

THE VALUE OF SPONSORED ADS FOR XML DOCUMENTS

Michel de Rougemont
University Paris II
LRI, CNRS
mdr@lri.fr

Abstract

We study how sponsored ads which are classically inserted in Web pages could also be inserted in XML documents such as Ebooks. Advertisers set prices for words and DTDs, and a Vickrey mechanism would select an ad for some price.

The approximate distance between a document and any DTD, which is easy to compute, allows to associate a value to a new document in such a market, as a combination of the price and the distance.

We describe a simple auction for one ad and its generalization to several ads.

Keywords: XML, distances, auctions, value of information

Introduction

There is a large source of new documents in normalized XML formats, which can be freely distributed : short novels, technical documents, catalogs. This is a new market for sponsored adds.

We consider an economic setting, where sponsored ads can be inserted into XML documents, in particular Ebooks viewed as specific XML documents, and where advertisers target DTDs (Data Type Definition, i.e. XML Schemas) which generalize the classical keywords and capture the structure of the documents. As for sponsored ads such as *GoogleAdWords* where the advertisers define an ad and set private prices for sets of keywords, the advertisers define a similar ad but set private prices for several DTDs. We study a mechanism, inspired by Vickrey auctions which will allocate an ad to any Ebook for some price.

We use an Edit distance between documents and DTDs, which is easy to approximate [1], even on Streaming documents and associate to any Ebook its relative distance $0 < d < 1$ to all the known DTDs of the central mechanism. If the Ebook is at distance d_i of DTD_{*i*}, valued for a price p_i , we define its private value as $(1 - d_i) \cdot p_i$. We then set a Vickrey auction [2] on these values: we allocate the ad with the highest value for a price of the second best value. We will show that this mechanism keeps the fundamental Vickrey properties.

This approach can be generalized to several ads in a given Ebook, and to several positions inside the document. In the first case, we allocate k ads instead of one ad, and in the second case we consider the distance as a dynamic value over the streaming document and record the positions which minimize the distances to several DTDs.

References

- [1] Fischer, E. , Magniez, F., de Rougemont, M. "Approximate Satisfiability and Equivalence". *IEEE Logic in Computer Science* , 2006: 421-430
- [2] Vickrey,W. "Counterspeculation, Auctions and Competitive Sealed Tenders", *Journal of Finance*, 16, 1961, pp 8-37.