Fuzzy Systems and Soft Computing
ISSN : 1819-4362

OPTIMIZING ONLINE PRICING IN PERSONAL DATA MARKETS WITH RESERVE PRICE CONSTRAINTS

[#1]THODETI AKANKSHA,
[#2]PINAMALLA NITHIN,
[#3]D.SHANTHI KUMAR, ***Assistant Professor,***
Department of Computer Science and Engineering,
SREE CHAITANYA INSTITUTE OF TECHNOLOGICAL SCIENCES, KARIMNAGAR, TS.

ABSTRACT: Modern society's unprecedented need for personal data is driving the expansion of data marketplaces. Data consumers are able to execute bespoke searches on datasets obtained from data proprietors through data brokers through these markets. In this study, we look at how a data broker could maximize its revenue by providing acceptable price for consecutive queries. Consequently, we supply a contextual dynamic pricing system that includes a ceiling price. Online optimization of linear and non-linear market value models with uncertainty can be accomplished via this approach. It also shows signs of being ellipsoid-shaped. The proposed pricing technique, when used under circumstances of little uncertainty, results in a cumulative regret that, in the worst-case scenario, is logarithmic and related to the number of searches. We have broadened our methods to cover more similar use cases, such as online advertising and hotel services. Using the MovieLens 20M dataset, the Avazu mobile ad click dataset, and Airbnb listings in key U.S. cities, we thoroughly test all three use cases. Our pricing mechanism has low memory overhead and latency, so it can support online applications. The results of the evaluation and analysis show that: (1) our pricing mechanism has minimal practical regret; and (2) a posted price mechanism can avoid a cold-start problem with the help of a reserve price, which reduces cumulative regret.

KEYWORDS:Online Pricing, Reserve Price Constraint, Personal Data Markets

1.INTRODUCTION

Companies collect a lot of different kinds of information about people in order to keep an eye on what they're doing. This includes medical records, travel routes, reviews, energy use, social networking data, and website cookies.

Because most data owners don't want to share their information for security, privacy, or business reasons, there are a lot of "data islands." While data is being separated, potential data users such as banks, hospitals, colleges, and businesses are not able to get private information. As the need for data brokers grows, more and more businesses have sprung up to make it easier for data owners and customers to share personal information. Some of the best known data providers in the business are Factual, DataSift, Datacoup, CitizenMe, and CoverUS. If data brokers want to encourage people to share data proactively, they must first make sure that data owners are fairly compensated for the privacy breaches that happen when they comply with data consumers' requests. So the data provider doesn't lose money by underpricing or overpricing, they have to charge online data users a fair price for searching sequentially through the datasets they've bought. In the academic world, the environment for moving data around is often called the "data market".

From the point of view of a data merchant who works in digital data markets, this piece looks at the best ways to trade personal information in order to make money.Here is a summary of the three main design problems. The difficulty of the optimization goal function is one of the main and most difficult problems to solve.

A data broker's main goal when it comes to data markets is to make as much money as possible over time. The amount of money this business makes is the difference between how much people pay to look for data and how much data owners get paid for privacy. We are going to look at a single data trade round right now. When it comes to privacy breaches and general compensation for privacy, the reserve price for a certain inquiry is pretty much set in stone. The data broker must set a price that takes into account both the reserve price and the market value of the question in order to make the most money.

Using past trade records and current queries, the data broker can only make educated guesses about the real market value. It's clear that wrong predictions will cause people to feel different amounts of regret. The question won't be able to be sold if the reserve price is higher than its market value, even if the data broker knows how much it's worth. In other words, the item should be worth more on the market than its stated price. Because of this, there is no sorrow. (2) If the reserve price is the same as or less than the market value, a small underestimation doesn't cause much regret. On the other hand, if the thing doesn't sell, a small overestimation causes a lot of stress.

So, stopping cumulative regret is the same thing as reaching the original goal of making the most money. This means reducing the difference between how much money the data broker makes when they use market price information and when they don't. It's hard to get better after the first round because the regret function is already very uneven and broken up.

Trying to guess how much customized data questions from clients are worth is another problem. To lower the chance of being unable to decide how much to charge for internet searches, data brokers must first fully understand what those searches are worth in the market. Every person who uses data is a buyer in the market for personal data. The data broker is not the only one who can figure out what the question or offering is. Every investigation usually uses a different way to look at the data and adds the right amount of noise to the real data .

Because of this, the data broker doesn't have much control over the many requests that come in from different data users. Traditional dynamic pricing methods, which try to target a wide range of similar or different goods, would not work in this case because of this unique feature.

Also, past research into the creation of the data market didn't look into whether data consumers agreed or disagreed with the price that was given. Instead, they focused on a single query or looked at the link between determinacy and multiple searches . So, even though they were going on at the same time as ours, these two projects failed to get the market prices for the queries.

Finding your way around the electronic reservation-price pricing scheme is the hardest part. The data broker is the only one who can figure out how much a query is worth on the market by looking at both new and old queries. This means that the price of following up is about the same as the price of taking an online course.

In addition to the usual two options of exploitation and research, our pricing problem has three clear aspects: A surprising lack of reaction follows the exchange of a single question. Standard methods for online learning aren't thought to work for these reasons: For reasons (1) the reserve price effectively caps the posted price at a value higher than the estimated market value; (2) the data broker can only tell if the posted price is higher than the market value; and (3) the relationship between the reserve price and the market value is still unknown.

No one has looked into how this lower limit affects the whole learning process, and there needs to be a good way to post prices for the online mode. That is, the only way for a data broker to get lower latency is to constantly review its understanding of the market value model along with the choice of every stated price.

2.SYSTEM ANALYSIS

EXISTINGSYSTEM

To begin, the problem of regular, unrefined data trade comes up. Most study on databases has been on the topic of arbitrage freeness in query pricing over relational databases (e.g., Koutris et al. Lin and Kifer . Due to arbitrage, people who want to buy data can get a query for less than the stated price by combining a number of different, cheaper queries. If data brokers want to make money, they need to get rid of arbitrage possibilities. Stahl et al. looked at a number of different pricing methods that work in real-world data markets. In later projects, the quality of the data affected the price, and people who bought the data were free to set their own rates.

Chawla et al. looked at the static income maximization problem even though the online environment is always changing. They did this by using what they already knew about data users' questions and reviews. Their main ways of setting fixed prices were item pricing and uniform package pricing. Agarwal et al. [48] came up with combinatorial auctions to make it easier for ML tasks to send data.

When scientists traded personal information, they often used the cost-plus pricing method. This means that after the data broker pays the data owners for breaching their privacy, the data user will have to pay more for their inquiries. Researchers looked at a number of different types of info that users were interested in.

Ghosh and Roth looked at one counting question. Li et al. did more research and changed the settings to include more chaotic linear searches.A study was done on how to ask questions about private and messy aggregate figures and data . A study by Hynes et al. looked into the market for training models. It was the study of Chen and his colleagues to see how searching through personal data is similar to pricing a learned model with different levels of noise disturbance. Assuming that the error demands and related valuations of the data consumers were known, they also looked at ways to maximize the data broker's static income.

3.PROPOSEDSYSTEM

The hardest part is figuring out how to use the automatic reservation-pricing system. The data broker is the only one who can figure out how much a query is worth on the market by looking at both new and old queries. This means that the cost of following questions is about the same as the cost of taking an online course. Our price problem is more complicated than just a choice between exploitation and research. It has three separate parts:

A very small amount of information is given out after the investigation begins. Data brokers can't correctly figure out what an item is really worth on the market, so traditional online learning algorithms can't tell if the listed price for an item is higher than its real value.

No matter how important the reserve price and market value are compared, the mentioned price can't go below the expected market value. This is what the reserve price is for. Also, no studies have looked at how this lessened limitation affects learning in general. Lastly, it is very important that we create a very effective method for setting prices online. To put it simply, after picking each price, the data broker needs to quickly change what it knows about the market value model.Because there are no set prices in the linear market value model, the system works better.

It makes the system work better when the Ellipsoid-Based Pricing Mechanism is used.

ADVANTAGES

- Adding experimental posted prices to the linear market value model makes the system work better.
- Because the Ellipsoid-Based Pricing Mechanism is used, the method works better.

4.IMPLEMENTATION

Architecture:

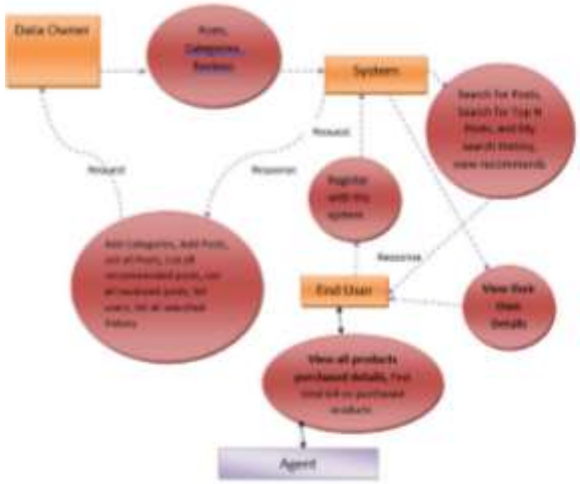

Fig-1:Architecture

Module

User: People buy and sell things and services over the internet, which is called e-commerce. Through the Internet, customers can go to internet stores and buy things without leaving their homes. An enormous selection of goods can be bought through e-commerce sites. Getting almost anything these days is possible through online sites.

Data Owner: A lot of stores used to care more about customers who bought things online.... It is suggested that customers get to know the seller and get their contact information before making any online purchases. Companies with a good name should make this information simple for people to find. You can also keep track of product details like the number of views or sales, as well as the number of happy customers. A trustworthy way to find out how good a company's service is is to ask friends, family, and past customers for suggestions.

Agent:The company Supply sends goods to many stores in a city. Providers also keep track of what customers think and which goods sell the most. Weighted set of common things Modern methods, like mining, make it easier to keep an eye on the quantity and demand of goods in the market.

5. CONCLUSION

This piece talks about a way to set prices that changes based on the starting situation and the reserve price's upper limit. This is meant to help data traders in the online markets for personal information make as much money as possible. Our pricing approach includes a number of different parts that work together to speed up and improve online optimization. It can handle some error when used with both non-linear and linear market value models. Along with that, we showed how to handle two more similar application cases and carefully examined all three usage examples using three real datasets. There is real-world evidence that the reserve price constraint is true and that our pricing method works and can be expanded.

REFERENCES

1. C.Niu,Z.Zheng,F.Wu,S.Tang,andG.Chen,“Onlinepricingwithreservepriceconstraintforpersonaldatamarkets,”inProc.ofICDE,2020,pp.1978–1981.
2. “Factual,”https://www:factual:com/,2008.
3. “DataSift,”https://datasift:com/,2010.
4. “Datacoup,”https://datacoup:com/,2012.
5. “CitizenMe,”https://www:citizenme:com/,2013.
6. “CoverUs,”https://coverus:health/,2018.
7. M. Balazinska, B. Howe,and D. Suciu,“Datamarketsinthe cloud:
8. An opportunity for the database community,” PVLDB,vol. 4,no.12,pp.1482–1485,2011.
9. Roth,“Technicalperspective:Pricinginformation (anditsimplications),”Communications of the ACM, vol. 60, no. 12, p. 78,2017.
10. Li,D.Y.Li,G.Miklau,andD.Suciu,“Atheory of pricing private data,” Communications oftheACM, vol. 60,no.12,pp.79–86, 2017.