On the Benefits of Keyword Spreading in Sponsored Search Auctions: An Experimental Analysis

Michele Budinich\(^1\), Bruno Codenotti\(^2\), Filippo Geraci\(^2\), and Marco Pellegrini\(^2\)

\(^1\) Institute for Advanced Studies, Lucca, Italy
m.budinich@imtlucca.it
\(^2\) Istituto di Informatica e Telematica, CNR – Consiglio Nazionale delle Ricerche, 56100 Pisa, Italy
\{bruno.codenotti,filippo.geraci,marco.pellegrini\}@iit.cnr.it

Abstract. Sellers of goods or services wishing to participate in sponsored search auctions (SSA) must define a pool of keywords that are matched on-line to the queries submitted by the users to a search engine. Sellers must also define the value of their bid to the search engine for showing their advertisements in case of a query-keyword match. In order to optimize its revenue a seller might decide to substitute a keyword with a high cost, thus likely to be the object of intense competition, with sets of related keywords that collectively have lower cost while capturing an equivalent volume of user clicks. This technique is called keyword spreading and has recently attracted the attention of several researchers in the area of sponsored search auctions. In this paper we describe an experimental benchmark that, through large scale realistic simulations, allows us to pin-point the potential benefits/drawbacks of keyword spreading for the players using this technique, for those not using it, and for the search engine itself. Experimental results reveal that keyword spreading is generally convenient (or non-damaging) to all parties involved.

1 Introduction

A very large fraction of consumers use search engines to find information on the web about goods and services before deciding whether to purchase them in the online markets. Search engines take advantage of their key position on the Web to sell advertising space to economic players on search result pages. Indeed, over the last few years, sponsored search advertising has become the dominant source of profits for search engines. Typically sponsored search results appear in two separate parts of the page above and to the right of the results returned by a search engine. Sponsored search results include a title, a short text, and a link referring to a Website. Advertising space comes in the form of slots, which are sold by auctions. When a user submits a given keyword in a query to a search engine, an auction is run among all the advertisers submitting bids for that keyword. The advertisers who wish to display their ads against the search for a keyword participate in the auction by specifying their valuation and a daily
budget to the search engine. The search engine could use various mechanisms for determining winners and payments, the most popular mechanism being the generalized second price (GSP) auction.

Although GSP looks similar to the classical Vickrey-Clarke-Groves (VCG) mechanism, its properties are very different, i.e., truth-telling is not an equilibrium in GSP. Over the last years, several papers of computational flavor have appeared, touching in different ways this paradigm of online advertising, see, e.g., [5,6,12,20,21]. From the viewpoint of a search engine, the adword problem consists of assigning a sequence of search keywords to a set of competing bidders, each with a daily spending limit, with the goal of maximizing the revenue generated by these keyword sales. This problem generalizes on-line matching, and this connection has been exploited in [23]. A central problem in adword markets from the point of view of a seller of goods and services is the generation of keywords. Advertisers typically prefer to bid for keywords that have high search volumes; however they may be very expensive, so that it might be reasonable to bid instead for several related and low volume, inexpensive terms that generate roughly the same amount of traffic altogether. Some preliminary work exploring this idea has been done in [1], where however, the emphasis is on the algorithmic aspects of keyword generation, not on the global market phenomena as in the present work.

In this paper we describe a large scale simulator for analyzing the effect of using synonyms for keyword spreading in sponsored search auctions (SSA), and collect a number of evidences about the effects of this strategy. Our simulations involve up to 2M agents bidding for words from a pool of 36K words and 3M queries per experiment (more details in Sections 2 and 3). Our experiments point to the following conclusions:

- using synonyms increases the revenues for all players in the market (Figs. 6a, 7a, 7b); in particular the early adopter agents benefit the most (Figs. 7a, 7b)
- using a VCG payment scheme decreases the agents’ benefits with respect to using GSP while not much changes for the search engine (data omitted for lack of space, see [8])
- as the fraction of agents using synonyms increases, the search engine revenues are not significantly affected (Fig. 6b) as well as the costs for the agents not using them (Fig. 8a) while the agents using synonyms have decreasing gains (Fig. 8b)
- budget depletion strategies are shown to rarely be beneficial for the agents, while always increasing the revenue for the search engine, even in presence of keyword spreading (data omitted for lack of space, see [8]).

A problem related to keyword spreading is that of keyword selection, where the economic players try to select at fixed rounds the subset of keywords that maximize revenues while trying to learn basic parameters (such as keyword click-through rates) during the repeated bidding processes. Note that here the viewpoint is that of a single player and that the market, as seen by the seller, is