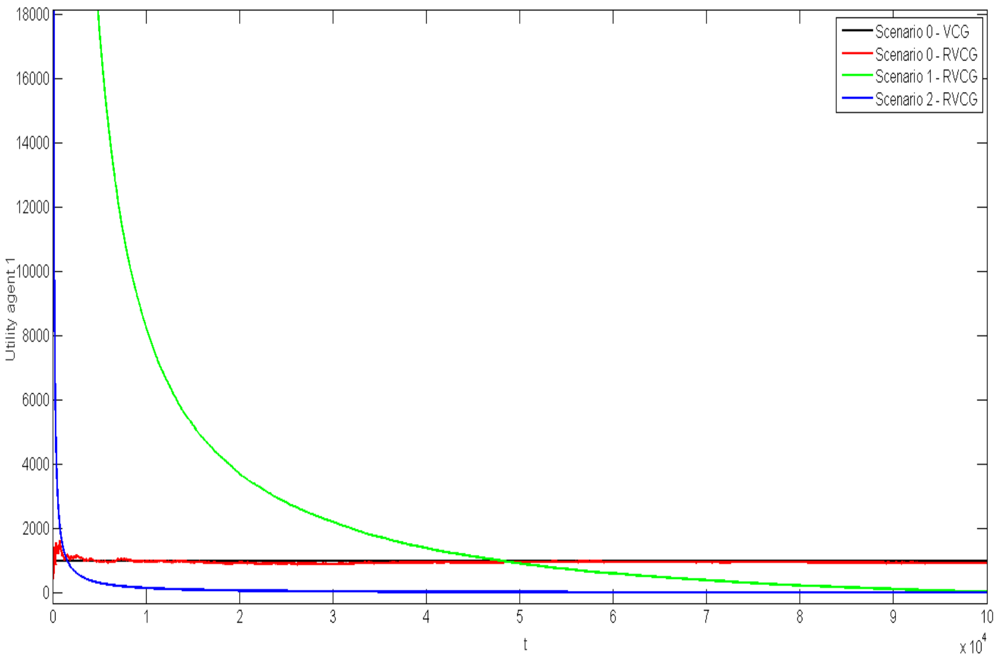

Figure 3.3: Utility of agent 1, $\mu = 10^{-2}$.

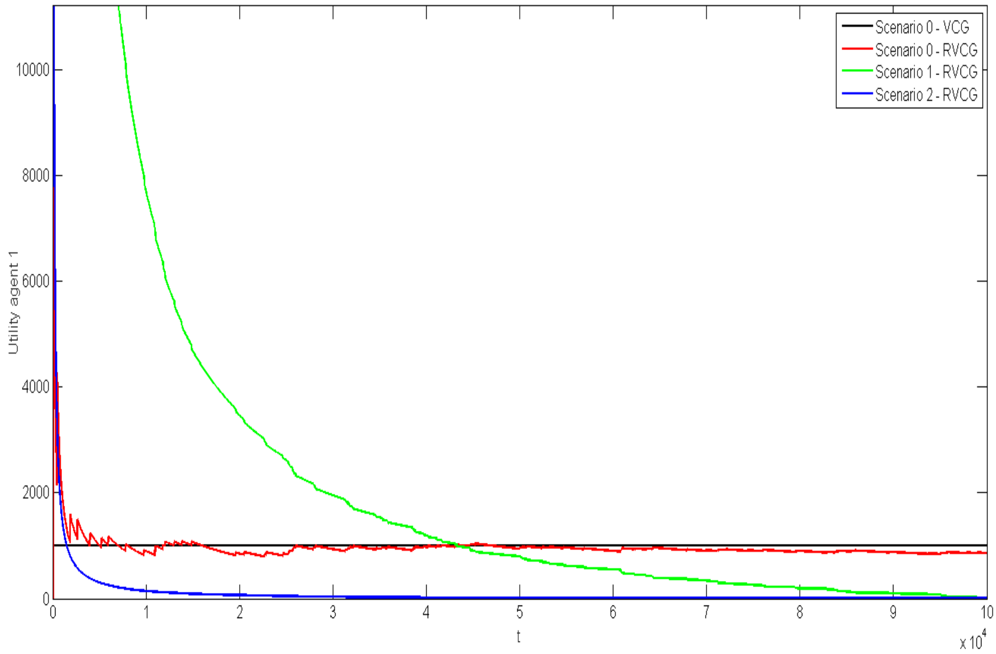

Figure 3.4: Utility of agent 1, $\mu = 10^{-3}$.

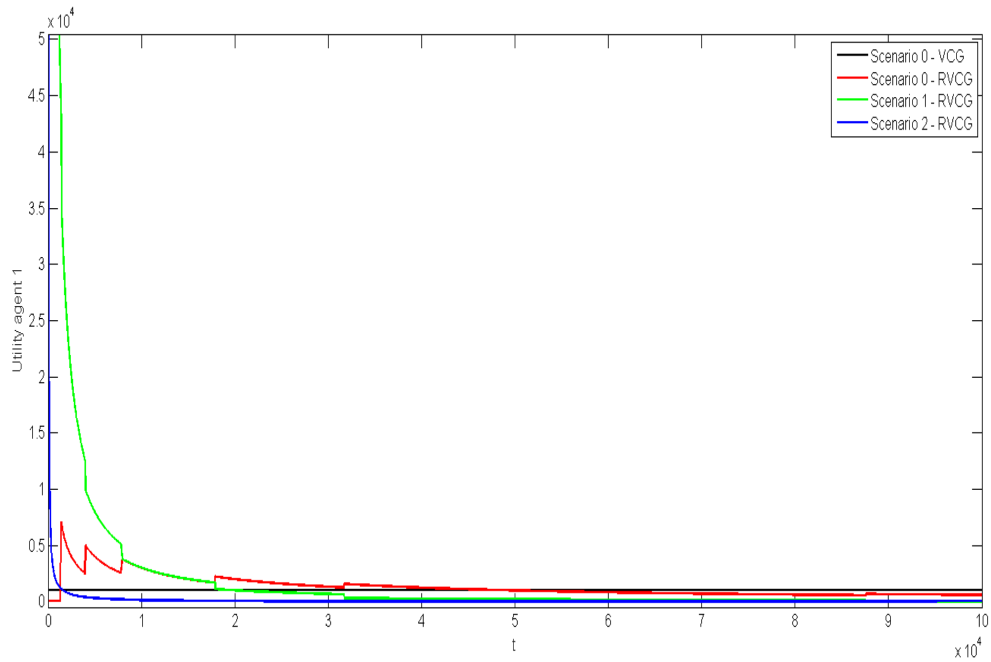

Figure 3.5: Utility of agent 1, $\mu = 10^{-4}$.

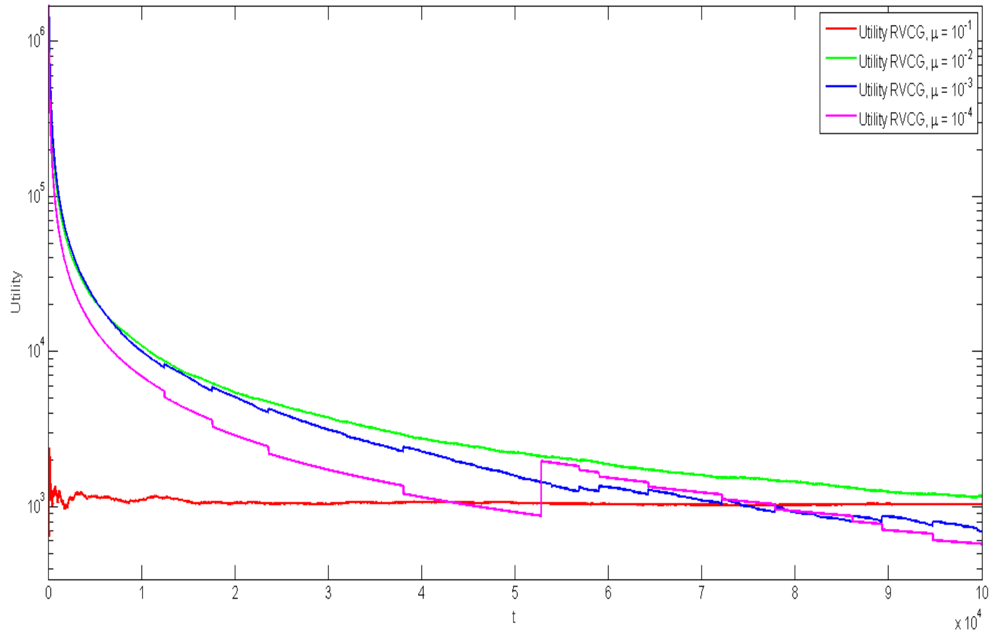

Figure 3.6: Comparison of utilities with different values of μ.

Individual Rationality

It is possible to demonstrate that the mechanism is ex-post IR and agents always receive a non-negative utility by applying the Definition11 to agents' utilities in a direct way.

In order to prove it we take into consideration the two possible cases:

1. ad a_i is not displayed: in this case, since the mechanism adopts a pay-per-click payment scheme, payment p_i is null and consequently also the relative utility u_i of agent i will be null;

2. ad a_i is displayed: the utility u_i of agent i is given by $\rho v_i - \rho b_i + \rho R_i$ when computed in ex-post. It is immediate to notice that $u_i \geq 0$ every time agent i bid truthfully ($v_i = b_i$).

Budget Balance

In order to study this property we are interested in how the revenue of the auctioneer changes during the time.

Recalling that M is the total number of experiment for each experimental setting and T is the number of rounds considered for each experiment m, $R_{m,t}$ represents the reward of the mechanism during the experiment m at round t.
For each round t we computed the mean revenue $\overline{R}_{m,t} = \frac{\sum_{t'=1}^{t} R_{m,t'}}{t}$ with respect to the first t rounds of the experiment m and then we built the table reported in Table 3.1 for each different experimental setting.

Table 3.1: Experiment supporting table.

$R_{1,1}$	$\frac{R_{1,1}+R_{1,2}}{2}$	...	$\frac{\sum_{t'=1}^{T} R_{1,t'}}{T}$
$R_{2,1}$	$\frac{R_{2,1}+R_{2,2}}{2}$	...	$\frac{\sum_{t'=1}^{T} R_{2,t'}}{T}$
...	...	...	...
$R_{M,1}$	$\frac{R_{M,1}+R_{M,2}}{2}$	...	$\frac{\sum_{t'=1}^{T} R_{M,t'}}{T}$

Considering Table 3.1 we defined two measures, necessary to evaluate the experiment:

- $\mathbb{E}(t) = \frac{\sum_{m'=1}^{M} \overline{R}_{m',t}}{M}$, that is the average of the values of column t;
- $\sigma(t) = \sqrt{\frac{\sum_{m'=1}^{M} (\overline{R}_{m',t} - \mathbb{E}(t))^2}{M}}$, that is the standard deviation of the values of column t.

For each experimental setting we plotted the following graphics:

- $\frac{1}{T} \sum_{t'=1}^{T} \mathbb{E}(t')$, for simplicity we will call such a function mean revenue;
- $\frac{1}{T} \sum_{t'=1}^{T} \mathbb{E}(t') + \sigma(t)$, for simplicity we will call such a function standard deviation revenue.

In Figure 3.7 it is possible to note how the results change depending on which value of μ is used. Since the revenue is strictly depending on the value of the rebate applied to the payment and, given that the smaller is the resampling probability μ and the bigger is the rebate, the difference between the pointwise revenue and the average revenue is higher when smaller values of μ are used. However, Figure 3.7 shows that the mean revenue achieved by RVCG mechanism is similar and not lower than the one obtained by using VCG only for $\mu = 10^{-4}$.
Since the difference between the pointwise revenue and the average revenue can not decrease, the only possible evaluations have to concern the average

revenue. It is not possible to obtain stability in the revenue despite the succession of rounds due to the structure of the mechanism itself. Finally, it is necessary to note that, in the optimal case in which $\mu = 10^{-4}$, the time needed to achieve a reasonable mean revenue is really high and greater than 100,000 rounds.

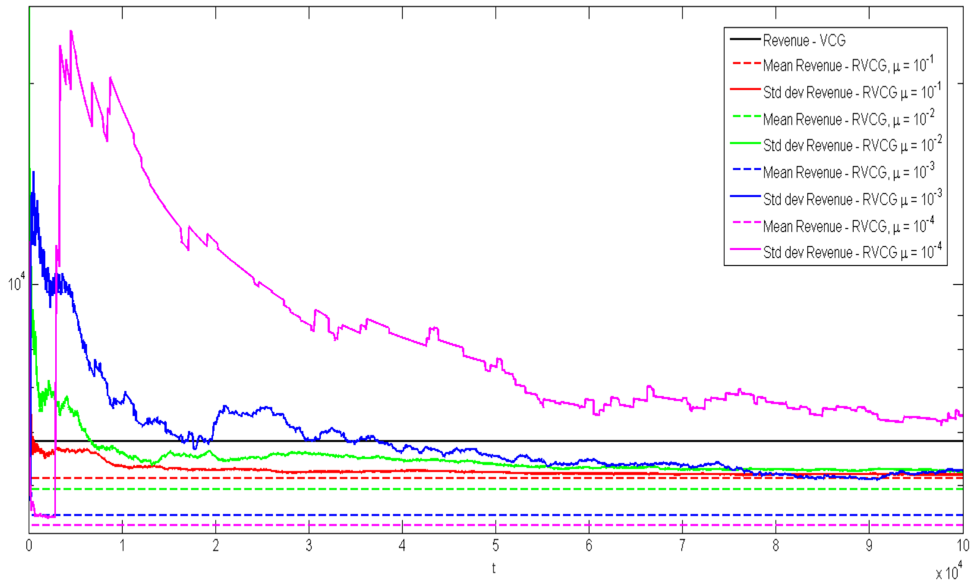

Figure 3.7: Revenue comparison: RVCG and VCG.

3.3.4 Theoretical bounds and experimental results

Inspired by the error bound over the estimator of the single payments described in [14], we built an error upper bound over the estimator of the revenue.

Revenue error upper bound

Statement: With probability at least $1 - \delta$, it holds

$$\left| \left(\frac{1}{t} \sum_{t'=0}^{t} \sum_{i=0}^{n} p_{i,t'} \right) - \left(\sum_{i=0}^{n} \mathbb{E}[p_i] \right) \right| \leq n \frac{v_{max}}{\mu} \sqrt{\frac{1}{2t} \log \frac{2}{\delta}} = \epsilon(t)$$

where $p_{i,t'}$ is the amount charge to agent i at time t' and v_{max} is the maximum value reported by the agents. The function ϵ does not depend only on t but, for simplicity of notation, we make explicit only this dependence.

Proof: applying the Hoeffding's bound [21]

$$Pr\{|\overline{X} - \mu| \geq \epsilon\} \leq \delta$$

in which $\delta = 2e^{\frac{-2\epsilon^2 T^2}{\sum_{i=1}^{T}(b_i - a_i)^2}}$ at round t, we have:

$$Pr\left\{\left|\left(\frac{1}{t}\sum_{t'=0}^{t}\sum_{i=0}^{n} p_{i,t'}\right) - \left(\sum_{i=0}^{n}\mathbb{E}[p_i]\right)\right| \geq \epsilon\right\} \leq \delta$$

with

$$\delta = 2e^{-\frac{2\epsilon^2 t^2}{\sum_{i=1}^{t}\left(\sum_{i=1}^{n} v_i - \sum_{i=1}^{n}(v_i - \frac{v_i}{\mu})\right)^2}} = 2e^{-\frac{2\epsilon^2 t^2}{\sum_{i=1}^{t} n^2\left(\frac{v_{max}}{\mu}\right)^2}} = 2e^{-\frac{2\epsilon^2 t^2}{tn^2\left(\frac{v_{max}}{\mu}\right)^2}} = 2e^{-\frac{2\epsilon^2 t}{n^2\left(\frac{v_{max}}{\mu}\right)^2}}.$$

Thus $2e^{-\frac{2\epsilon^2 t}{n^2\left(\frac{v_{max}}{\mu}\right)^2}} = \delta$, $\log e^{-\frac{2\epsilon^2 t}{n^2\left(\frac{v_{max}}{\mu}\right)^2}} = \log\frac{\delta}{2}$, $-\frac{2\epsilon^2 t}{n^2\left(\frac{v_{max}}{\mu}\right)^2} = \log\frac{\delta}{2}$, $\epsilon^2 = n^2\left(\frac{v_{max}}{\mu}\right)^2\frac{1}{2t}\log\frac{2}{\delta}$, and $\epsilon = n\frac{v_{max}}{\mu}\sqrt{\frac{1}{2t}\log\frac{2}{\delta}}$.

Comparison with theoretical bounds

We finally compare the results obtained through our simulation with the theoretical bounds achieved above. Figures 3.8, 3.9, 3.10 and 3.11 show the revenue upper bound $\epsilon(t)$ for three different values of δ and compare them with the mean revenue for each value of μ. Recalling that δ is the probability that the bound does not hold, it is possible to note that the lower is δ, the less strict the bound is.

Let $\Omega_t = [\ \overline{R}_{m,t} \mid m = 1, ..., M\ ]$, that is the list of values of column t in Table 3.1. In Figures 3.12, 3.13, 3.14 and 3.15 we use boxplots in order to represent the distribution of Ω_t for 10 significant values of t taking into account the total number of rounds T. It is possible to note that the lower the resampling probability μ is, the faster the range of values in which the mean revenues of the M experiments fall becomes small, i.e. the mean revenue of the auction after a sufficient amount of time does not depend heavily on the specific experiment. This range does not strictly become smaller since, if at round t the resampling procedure is activated for some experiment m, the mean revenue will undergo a perturbation. Of course, this perturbation will be less significant as the times goes by. Such a phenomenon is clearly visible when $\mu = 10^{-4}$.

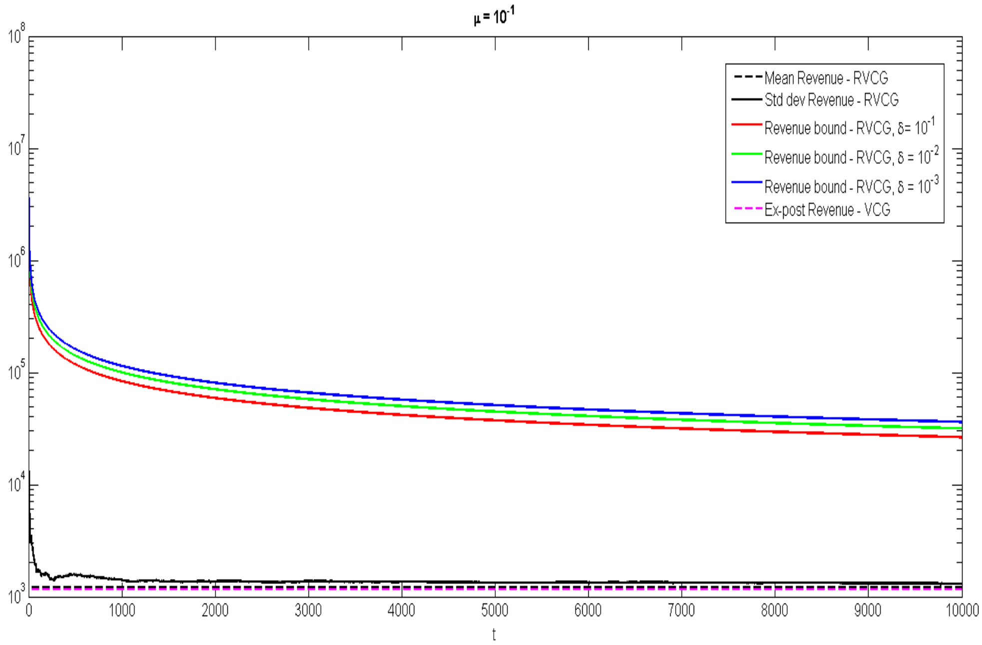

Figure 3.8: Revenue RVCG and theoretical bounds, $\mu = 10^{-1}$.

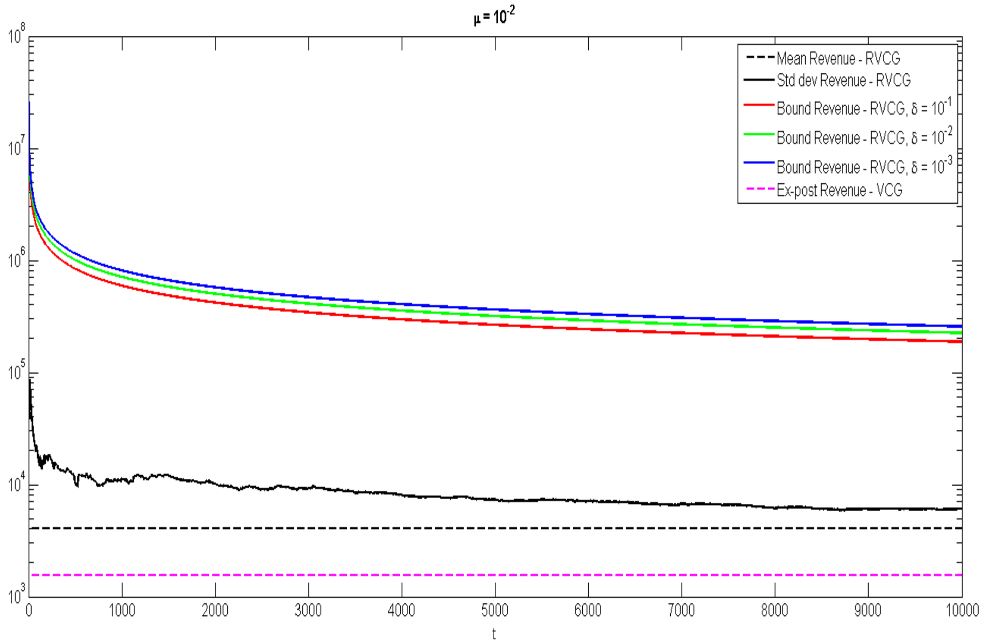

Figure 3.9: Revenue RVCG and theoretical bounds, $\mu = 10^{-2}$.

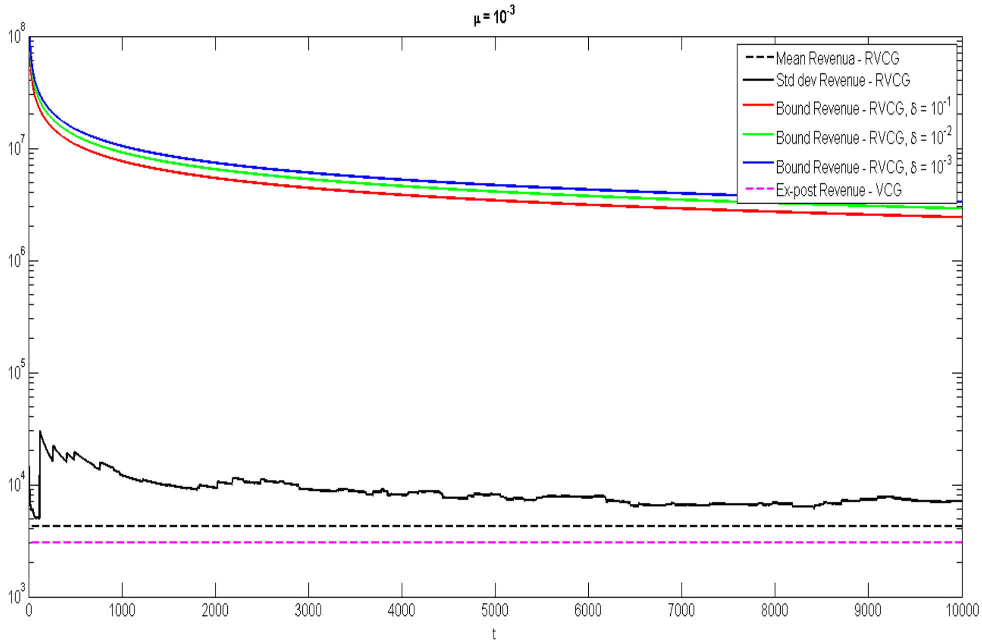

Figure 3.10: Revenue RVCG and theoretical bounds, $\mu = 10^{-3}$.

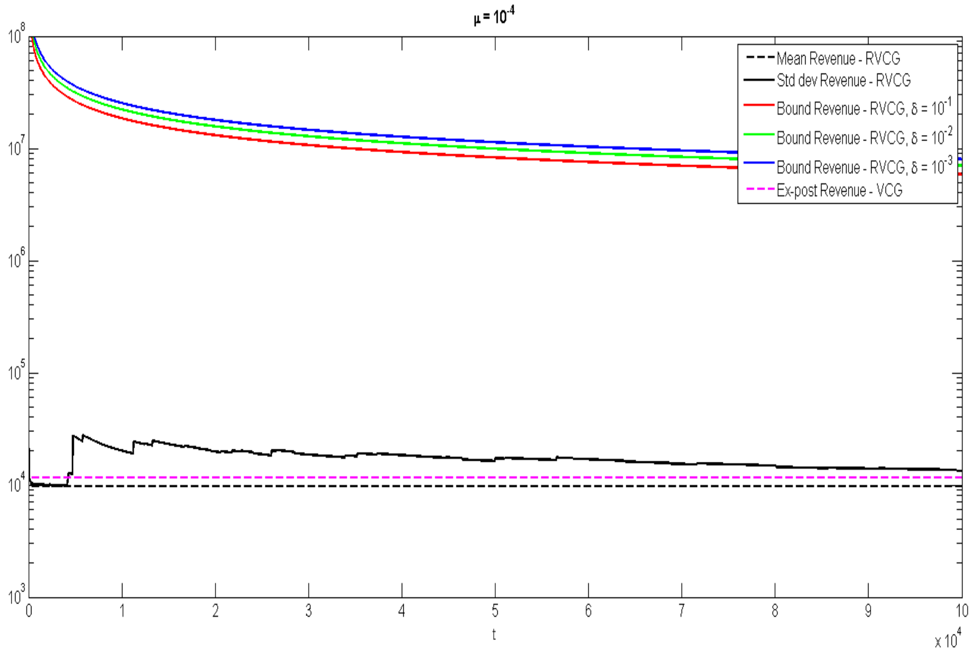

Figure 3.11: Revenue RVCG and theoretical bounds, $\mu = 10^{-4}$.

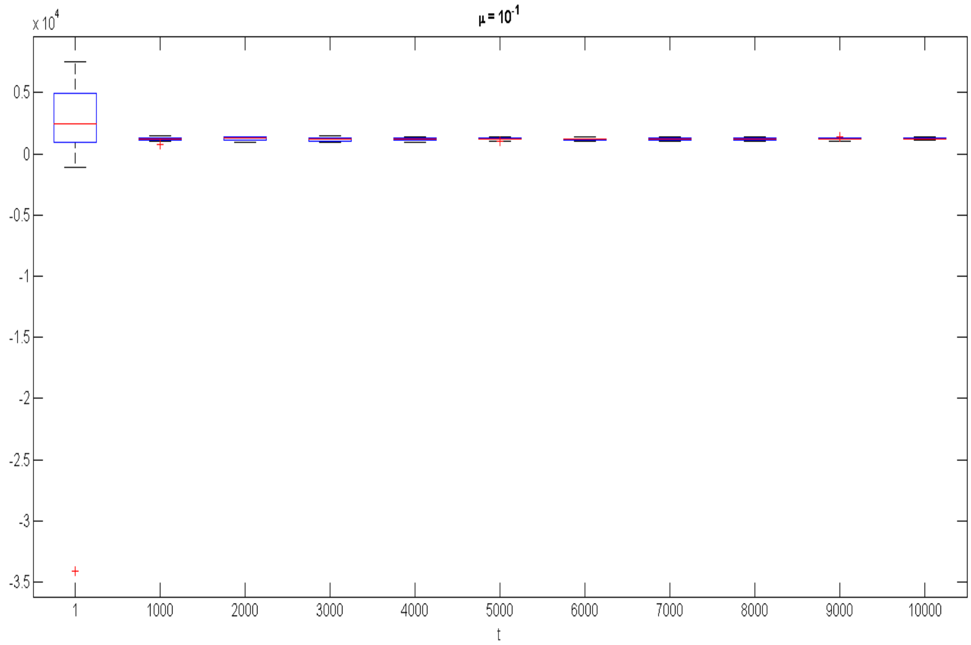

Figure 3.12: Mean revenue distribution - RVCG, $\mu = 10^{-1}$.

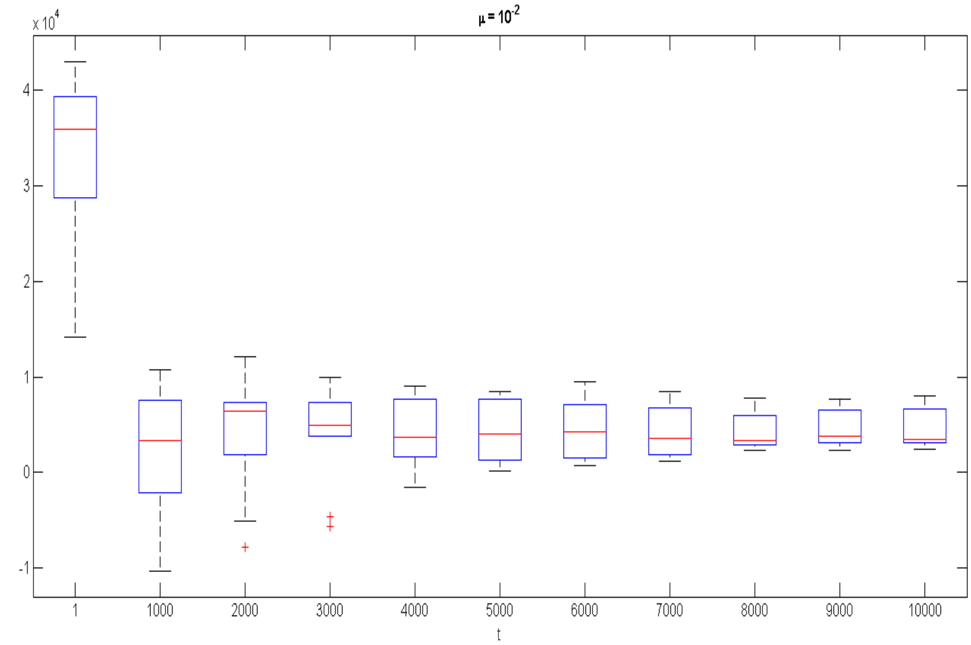

Figure 3.13: Mean revenue distribution - RVCG, $\mu = 10^{-2}$.

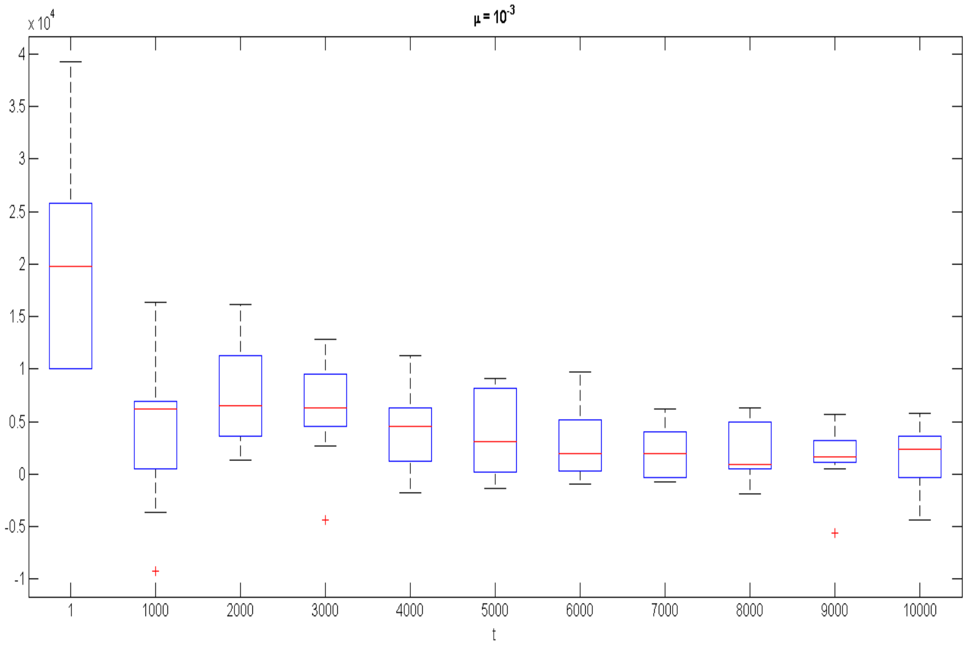

Figure 3.14: Mean revenue distribution - RVCG, $\mu = 10^{-3}$.

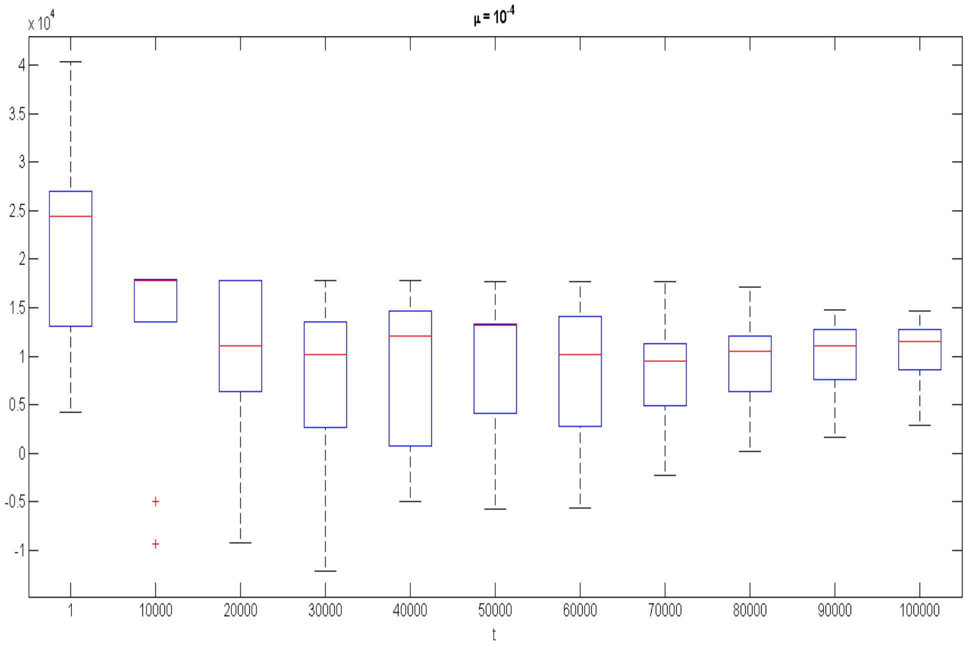

Figure 3.15: Mean revenue distribution - RVCG, $\mu = 10^{-4}$.

Performance

In order to evaluate the performance of our algorithm, we tried to compute the minimum number of rounds T necessary to achieve a reasonable precision.

We tested the algorithm with three different realistic values of the error threshold $\bar{\epsilon}$ ($\epsilon(t) \leq \bar{\epsilon}$) and of the failure probability δ of the bound. We recorded the minimum number of rounds T needed to satisfy the conditions imposed in the possible scenario of 3 agents and 2 slots. We did it for three different values of the resampling probability μ. The results are reported in Tables 3.2, 3.3 and 3.4.

Table 3.2: Minimum T to obtain $\epsilon(t) \leq \bar{\epsilon}$, $\mu = 10^{-1}$.

	$\delta = 10^{-1}$	$\delta = 10^{-2}$	$\delta = 10^{-3}$	$\delta = 10^{-4}$	
$\bar{\epsilon} = 10^{-1}$	0.0000	0.0001	0.0001	0.0001	
$\bar{\epsilon} = 10^{-2}$	0.0041	0.0064	0.0087	0.0110	$\cdot 10^9$
$\bar{\epsilon} = 10^{-3}$	0.4094	0.6397	0.8700	1.1002	

Table 3.3: Minimum T to obtain $\epsilon(t) \leq \bar{\epsilon}$, $\mu = 10^{-2}$.

	$\delta = 10^{-1}$	$\delta = 10^{-2}$	$\delta = 10^{-3}$	$\delta = 10^{-4}$	
$\bar{\epsilon} = 10^{-1}$	0.0000	0.0001	0.0001	0.0001	
$\bar{\epsilon} = 10^{-2}$	0.0041	0.0064	0.0087	0.0110	$\cdot 10^{11}$
$\bar{\epsilon} = 10^{-3}$	0.4094	0.6397	0.8700	1.1002	

Table 3.4: Minimum T to obtain $\epsilon(t) \leq \bar{\epsilon}$, $\mu = 10^{-3}$.

	$\delta = 10^{-1}$	$\delta = 10^{-2}$	$\delta = 10^{-3}$	$\delta = 10^{-4}$	
$\bar{\epsilon} = 10^{-1}$	0.0000	0.0001	0.0001	0.0001	
$\bar{\epsilon} = 10^{-2}$	0.0041	0.0064	0.0087	0.0110	$\cdot 10^{13}$
$\bar{\epsilon} = 10^{-3}$	0.4094	0.6397	0.8700	1.1002	

Chapter 4

Conclusions and future works

In this thesis we focused on the problem of auctions for sponsored links (Sponsored Search Auctions, SSA). It is a problem related to online advertising, in which advertisements are displayed alongside to the results of search engines. These advertisements are chosen taking into account the keywords submitted by the users. These systems generate a revenue of dozens of billions of euros every year. Improving their functioning could significantly increase the profit of the companies.

From a scientific point of view, the problem of building an advertising system for SSA is formalized as a game theory problem, in which the objective is to design an economic mechanism that induces all the participants to truthfully reveal their own evaluations over the benefits received when their ads are clicked. The peculiarity of this scenario is that payments respect the pay-per-click scheme, which means that advertisers pay the mechanism only when their ads are clicked by the users. In such a context, the click probability is an information that is usually considered to be known by the mechanism. Pay-per-click payments and the click probability estimation make the problem of designing an economic mechanism non trivial. In particular, while this problem is well understood in single slot environments, in a more real scenario, i.e. the multi-slot environment, the problem is still open.

In this thesis, mechanisms have been studied in terms of theoretical properties and evaluated through experimental analysis.

From the theoretical point of view, it has been discovered that the concept of ex-post truthfulness, commonly used for the single-slot environment, can not be extended to the multi-slot one. In fact, any simple extension leads a non-randomized mechanism to be non ex-post truthful with respect to clicks,

but only in expectation.
It is possible to design a randomized ex-post truthful mechanism with respect to clicks, but still in expectation with respect to the randomization of the mechanism. However, this solution leads to a very high variance of the payments so much so it makes the mechanism not appealing for a search engine. In particular we experimentally observed that the time needed to obtain a negligible regret for the auctioneer is extremely high (more than $100,000$ rounds).

Our research shows that, from the theoretical point of view, ex-post truthfulness with respect to clicks is a solution concept too hard to fulfill. Moreover, we retain that it is extremely strict also from the practical point of view, since it requires that an advertiser can not obtain a better payoff, by bidding untruthfully, also when he knows if his ad is clicked or not.
In fact, an advertiser cannot know if his ad will receive a click, unless it is the advertiser himself that generates it. However, in this last case the advertiser does not obtain any profit, since there is no real customer who is interested in the subject of the advertisement.

As future works, we retain fundamental to try to formally demonstrate that it is not possible to achieve ex-post truthfulness in multi-slot environment unless a randomized mechanism is adopted.
Such a demonstration, together with the impossibility of using randomized mechanisms in real applications, would show that truthfulness in expectation with respect to click represents the most appropriate solution concept.
Other future works are related to the estimation of the parameters that constitute the click probability, since for these problems the literature shows that it is not possible to avoid the use of randomized economic mechanism.
Furthermore the mechanism, with low probability, generates potentially high negative payments, i.e. it gives money to the advertisers. This gives rise to risk seeking behaviours by fake advertisers that are not interested in having their ads to be displayed but instead they just want to exploit the auction in order to earn money. This behavior should be analyzed in detail in future.

Bibliography

[1] IAB, "Iab internet advertising revenue report, 2014 full year results," 2015.

[2] M. Babaioff, R. D. Kleinberg, and A. Slivkins, "Truthful mechanisms with implicit payment computation," *CoRR*, vol. abs/1004.3630, 2010.

[3] D. Garg, Y. Narahari, and S. Gujar, "Foundations of mechanism design: A tutorial part 1-key concepts and classical results," *Sadhana*, vol. 33, no. 2, pp. 83–130, 2008.

[4] D. Garg, Y. Narahari, and S. Gujar, "Foundations of mechanism design: A tutorial part 2-advanced concepts and results," *Sadhana*, vol. 33, no. 2, pp. 131–174, 2008.

[5] A. Gibbard, "Manipulation of voting schemes: A general result," *Econometrica*, vol. 41, no. 4, pp. pp. 587–601, 1973.

[6] M. A. Satterthwaite, "Strategy-proofness and arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions," *Journal of Economic Theory*, vol. 10, no. 2, pp. 187 – 217, 1975.

[7] J. R. Green and J.-J. Laffont, *Incentives in Public Decision Making.* Amsterdam: North-Holland, 1979.

[8] B. Holmstrm, "Groves' scheme on restricted domains," *Econometrica*, vol. 47, no. 5, pp. pp. 1137–1144, 1979.

[9] A. Archer and E. Tardos, "Truthful mechanisms for one-parameter agents," in *Foundations of Computer Science, 2001. Proceedings. 42nd IEEE Symposium on*, pp. 482–491, Oct 2001.

[10] A. Archer and R. Kleinberg, "Truthful germs are contagious: A local to global characterization of truthfulness," in *Proceedings of the 9th ACM*

Conference on Electronic Commerce, EC '08, (New York, NY, USA), pp. 21–30, ACM, 2008.

[11] G. Aggarwal, J. Feldman, S. Muthukrishnan, and M. Pál, "Sponsored search auctions with markovian users," in *Internet and Network Economics*, pp. 621–628, Springer, 2008.

[12] D. Kempe and M. Mahdian, "A cascade model for externalities in sponsored search," in *Proceedings of the 4th International Workshop on Internet and Network Economics*, WINE '08, (Berlin, Heidelberg), pp. 585–596, Springer-Verlag, 2008.

[13] Y. Narahari, D. Garg, R. Narayanam, and H. Prakash, *Game Theoretic Problems in Network Economics and Mechanism Design Solutions*. Springer Publishing Company, Incorporated, 1 ed., 2009.

[14] N. Gatti and M. Rocco, "Which mechanism for sponsored search auctions with externalities?," in *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems*, AAMAS '13, (Richland, SC), pp. 635–642, International Foundation for Autonomous Agents and Multiagent Systems, 2013.

[15] N. Gatti, A. Lazaric, M. Rocco, and F. Trov, "Truthful learning mechanisms for multi-slot sponsored search auctions with externalities," *Artificial Intelligence*, vol. 227, pp. 93–139, 2015.

[16] T. L. Lai and H. Robbins, "Asymptotically efficient adaptive allocation rules," *Advances in Applied Mathematics*, vol. 6, no. 1, pp. 4–22, 1985.

[17] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Machine Learning*, vol. 47, no. 2-3, pp. 235–256, 2002.

[18] M. Babaioff, Y. Sharma, and A. Slivkins, "Characterizing truthful multi-armed bandit mechanisms," *CoRR*, vol. abs/0812.2291, 2008.

[19] N. R. Devanur and S. M. Kakade, "The price of truthfulness for pay-per-click auctions," in *ACM Conference on Electronic Commerce*, Association for Computing Machinery, Inc., 2009.

[20] D. Mandal and Y. Narahari, "A novel ex-post truthful mechanism for multi-slot sponsored search auctions," in *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems*, AAMAS '14, (Richland, SC), pp. 1555–1556, International Foundation for Autonomous Agents and Multiagent Systems, 2014.

[21] W. Hoeffding, “Probability inequalities for sums of bounded random variables,” *Journal of the American Statistical Association*, vol. 58, no. 301, pp. pp. 13–30, 1963.