A Double Auction for Querying the Web of Data

Moor, Dmitry ; Grubenmann, Tobias ; Seuken, Sven ; Bernstein, Abraham

A Double Auction for Querying the Web of Data

[Extended Abstract][*]

Dmitry Moor, Tobias Grubenmann, Sven Seuken, Abraham Bernstein
Department of Informatics, University of Zurich
{dmitry.moor, grubenmann, seuken, bernstein}@ifi.uzh.ch

Categories and Subject Descriptors

J.4 [**Social and Behavioral Sciences**]: Economics

General Terms

Algorithms, Design, Economics

Keywords

Web of Data; Mechanism Design; Double Auction.

The *Web of Data (WoD)* is to computers what the traditional Web is to humans. The goal is to expose data in a semantic format such that machines can easily find the information they are looking for. If implemented properly, a *semantic search* would lead to the desired results much faster than using a traditional search engine. However, a big practical challenge is the adoption of WoD technologies (Antoniou and van Harmelen, 2004). In particular, the majority of data is not available in a suitable WoD format. The primary reason for this is a lack of incentives for data providers: the WoD is usually queried via algorithms rather than browsed by people and thus, *advertising*, the main source of income for search engines on the traditional web, does not work.

It is of course possible to directly charge users to search through high-quality data. For example, *Bloomberg*, *Lexis-Nexis*, and *Thomson Reuters* charge customers high fees for accessing their data. However, none of these companies provide the means to join multiple data sets, thereby forgoing the complementarities the WoD would enable. To address this shortcoming, we suggest using a market which allows buyers to benefit from complementary data sources while providing incentives to sellers to adopt WoD technologies.

Concretely, we propose to use a double auction, designed to provide market participants with good incentives to reveal their values and costs such that we can make (approximately) efficient allocations. Interestingly, in our domain, it is non-trivial to find a payment rule that satisfies the *participation* constraint (all participants have non-negative utility) and the *budget-balance* constraint (payments made by the buyers are equal to the payments received by the sellers).

[*] Working paper available at www.ifi.uzh.ch/ce/files/QM.pdf

The main complicating feature of our domain is that the sellers are selling their *data* instead of just *computational resources*. While the sellers can provide some statistics about their data, it is impossible to know a priori what data is available to answer a specific query, because this would require running the whole query and incurring the full costs. Thus, there is *uncertainty about which goods exist in the market*, and this may sometimes lead to inefficient allocations. If not handled properly, these inefficient allocations can make the payment rule infeasible because the *participation* constraint and the *budget balance* constraint may be violated. For example, the *Threshold* rule (Parkes et al., 2001), a classic double auction which always satisfies *participation* and *budget-balance* in a domain without uncertainty, is not guaranteed to work in our domain.

Our Contributions

1. We propose the "Query Market," a double auction mechanism for querying the WoD.

2. We show that the *Threshold* rule does not work in our domain due to the uncertainty about the sellers' data.

3. We introduce three *payment correction* rules to design payment rules that always satisfy the *participation* constraint despite the uncertainty in our domain.

4. We evaluate the efficiency of the three payment rules via a computational Bayes-Nash equilibrium analysis (Lubin et al., 2015) for multiple market scenarios.

We compare the performance of our new payment rules in a domain *with* uncertainty to the performance of the *Threshold* rule in a domain *without* uncertainty. We find that our rules provide similar incentives and achieve similar efficiency as *Threshold*, despite the additional challenge of having to deal with the uncertainty about the sellers' data. Furthermore, we show that even our rules cannot guarantee *budget balance* due to the occasional inefficient allocations, but we demonstrate that the average amount of deficit is small.

References

G. Antoniou and F. van Harmelen. *A Semantic Web Primer*. Cambridge, MA, London, England: The MIT Press, 2004.

B. Lubin, B. Bünz, and S. Seuken. New core-selecting payment rules with better fairness and incentive properties. Working Paper, August 2015.

D. C. Parkes, J. Kalagnanam, and M. Eso. Achieving budget-balance with vickrey-based payment schemes in exchanges. In *Proceedings of the 17th International Joint Conference on Artificial Intelligence (IJCAI)*, Seattle, WA, August 2001.

.