Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence

Mechanism Design for Federated Sponsored Search Auctions

Sofia Ceppi and **Nicola Gatti**
Dipartimento di Elettronica e Informazione
Politecnico di Milano, Italy

Enrico H. Gerding
Electronics and Computer Science
University of Southampton, UK

Abstract

Recently there is an increase in smaller, domain–specific search engines that scour the deep web finding information that general–purpose engines are unable to discover. These search engines play a crucial role in the new generation of search paradigms where *federated search engines* (FSEs) integrate search results from heterogeneous sources. In this paper we pose, for the first time, the problem to design a revenue mechanism that ensures profits both to individual search engines and FSEs as a mechanism design problem. To this end, we extend the sponsored search auction models and we discuss possibility and impossibility results on the implementation of an incentive compatible mechanism. Specifically, we develop an execution–contingent VCG (where payments depend on the observed click behavior) that satisfies both individual rationality and weak budget balance in expectation.

Introduction

Sponsored search or "pay–per–click" auctions, where search engines auction off slots to advertisers to display targeted sponsored links alongside the search results, account for a significant part of the engines' revenue. In the first half of 2010, revenue from online advertising totalled $12.1 billion in the U.S. alone, of which search revenue accounted for 47%, dominating display ads, the second–largest revenue source (IAB 2010). In addition to the major search engines Google, Microsoft, and Yahoo!, there is an increase in smaller, domain–specific search engines that scour the *deep web*, which find information (hidden in e.g. databases) that current general–purpose engines are unable to discover (Yu et al. 2008). This has resulted in a new generation of search paradigms where *federated search engines* (FSEs) integrate search results from heterogeneous sources (Braga et al. 2008; Martinenghi and Tagliasacchi 2010). However, in order to use their results for free, search engines require the publishers of the information to also show their ads. For example, Google allows other webpages to display their search results, but only together with the ads (using AdSense). The problem faced by an FSE, however, is that displaying a separate list from each engine is impractical, and so it has to integrate the results by merging a selection of the ads into a coherent list. To do this effectively, it requires detailed information about the ads, such as their quality, from the search engines. Moreover, existing revenue sharing agreements between search engines and publishers, where each receives a fixed share of the profit, are no longer appropriate. Here, we propose a solution to this problem using mechanism design.

There is considerable literature on sponsored search auctions. Many papers focus on equilibrium bidding strategies for the *generalized second price* (GSP) auction, which is the most commonly used format by search engines today (Edelman, Ostrovsky, and Schwarz 2007; Varian 2007). While GSP is shown not to be incentive compatible in (Edelman, Ostrovsky, and Schwarz 2007), (Narahari et al. 2009) has investigated a generalization of the Vickrey–Clarke–Groves (VCG) mechanism for the setting with sponsored search auctions. Other papers focus instead on modeling the user behavior and learning the click probability of the ad (Devanur and Kakade 2009). Among these models the *cascade* model, which assumes that the user scans the links from the top to the bottom in a Markovian way (Aggarwal et al. 2008; Kempe and Mahdian 2008), is the most commonly used. These models introduce negative externalities in the auction whereby the click probability, and therefore an advertiser's profit, depends on which other ads are shown in conjunction.

However, none of these papers consider the problem faced by an FSE. Specifically, in order for the FSE to effectively select which ads to show and in what order, it needs to elicit information about the advertisers' bids as well as (unlike classical sponsored search) the quality score from the search engines (where the quality score reflects the likelihood that an ad is going to be clicked). In this case, existing revenue sharing agreements, whereby the search engines and the publishers receive a fixed share, are no longer appropriate since the search engine can inflate the score and/or decrease the bid to receive a better position at a lower cost. To address this problem, in this paper we design a mechanism that incentivizes the search engines to provide this information truthfully, while satisfying *allocative efficiency* (the optimal set of ads are selected in an optimal order), *budget balance* (the FSE makes no loss), and *individual rationality* (the domain–specific search engines have an incentive to participate). The challenge here is to design such a mechanism for a setting with externalities caused by the user model. Furthermore, we would like these properties to hold *ex–post*, i.e., no matter whether the ads are clicked or not. However, we prove that these properties cannot be satisfied simultane-

Copyright © 2011, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

ously in *ex–post*. We then go on to show that, by only requiring the latter two properties to be satisfied in *expectation* (at the *ex–interim* stage), we can apply an execution–contingent VCG mechanism akin to those in (Porter et al. 2008; Gerding et al. 2010) for service–oriented domains, where the aim is to elicit the quality as well as the costs of a service in settings with interdependent valuations. In more detail, our contributions to the state–of–the–art are as follows.

- We extend existing models of sponsored search auctions to the federated case in a way compliant with the current commercial exploitation of search results, and we show how this can be posed as a mechanism design problem.
- We show, for the first time, that for our setting *without* externalities, there exists no efficient incentive compatible (in *ex–post* Nash) mechanism that satisfies individual rationality in *ex–post*, while there exists a mechanism that satisfies individual rationality in *ex–interim*.
- Furthermore, we show, for our setting *with* externalities, that there exists no efficient, incentive compatible (*ex–post* Nash) *standard* mechanism (i.e., where the payments are not conditional on actual click behavior).
- We then introduce an execution–contingent VCG, where payments depend on the observed click behavior, and show that it satisfies all properties but individual rationality and weak budget balance only in *ex–interim*.

Federated Search Engines Model

In this section we initially describe the general model of sponsored search auctions for federated search engines. We then review and extend the most commonly employed user model. Finally, we pose the problem to design federated sponsored search auctions as a mechanism design problem.

The Ads, Agents, and FSE

Our model consists of a number of domain–specific search engines, henceforth called *agents*, and a single FSE that integrates the sponsored search results of the agents (we do not consider the organic search results here). We let $S = \{s_1, \ldots, s_n\}$ denote the set of agents (search engines). Each agent, in turn, has a number of advertisers that place bids to show their ads in the search engine. We let $A = \{a_1, \ldots, a_m\}$ be the set of all ads, and $A_s \subset A$ the set of ads belonging to agent $s \in S$. Ads are partitioned over the agents, i.e. $\cup_{s \in S} A_s = A$ and $A_s \cap A_{s'} = \emptyset$ for all $s \neq s'$ (each ad belongs to exactly one agent).

Each ad $a \in A_s$ is characterized by its bid, b_a, that has been submitted to agent s by the advertiser, and a quality score, q_a, which is determined by agent s (the quality can be viewed as the probability of being clicked, in the absence of any other ads). Now, given the bids received and the qualities of the ad, the agent computes the advertiser's *pay–per–click* (PPC). In the following, we refer to the PPC as the agent's *value*, v_a, for ad a, since this is the payment she receives when the ad is clicked by a user. This value can be computed, for example, using the generalized second–price auction (GSP). However, in this paper, since we focus on the interaction between search engines and the FSE, we are not concerned with *how* the search engines compute this value, and simply take this as given.

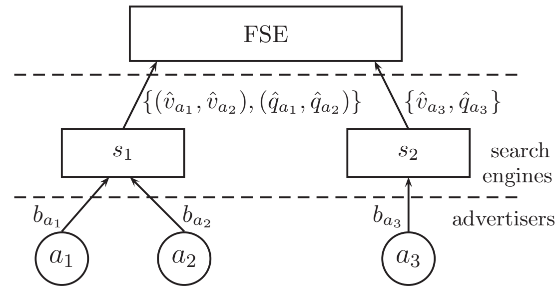

Figure 1: The three levels involved in the FSE model.

We denote by V and Q the set of all values and qualities respectively, and V_s, Q_s the values and qualities of the ads belonging to agent $s \in S$. This information is *privately* known by the agents and we refer to $\theta_s = \{V_s, Q_s\}$ as agent s's *type*. Moreover, $\theta = (\theta_{s_1}, \ldots, \theta_{s_n})$. We will often use $-s$ to denote the set with all agents except s. For example, $A_{-s} = \cup_{s' \in S \setminus \{s\}} A_{s'}$ denotes all ads except those of s.

The agents are asked to report their *types* to the FSE, so that it can allocate its own slots to the ads supplied by the agents, and calculates appropriate renumeration for the agents. We assume that agents are self–interested, and so they may misreport if it's in their best interest to do so. To this end, we let $\hat{\theta}_s$ denote agent s's reported type, and similarly denote by $\hat{v}_a, \hat{q}_a$ the reported value and quality of an ad $a \in A_s$. Moreover, $\hat{\theta} = (\hat{\theta}_{s_1}, \ldots, \hat{\theta}_{s_n})$. An example with 2 agents and 3 ads is depicted in Figure 1. Note that the model consists of three levels: the advertisers, the agents and the FSE. However, in this work we only consider the latter 2 levels and treat the bids and how the search engines calculate the PPC (value) as given (we note that the advertisers are motivated to bid strategically, but this does not affect the interaction between the search engines and the FSE).

The FSE's User Model

The aim of the FSE is to calculate an outcome which is *efficient* (as defined below). To this end, the FSE requires a user model which determines the probability of an ad being clicked, which depends on its quality score, but also, depending on the user model, its position and (in the case of externalities) the quality of other ads which are shown in conjunction. For example, in the well–known cascade model (Aggarwal et al. 2008; Kempe and Mahdian 2008), the top positions are most likely to be clicked, and the probability of ads being clicked also depends on the quality of ads in higher positions. In this paper we do not consider a specific user model, but rather introduce a general model so that our results apply to a wide range of user models, e.g. including models with multiple advertising slates, where each slate has different properties (Kempe and Mahdian 2008).

More formally, we use the function $\alpha_x(\mathbf{x}, Q) \in [0, 1]$ to denote the *click–through rate* (CTR) of ad $x \in A$, i.e. probability of a being clicked, given the qualities, Q, and the *allocation*, $\mathbf{x}$, where $\mathbf{x} = \langle x_1, x_2, \ldots \rangle$ is a vector denoting which ad $x_i \in A$ is shown in position i, and $|\mathbf{x}|$ is the total number of slots. Note that $\alpha_x = 0$ for $x \notin \mathbf{x}$.

We say that a user model has *externalities* if the CTR of an ad a depends on quality score of ads other than a. For example, in the case of the cascade model, the function becomes

$\alpha_{x_i}(\mathbf{x}, Q) = q_{x_i} \prod_{j=1}^{i-1} c_{x_j}$, where c_{x_j} is the so–called *continuation probability* of the ad $x_j \in A$ in position j. Now, when all c_{x_j} are constant, the model reduces to classical sponsored search without externalities. However, if c_{x_j} is a function of q_{x_j}, then the model displays externalities.

The Mechanism Design Problem

We are now ready to define the mechanism design problem, which involves finding appropriate allocations and payments. Importantly, we consider *execution–contingent* mechanisms, where the payments are contingent on the realization of events, in this case which ads are actually clicked. We choose this type of mechanism because, if the user model exhibits externalities, the expected utilities of the agents are *interdependent* (the expected utility of an agent not only depends on the allocation, but also on the types, in this case the ad qualities, of other agents). For such settings, it has been shown that, under certain conditions, there exist no *standard* mechanisms which are both efficient and incentive compatible (Jehiel and Moldovanu 2001) (we discuss these properties in detail below). However, this impossibility result is avoided by using contingent payments in cases such as this one, where the interdependency is caused by uncertainty of events occurring (Mezzetti 2004).

In more detail, a mechanism is formally characterized by a tuple $M(S, A, X, \Theta, f, \langle p_s \rangle_{s \in S})$, where S and A are defined as before, X is the set of possible outcomes, Θ is the set of the possible agents' types, $f : \Theta^n \to X$ is the decision function mapping agents *reported* types to outcomes, and $p_s : \Theta^n \times \Omega \to \mathbb{R}$ specifies the payment from agent $s \in S$ to the FSE, given the (reported) types and the realization of the event. In this case, an event represents all ads that are actually clicked. Specifically, $\Omega = \mathcal{P}(A)$ specifies the set of possible events, and $\omega \in \Omega, \omega \subseteq A$ its realization.

We assume that all agents are expected utility maximizers. The utility of an agent s with true type θ_s when reporting $\hat{\theta}_s$ to the mechanism, and given realization ω is: $u_s(\hat{\theta}, \theta|\omega) = \sum_{a \in \omega \cap A_s} v_a - p_s(\hat{\theta}|\omega)$. This utility function implicitly assumes a revenue sharing agreement whereby the advertisers' payments go to the search engines, and the FSE receives payments from the search engines as compensation (note that, as we will see, payments can also be negative, which means that the FSE has to pay out). Thus the utility of the FSE is: $u_F(\hat{\theta}|\omega) = \sum_{s \in S} p_s(\hat{\theta}|\omega)$.

Then, an agent's *expected* utility (w.r.t. Ω) is given by:

$$E[u_s(\hat{\theta}, \theta)] = \sum_{a \in A_s} \alpha_a(f(\hat{\theta}), Q) \cdot v_a - E[p_s(\hat{\theta})]$$

Given this, we would like to design f and p such that the following properties are satisfied:

1. **Incentive Compatibility** (IC) This assures that reporting truthfully maximizes an agent's expected utility. We consider two variations: dominant strategy IC, where reporting truthfully maximizes expected utility regardless of the reports of others, and the slightly weaker version *ex–post* IC (Bergemann and Morris 2008), where each agent has an incentive to be truthful if other agents are truthful (note that this is still stronger than Bayes–Nash IC where this decision depends on the types of other agents).
2. **Individual Rationality** (IR) This means that each agent $s \in S$ receives non–negative utility from taking part (given truthful reporting). We consider *ex–interim* IR, where *expected utility*, $E[u_s(\langle \theta_s, \hat{\theta}_{-s} \rangle, \theta)]$ is always positive, as well as the stronger notion of *ex–post* IR, where the utility is positive for every realization of events.
3. **Weak Budget Balance** (WBB) This assures that the FSE does not run a deficit, i.e. that the sum of payments received by the agents is positive. We distinguish between *ex–interim WBB*, where $\forall \hat{\theta} : E[u_F(\hat{\theta})] \geq 0$, and the stronger *ex–post* WBB, where $\forall \hat{\theta}, \omega : u_F(\hat{\theta}|\omega) \geq 0$
4. **Allocative Efficiency** (AE) Finally, as is common in mechanism design, we would like the allocation function f to be *efficient*, that is, to maximize *social welfare* (the sum of the utility of all agents as well as the FSE). To this end, the allocation is given by:

$$f(\hat{\theta}) = \arg\max_{\mathbf{x} \in X} \sum_{i=1}^{|\mathbf{x}|} \alpha_{x_i}(\mathbf{x}, \hat{Q}) \cdot \hat{v}_{x_i} \qquad (1)$$

Ideally, we would like the mechanism to satisfy the strongest properties mentioned above. However, as we will see, this is not always possible, which necessitates using slightly weaker conditions. In the next sections, we derive possibility and impossibility results for our setting both with and without externalities.

Mechanism Design without Externalities

We first focus on a setting with no externalities where the CTR of an ad a depends only on quality q_a and whether it is in $\mathbf{x}$. In particular, we assume $\alpha_a(\mathbf{x}, q_a) = q_a$ whenever $a \in \mathbf{x}$, and $\alpha_a(\mathbf{x}, q_a) = 0$ otherwise. Differently from classical sponsored auctions (Narahari et al. 2009), we will show that in the federated case no efficient incentive compatible (in *ex–post* implementation) mechanism satisfies individual rationality in *ex–post*. Furthermore, we will show that the VCG mechanism satisfies *ex–interim* individual rationality. The main reason underlying this impossibility result is that, in addition to the value, v_a, the payment of an agent must depend on its CTR, and this value is private information. The practical implications of this result are that the pay–per–click scheme, commonly used in the sponsored search auctions, cannot be used when search engines are federated. In the following, we discuss our arguments.

Theorem 1. *For the FSE model without externalities there exists no mechanism that is AE,* ex–post *IC,* ex–post *IR, and* ex–post *WBB.*

Proof. The proof is by counterexample. Consider a situation with two agents (search engines) s_1 and s_2, each one with a single ad, and only one available slot (i.e., $|\mathbf{x}| = 1$). Let $A_{s_1} = \{a_1\}$ and $A_{s_2} = \{a_2\}$ and call α_{a_1}, v_{a_1}, and α_{a_2}, v_{a_2} the report related to the ads. For the sake of simplicity, the proof is split into two parts.

Part I We introduce an additional constraint (relaxed in the second part of the proof): the payment of the agent whose ad is not displayed is zero. Call s^* the agent whose ad a^* is displayed on the unique available slot and $-s^*$ the other agent (whose ad is denoted by $-a^*$). By AE, *ex–post*

IC, *ex–post* IR, and *ex–post* WBB, we have that the payments of agent s^* must satisfy the following constraints:

$$\begin{cases} p_{s^*}(\hat{\theta}) = 0 & \text{if ad } a^* \text{ is not clicked} \\ p_{s^*}(\hat{\theta}) \leq q_{-a^*} v_{-a^*} & \text{if ad } a^* \text{ is clicked} \end{cases}.$$

To see this, note that if the link is not clicked, the valuation of agent s^* is zero and then, by *ex–post* IR, the payment must be equal to or smaller than zero. By *ex–post* WBB (given that we assume $p_{-s^*}(\hat{\theta}) = 0$ for every $\hat{\theta}$), $p_{s^*}(\hat{\theta})$ must be equal to or larger than zero, and then $p_{s^*}(\hat{\theta}) = 0$. Consider instead the case in which ad a^* is clicked. By AE, $q_{a^*} v_{a^*} \geq q_{-a^*} v_{-a^*}$ and, by *ex–post* IC, $p_{s^*}(\hat{\theta})$ must depend only on the report of agent $-s^*$. To satisfy *ex–post* IR, the largest value that $p_{s^*}(\hat{\theta})$ can assume is $q_{-a^*} v_{-a^*}$, otherwise when, e.g., $q_{a^*} v_{a^*} = q_{-a^*} v_{-a^*}$ and $q_{a^*} = 1$ agent s^*'s utility may be negative and therefore *ex–post* IR is not satisfied.

To prove that *ex–post* IC does not hold, we consider the specific case in which $v_{a_1} \geq q_{a_2} v_{a_2}$ and $q_{a_1} v_{a_1} < q_{a_2} v_{a_2}$. When both agents report truthfully their report, we have $E[u_{s_1}(\theta)] = 0$ because s_1's ad will not be displayed and the payment is zero, i.e., $p_{s_1}(\theta) = 0$. When agent s_1 misreports her type, reporting $\hat{q}_{a_1} = 1$, we have $E[u_{s_1}(\hat{\theta})] \geq q_{a_1}(v_{a_1} - q_{a_2} v_{a_2})$ (due to the above constraint on $p_{s^*}(\hat{\theta})$) that is, by definition, strictly positive. As a result, agent $s1$ strictly prefers to misreport her type and therefore there cannot be an incentive compatible mechanism given AE, *ex–post* IR, and *ex–post* WBB.

Part II. We relax the constraint $p_{-s^*}(\hat{\theta}) = 0$ and we show that it always holds under the hypotheses of the theorem. We observe that the constraint $q_{a_2} v_{a_2} \geq p_{s_1}(\hat{\theta})$ holds when the link of agent s_1 is not clicked. By *ex–post* WBB, $p_{s_1}(\hat{\theta}) \geq -p_{s_2}(\hat{\theta})$ and, by *ex–post* IR, $-p_{s_2}(\hat{\theta}) \geq 0$. Since, by *ex–post* IC, $p_{s_2}(\hat{\theta})$ cannot depend on agent s_2's report and $q_{a_2} v_{a_2} \geq p_{s_1}(\hat{\theta}) \geq -p_{s_2}(\hat{\theta})$ must hold for all the agents' report (even when $q_{a_2} = 0$), we have that $p_{s_2}(\hat{\theta}) = 0$. That is, the hypotheses of the theorem makes $p_{-s^*}(\hat{\theta}) = 0$ always true thereby completing the proof. □

Furthermore we show that the impossibility result holds even when WBB is in *ex–interim*.

Theorem 2. *For the FSE model without externalities there is no mechanism that is AE,* ex–post *IC,* ex–post *IR, and* ex–interim *WBB.*

Proof. The proof is a straightforward generalization of the second part of the proof of Theorem 1. First, we observe that the constraint $p_{s^*} \leq q_{-a^*} v_{-a^*}$ holds also when WBB is in *ex–interim*. Then, by *ex–interim* WBB, we have $q_{a^*} q_{-a^*} v_{-a^*} \geq q_{a^*} p_{s^*}(\hat{\theta}) \geq -E[p_{-s^*}(\hat{\theta})]$. Since, by *ex–post* IC, $E[p_{-s^*}(\hat{\theta})]$ cannot depend on agent $-s^*$'s report and $q_{a^*} q_{-a^*} v_{-a^*} \geq q_{a^*} p_{s^*}(\hat{\theta}) \geq -E[p_{-s^*}(\hat{\theta})]$ must hold for all the agents' reports (even when $q_{-a^*} = 0$), we have that $E[p_{-s^*}(\hat{\theta})] = 0$. It is easy to see that, if $E[p_{-s^*}(\hat{\theta})] = 0$, then the first part of the proof of Theorem 1 can be applied also when WBB is in *ex–interim* and so the theorem is proven. □

We are now ready to show that the critical property is *ex–post* individual rationality. Indeed, it is possible to design a mechanism when individual rationality is in *ex–interim*.

Theorem 3. *For the FSE model without externalities, VCG with Clarke pivot rule is AE, dominant–strategy IC,* ex–interim *IR and* ex–post *WBB.*

Proof. The proof is trivial by considering the type θ_s of each search engine s as a single–dimensional fictitious type $\theta_s = q_s v_s$. □

Mechanism Design with Externalities

We now focus on the general setting in which there are externalities and therefore $\alpha_a = \alpha_a(\mathbf{x}, \hat{Q})$ not only depends on q_a, but also on (with abuse of notation) Q_{-a} and $\mathbf{x}$. It is easy to see that the results discussed in the previous section hold also in the presence of externalities. Therefore, with externalities no incentive compatible efficient mechanism can be implemented when individual rationality is in *ex–post*.

In addition, however, as mentioned earlier, with externalities the expected utility of an agent depends on the private information of other agents. These settings are known as settings with *interdependent types*, and a number of works show impossibility results on implementing incentive compatible standard mechanisms in presence of interdependencies (Jehiel and Moldovanu 2001). Here, we prove the impossibility to design an efficient *standard* mechanism (where payments depend only on reported types) for FSE with externalities implemented in *ex–post*, even when IR and WBB are in *ex–interim*, and the cascade model is used.

Theorem 4. *For the FSE model with externalities there is no standard mechanism that is AE,* ex–post *IC,* ex–interim *WBB, and* ex–interim *IR.*

Proof. We prove the theorem by counterexample. Consider a situation with two agents (search engines) s_1 and s_2, each one with a single ad, and two available slots (i.e., $|\mathbf{x}| = 2$). Let $A_{s_1} = \{a_1\}$ and $A_{a_2} = \{a_2\}$. Assume the user model to be the cascade model, in which a's continuation probability c_a is a function of q_a, i.e., $c_a = c_a(q_a)$. Let $\mathbf{x}_1 = \langle a_1, a_2 \rangle$ and $\mathbf{x}_2 = \langle a_2, a_1 \rangle$, s^* be the agent whose ad a^* is allocated in the first slot and $-s^*$ be the agent whose ad $-s^*$ is allocated in the second slot. CTRs are:

$$\alpha_{a_1}(\mathbf{x}_1, \hat{Q}) = q_{a_1} \qquad \alpha_{a_1}(\mathbf{x}_2, \hat{Q}) = c_{a_2} q_{a_1}$$
$$\alpha_{a_2}(\mathbf{x}_1, \hat{Q}) = c_{a_1} q_{a_2} \qquad \alpha_{a_2}(\mathbf{x}_2, \hat{Q}) = q_{a_2}$$

By AE, *ex–post* IC, *ex–interim* WBB, and *ex–interim* IR we have that the payment of agent s^* is $p_{s^*}(\hat{\theta}) = 0$ for every $\hat{\theta}$. To see this, note that, by *ex–interim* IR, $p_{s^*}(\hat{\theta}) \leq q_{a^*} v_{a^*}$ (given that $E[u_{s^*}(\hat{\theta})] = q_{a^*} v_{a^*} - p_{s^*}(\hat{\theta})$) and by AE, $q_{a^*} v_{a^*} \geq \frac{q_{-a^*} v_{-a^*}(1 - c_{-a^*})}{1 - c_{-a^*}}$ (given that $q_{a^*} v_{a^*} + c_{a^*} q_{-a^*} v_{-a^*} \geq q_{-a^*} v_{-a^*} + c_{-a^*} q_{a^*} v_{a^*}$). To satisfy *ex–interim* IR, the largest value that $p_{s^*}(\hat{\theta})$ can assume is zero, otherwise when, e.g., $c_{a^*} = 1$ the $-s^*$'s expected utility may be negative. By *ex–interim* WBB (given that $p_{s^*}(\hat{\theta}) = 0$ for every $\hat{\theta}$) $p_{-s^*}(\hat{\theta}) \geq 0$, and by *ex–post* IC $p_{-s^*}(\hat{\theta}) \leq 0$, therefore $p_{-s^*} = 0$.

To prove that the *ex–post* IC does not hold, consider the specific case in which $\mathbf{x}^* = \mathbf{x}_1$. In the case both agents report truthfully their type, s_2's expected utility is $E[u_{s_2}(\theta)] = c_{a_1} q_{s_2} v_{s_2}$ (given that her payment is zero). In the case s_2 misreports her report such that the resulting efficient allocation is $\mathbf{x}_2$, her expected utility is $E[u_2(\hat{\theta})] = q_{s_2} v_{s_2}$ (given that her payment is zero). Since $q_{s_2} v_{s_2} \geq c_{a_1} q_{s_2} v_{s_2}$, s_2 strictly prefers to misreport her type and therefore there cannot be an incentive compatible mechanism given AE, *ex–interim* WBB, and *ex–interim* IR. □

To overcome the impossibility, we propose a VCG execution–contingent mechanism in which the payments depend on the actual realization of the events. More precisely, the allocation function is efficient and is defined as in Equation (1) and the payments are defined as:

$$p_s(\hat{\theta}|\omega) = \sum_{a \in A \setminus A_s} \alpha_a(f(\hat{\theta}_{-s}), \hat{Q}) \cdot \hat{v}_a - \sum_{a \in A \setminus A_s, a \in \omega} \hat{v}_a$$

Notice that the payments depend on the ads actually clicked, captured in the formula by $\omega \in \Omega$.

Theorem 5. *For the FSE model with externalities the above execution–contingent VCG is AE,* ex–post *IC,* ex–interim *IR, and* ex–interim *WBB.*

Proof. Assume that all the agents but s truthfully report their report values, i.e., $V_{-s} = \hat{V}_{-s}$ and $Q_{-s} = \hat{Q}_{-s}$. Then, if agent s report $\hat{V}_s$ and $\hat{Q}_s$, its expected utility is:

$$E[u_s(\hat{\theta}, \theta)] = \sum_{a \in A} \alpha_a(f(\hat{\theta}), (\hat{Q}_s, Q_{-s})) \cdot v_a - \sum_{a \in A / A_s} \alpha_a(f(\hat{\theta}_{-s}), (\hat{Q}_s, Q_{-s}))\hat{v}_a \geq 0$$

First, we note that there is nothing that agent s can do to manipulate the second term of the expected utility, given the reports of other agents. The reason is that this part of the expected utility of s is independent of her reported values. Secondly, we note that the first term of expected utility is computed after the observation of the realization of the event. While the selection of $\mathbf{x}^*$ depends on $\hat{V}$ and $\hat{Q}$, the actual outcome upon the realization depends on V and Q. As a result, note that, by definition of $f(\hat{\theta})$:

$$\sum_{a \in A} \alpha_a(f(\hat{\theta}), (Q_s, Q_{-s})) \cdot v_a \geq \sum_{a \in A} \alpha_a(f(\hat{\theta}), (\hat{Q}_s, Q_{-s})) \cdot v_a$$

Thus, if all agents $-s$ truthfully report their valuations and qualities, agent s is also best off revealing its information truthfully, since this will result in a mechanism selecting the allocation which maximize the social welfare in expectation. This leads to the expected utility maximization of agent s.

The execution–contingent mechanism is *ex–interim* IR, since $\sum_{a \in A} \alpha_a(f(\hat{\theta}), Q) \cdot v_a \geq \sum_{a \in A} \alpha_a(f(\hat{\theta}_{-s}), Q_{-s}) \cdot v_a$, implying that $E[u_s(\hat{\theta}, \theta)] \geq 0$. It can be easily seen that it is also *ex–interim* WBB. □

We now show that the proposed mechanism cannot guarantee *ex–post* IR and *ex–post* WBB.

Example. *Consider a situation with two agents (search engines) s_1 and s_2, and two available slots (i.e., $|\mathbf{x}| = 2$). Let $A = \{a_1, a_2, a_3\}$, and $A_{s_1} = \{a_1, a_2\}$ and $A_{s_2} = \{a_3\}$. We compute the agents' payments for each possible realization according to the above mechanism:*

$$p_{s_1}(\hat{\theta}|\emptyset) = \alpha_{a_3}(f(\hat{\theta}_{s_2}), \hat{Q})\hat{v}_{a_3}$$
$$p_{s_1}(\hat{\theta}|\{a_3\}) = \alpha_{a_3}(f(\hat{\theta}_{s_2}), \hat{Q})\hat{v}_{a_3} - \hat{v}_{a_3}$$
$$p_{s_2}(\hat{\theta}|\emptyset) = \alpha_{a_1}(f(\hat{\theta}_{s_1}), \hat{Q})\hat{v}_{a_1} + \alpha_{a_2}(f(\hat{\theta}_{s_1}), \hat{Q})\hat{v}_{a_2}$$
$$p_{s_2}(\hat{\theta}|\{a_1\}) = \alpha_{a_1}(f(\hat{\theta}_{s_1}), \hat{Q})\hat{v}_{a_1} + \alpha_{a_2}(f(\hat{\theta}_{s_1}), \hat{Q})\hat{v}_{a_2} - \hat{v}_{a_1}$$
$$p_{s_2}(\hat{\theta}|\{a_2\}) = \alpha_{a_1}(f(\hat{\theta}_{s_1}), \hat{Q})\hat{v}_{a_1} + \alpha_{a_2}(f(\hat{\theta}_{s_1}), \hat{Q})\hat{v}_{a_2} - \hat{v}_{a_2}$$
$$p_{s_2}(\hat{\theta}|\{a_1, a_2\}) = \alpha_{a_1}(f(\hat{\theta}_{s_1}), \hat{Q})\hat{v}_{a_1} + \alpha_{a_2}(f(\hat{\theta}_{s_1}), \hat{Q})\hat{v}_{a_2} - \sum_{i=1}^{2} \hat{v}_{a_i}$$

Consider the joint realization in which no displayed ads is clicked, i.e., $\omega = \emptyset$. The utility of s_1 is:

$$u_{s_1}(\hat{\theta}, \theta|\omega) = -\alpha_{a_3}(f(\hat{\theta}_{s_2}), \hat{Q})\hat{v}_{a_3}$$

It is actually negative and therefore ex–post *IR is not satisfied. Consider the joint realization in which all the displayed ads are clicked, i.e., $\omega = \{a_1, a_2, a_3\}$. The utility of FSE is:*

$$E[u_F(\hat{\theta})] = \sum_{s \in S} \sum_{a \in S_s} \hat{v}_a \cdot (1 - \alpha_a(f(\hat{\theta}_{-s}), \hat{Q}))$$

Both agents' payments are strictly negative (except when CTRs are zero) and therefore ex–post *WBB is not satisfied.*

The above example shows that, in addition to *ex–post* IR (as shown by Theorem 2), also *ex–post* WBB may be not satisfied. This may not be desirable because the FSE could, in specific instances, run a deficit. However, we show that this limitation is not specific to our mechanism but due to the nature of the problem. To prove this, we show in the next theorem that no mechanisms with WBB in ex-post can be implemented.

Theorem 6. *For the FSE model with externalities there is no execution–contingent mechanism that is AE,* ex–post *IC,* ex–post *WBB, and* ex–interim *IR.*

Proof. It can be easily observed that only efficient execution–contingent mechanisms that compute *ex–post* payments (w.r.t. realizations) such that the payments in expectation are those prescribed by Groves mechanisms can be incentive compatible. We recall that Groves payments are defined as $E[p_s(\hat{\theta})] = h_s(\hat{\theta}_{-s}) - \sum_{a \in A \setminus A_s} \alpha_a(f(\hat{\theta}_s), \hat{Q}) \cdot \hat{v}_a$ where h_s is an arbitrary function, while execution–contingent payments are defined as $E[p_s(\hat{\theta})] = \sum_{\omega \in \Omega} p_s(\hat{\theta}|\omega) \cdot \rho(\omega)$ where $\rho(\omega)$ is the probability associated to the realization ω. Thus, in order to prove the theorem, we just need to show that there is not any efficient execution–contingent mechanism whose payments are WBB in *ex–post* and are equivalent in expectation to those of some Groves mechanism. We show it by a counterexample.

Consider a situation with two search engines $S = \{s_1, s_2\}$ and two available slots. Let $A_1 = \{a_1\}$, $A_2 = \{a_2\}$. Consider the specific case of the cascade model. The report are: $q_{a_1}, c_{a_1}, v_{a_1}, q_{a_2}, c_{a_2}, v_{a_2}$ where c_{a_i} is a function of allocation $\mathbf{x}$ and qualities Q.

To prove that *ex–post* WBB does not hold, consider the case $q_{a_1} v_{a_1} + c_{a_1} q_{a_2} v_{a_2} \geq q_{a_2} v_{a_2} + c_{a_2} q_{a_1} v_{a_1}$.

We have the following constraints over the expected payment of s_1: (Groves payments) $E[p_{s_1}(\hat{\theta})] = h_{s_1}(\hat{\theta}_{s_2}) - c_{a_1}q_{a_2}v_{a_2}$, (execution–contingent payments) $E[p_{s_1}(\hat{\theta})] = p_{s_1}(\hat{\theta}_{s_2}|\{a_2\}) \cdot c_{a_1}q_{a_2} + p_{s_1}(\theta|\emptyset) \cdot (1 - c_{a_1}q_{a_2})$, (*ex–interim* IR) $E[p_{s_1}(\hat{\theta})] \leq q_{a_1}v_{a_1}$. (It can be observed that payments of agent i cannot depend on the actual click on the agent i's ads; the proof is trivial.) We have the following constraints over the expected payment of s_2: (Groves payments) $E[p_{s_2}(\hat{\theta})] = h_{s_2}(\hat{\theta}_{s_1}) - q_{a_1}v_{a_1}$, (execution–contingent payments) $E[p_{s_2}(\hat{\theta})] = p_{s_2}(\hat{\theta}_{s_1}|\{a_1\}) \cdot q_{a_1} + p_{s_2}(\theta|\emptyset) \cdot (1 - q_{a_1})$, (*ex–interim* IR) $E[p_{s_2}(\hat{\theta})] \leq c_{a_1}q_{a_2}v_{a_2}$.

To see that at least for one realization the payments of both the agents are negative, we derive $p_{s_1}(\hat{\theta}|\emptyset) = h_{s_1}(\hat{\theta}_2)$ and $p_{s_1}(\hat{\theta}|\{a_2\}) = h_{s_1}(\hat{\theta}_2) - v_{a_2}$ from the above constraints. Analogously, $p_{s_2}(\hat{\theta}|\emptyset) = h_{s_2}(\hat{\theta}_1)$ and $p_{s_2}(\hat{\theta}|\{a_1\}) = h_{s_2}(\hat{\theta}_1) - v_{a_1}$. Notice that, in order for both payments of s_1 (s_2) to be non–negative, we need $h_{s_1}(\hat{\theta}_2) \geq v_{a_2}$ ($h_{s_2}(\hat{\theta}_1) \geq v_{a_1}$). By applying the *ex–interim* IR constraint, we can observe that (e.g., with $q_{a_1} = 0.4$, $v_{a_1} = 1$, $q_{a_2} = 0.5$, $v_{a_2} = 1$) payments $p_{s_1}(\hat{\theta}_{s_2}|\{a_2\})$ and $p_{s_2}(\hat{\theta}_{s_1}|\{a_1\})$ are strictly negative. Therefore, there is at least a joint realization such that for all the possible h_{s_1} and h_{s_2} the FSE's payment is strictly negative in *ex–post*. □

To summarize, the results show that only an execution–contingent mechanisms can be designed for the FSE model with externalities (Theorems 4 and 5), and there exists no efficient and *ex–post* incentive compatible mechanism that can be *ex–post* IR (Theorem 2) or *ex–post* WBB (Theorem 6). The implication of the above results is that the execution–contingent VCG mechanism we propose is the strongest mechanism for the FSE model with interdependent types.

Conclusions and Future Works

It is commonly believed in the web technology field that the new generation of general–purpose search engines will be based on the integration of multiple sources, each being a domain–specific search engine. These *federated* search engines allow one to scour the deep web, finding information that current general–purpose engines are unable to discover. It is also expected that the main revenue stream for these federated search engines will be derived from advertising.

To this end, for the first time, in this paper we posed the problem of developing a revenue mechanism for federated search engines using mechanism design. In doing so, we extended the state–of–the–art models for sponsored search auctions to the federated case. Our work is essentially theoretical and provides possibility and impossibility results on the implementation of stable mechanisms. Specifically, we showed that efficient incentive compatible mechanisms cannot be individually rational in *ex–post*, but only in *ex–interim* and therefore the well–known pay–per–click scheme cannot be used in the federated case. We furthermore showed that, when user models are employed where ads incur externalities (such as the well–known cascade model), only execution–contingent mechanisms can be implemented and they cannot be weak budget balanced in *ex–post*, but only in *ex–interim*.

In our current work, we mainly focused on the interaction between the domain–specific search engines and the federated search engine. In future work, we would like to expand this model by also considering the strategic behavior of advertisers. Furthermore, we plan to consider computational aspects related to our mechanisms, such as approximation schemes for maximizing the social welfare. Finally, we consider the use of automated mechanism design techniques to get around the *ex–post* impossibility results by designing mechanisms for specific cases.

Acknowledgments

This research is part of the "Search Computing" (SeCo) project, funded by ERC, under the 2008 Call for "IDEAS Advanced Grants".

References

Aggarwal, G.; Feldman, J.; Muthukrishnan, S.; and Pál, M. 2008. Sponsored search auctions with markovian users. In *WINE*, 621–628.

Bergemann, D., and Morris, S. 2008. Ex post implementation. *Games and Economic Behavior* 63(2):527–566.

Braga, D.; Ceri, S.; Daniel, F.; and Martinenghi, D. 2008. Mashing up search services. *IEEE Internet Computing* 12(5):16–23.

Devanur, N. R., and Kakade, S. M. 2009. The price of truthfulness for pay-per-click auctions. In *ACM EC*, 99–106.

Edelman, B.; Ostrovsky, M.; and Schwarz, M. 2007. Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *American Economic Review* 97(1):242–259.

Gerding, E. H.; Stein, S.; Larson, K.; Rogers, A.; and Jennings, N. R. 2010. Scalable mechanism design for the procurement of services with uncertain durations. In *AAMAS*, 649–656.

IAB. 2010. IAB internet advertising revenue report. 2010 first half-year results.

Jehiel, P., and Moldovanu, B. 2001. Efficient design with interdependent valuations. *Econometrica* 69(5):1237–1259.

Kempe, D., and Mahdian, M. 2008. A cascade model for externalities in sponsored search. In *WINE*, 585–596.

Martinenghi, D., and Tagliasacchi, M. 2010. Proximity rank join. *The Proc. of the VLDB Endowment* 3(1):352–363.

Mezzetti, C. 2004. Mechanism design with interdependent valuations: Efficiency. *Econometrica* 72(5):1617–1626.

Narahari, Y.; Garg, D.; Narayanam, R.; and Prakash, H. 2009. *Game Theoretic Problems in Network Economics and Mechanism Design Solutions*. Springer.

Porter, R.; Ronen, A.; Shoham, Y.; and Tennenholtz, M. 2008. Fault tolerant mechanism design. *Artificial Intelligence* 172(15):1783–1799.

Varian, H. R. 2007. Position auctions. *International Journal of Industrial Organization* 25(6):1163–1178.

Yu, J.; Benatallah, B.; Casati, F.; and Daniel, F. 2008. Understanding mashup development. *IEEE Internet Computing* 12:44–52.