CRYPT'AD: Privacy Focused Reward based Ad System

Chirag Wadhwa*, Heena Jeswani, Tamanna Rohra, Lifna C.S.

Department of Computer Engineering, Vivekanand Education Society's Institute of Technology, Mumbai, Maharashtra, India

Publication Info

Article history:
Received : 13 February 2020
Accepted : 18 May 2020

Keywords:
Ad network, Bidding, Digital Currency, Privacy, Rewards, Recommender System.

Corresponding author:
Chirag Wadhwa
e-mail: *2016.chirag.wadhwa@ves.ac.in*

Abstract

Ad networks are an essential monetization element for an Advertising Ecosystem. These networks act as a broker between Advertisers and Publishers. The paper illustrates a system that contains an Ad Server that incorporates bidding between the selected Advertisers for a particular visitor. The selection process of Advertisers includes a Recommender System based on end-user Surveys. Whenever a visiting end-user takes a survey, the publisher will be rewarded with Digital Currency. With these rewards, Publishers can start their own ad campaign.

1. INTRODUCTION

Online Advertising is one of the fastest-growing areas in the IT Industry. Promotions through the Internet is increasingly accessible for every business and is simply just as progressively productive to arrive at their intended interest group. Almost every business has some kind of internet presence so that they can reach a large number of clients with the least speculation and most extreme benefit, which can be accomplished through Internet Advertising.

This paper illustrates an advertising system that will be focusing on the privacy of users and it will also be useful for targeted advertising like ads shown in web pages, targeting based on demographics, surveys etc. This paper explains a system in which publishers can monetize their content and ad advertisers can display their ads to the publisher's site. The system contains an ad server that incorporates bidding between many advertisers the highest bidder that would be decided based on some parameters like initial-Final bid range, time and highest bidder's ad will be served on to the publisher's site. Publishers will log in and share their ad inventory; they can set the priority of the ad running on their site. Publishers will also be given a revshare based on the ad traffic on their site. Publishers will be rewarded based on our surveys to get the user information and with their rewards, publishers can start their campaign.

There will be a different Survey area i.e. Survey will be led to pick up the data about the client and dependent on the given Information the Targeted Ad will have appeared to the particular clients with the goal that the client gets promotions of their enthusiasm as the ad will be indicated depending on their review. When clients associated with the study appear on the distributor's site, the distributer will be compensated.

2. LITERATURE SURVEY

This section, discuss the survey performed during the conceptualization of the concept. Survey started with a detailed study on the state-of-the-art techniques on Auctions, Bids, Data Collection techniques, User Identification and Recommendation Systems.

2.1. Ad Auction

As discussed in paper [1], auctions are used to determine the User Ads and its order. Paper [2] mentions the role of Advertiser Bids that control Ads' regularity and position to the Publishers Website. As discussed in paper [1] a simple auction can be summarized as follows: (1) Initially the dealer floats the action between Ads from Advertisers that are targeted to a given Publisher. Auction takes place when the browser requests a set of Ads. (2) Generates a rank among the Ads. (3) Maintain the ranks and filter these ads based on their ranks. (4) The ranks are then updated based on the factors like demographics. (5) Finally, Ads are shown to the user according to the updated ranks.

2.2. Surveys for Data collection

End-Users fill the Web surveys through their devices like cell Phones, desktop, Tablets via internet Connectivity. Social and demographic factors affect the participation in surveys, Gender is the one of the main factors that determine the participation rate like for example females participate in huge numbers compared to males. As mentioned in paper,[3] designing the surveys also plays an important

role in users' participation. The respondents to pop-up Web Surveys are good if designed with proper Graphics, different font types, font sizes, and pictures.

2.3. User Identification

Web storage API provides a method for storing client-side data instead of using cookies, it also provides a method to store data in the browser. The data is only accessible in the browser, it is not sent to a server like cookies do.It gives two storage mechanisms: session storage and Local Storage. Both of these give private areas for data. Session storage stores data for the duration of the page session,if the window or tab is closed the session storage for that particular tab is cleared. At the same time, local storage persists the data until it is explicitly removed. Local storage carries some edge over the use of cookies.The amount of storage available may differ most arbitrary limit is 5MB per origin.[4] Domains can store a huge amount of data. This huge amount of data can also be used to store user-edited data such as documents or file.

2.4. Recommendation System

Nowadays, we are available with the facility to filter and prioritize the information delivered to users. With the help of recommendation systems we can search from a huge amount of data produced for the users to give customized ads.[5]

2.4.1. Information Gathering Phase

Relevant information of users is collected from user behavior or through contents of the resources user accesses.

2.4.2. Learning Phase

It applies the algorithms to filter users' features from the previous phase's information.

2.4.3. Prediction/Recommendation Phase

It provides predictions/recommendations about what kind of items users may prefer. There are various algorithms available for recommendation systems; one of them is alternating least squares algorithm.

2.4.4. Alternating Least Square (ALS) Algorithm

Matrix Factorization is a mathematical operation for matrices. In this operation, the matrix is divided into a product of matrices. Als is a matrix factorization algorithm. ALS divides the user-item matrix into the user-feature matrix and item-feature matrix.ALS runs in a parallel manner.ALS is implemented using spark ML and it is used for collaborative filtering problems. ALS minimizes loss functions by taking the matrices alternatively i.e, by holding one matrix fixed and computing another. First, it fixes the user matrix and performs computation on the item matrix then, it fixes the item matrix and performs computation on the user matrix.

3. PROPOSED SYSTEM

CRYPT'AD focuses on the privacy of users as well as completes the basic market needs of targeted advertising. As there are a number of advertisers interested in showing their advertisement on a particular publisher's website, So filtration is required. Initially, we match the advertisers product categories with the publisher's categories so that relevant advertisements are shown on the publisher's website. After this filtration, We match the interest of users of the publisher with the advertiser's advertisement. To do this i.e, to get the recommendations, We use Alternating Least Squares Algorithm. After this, bidding between many selected advertisers takes place and finally, the winner, i.e the Ad of highest bidder, is shown on the publisher's website. But, to get the users interested, we require some information but keeping users' privacy in mind. For this, we take surveys from the user to get their interests and do not store any other user's personal information. The surveys are taken on the publisher's website, for each survey shown on the publisher's website the publisher gets rewarded. Using these rewards publishers can start their own Ad campaign using our platform.

With referenve to figure 1 The whole process can be explained as follows:

- Initially, advertisers will select their categories, set their initial bids and budget for each ad campaign.
- Publishers will share their ad inventory and will select categories for the type of advertisements.
- Filtration of advertisers and publishers according to their categories and recommendations generated will be done.
- Bidding between many selected advertisers will be done and the system will serve the highest bidder's advertisement to the publisher's website.

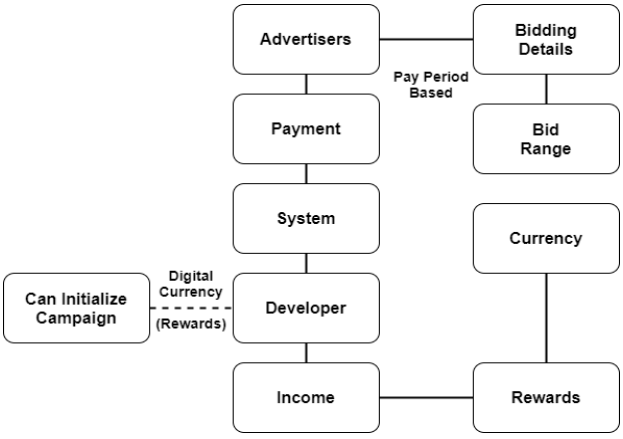

Figure 1: Proposed System Architecture

2010-2019 S-JPSET : Vol. 12, Issue Supplementary-1, ISSN : 2229-7111 (Print) and ISSN : 2454-5767 (Online)

- Publishers will be given revshare based on the ad traffic on their websites.
- Publishers will be rewarded based on our surveys to get the user information.
- With these rewards,publishers can start their own ad campaign.

4. IMPLEMENTATION

A website that connects Advertisers and Publishers. Many Advertisers register and provide information about their products, Initial bid to start their campaign. Publishers also register and provide relevant information about their website.

The website uses the technologies : NodeJs, Express Framework, Ejs as View-Engine, Passport for Authentication, MongoDB for database.

The advertiser inputs the following information: name of their campaign, Start and the end date of the campaign, start time, product's image, Category of the product, color of the product, name of the product, price range, other related keywords, URL of the product's site and bidding range.

Bidding between many selected advertisers takes place. And the ad of highest bidder will be shown on the publishers website. The selection process of advertisers includes a We have used node js timers to implement bidding on each request set_Interval timers are attached to each advertiser and one set_Timeout which will stop the bidding and data of the winner will be sent to the client.

Bidding algorithm as follows:

1. For all the advertisers define function
2. Call SetInterval function
3. Bids[i] <- max_bid + bids[i]/2
4. Set Highest_bid <- bids[i]
5. Define display function
6. Call clearInterval function
7. Display Winner name
8. Call setTimeout function

4.1. 1 setTimeout(callback, delay[, arguments])

callback(function) The call function when a timer is running out.

Delay (digit) The digit of milliseconds to wait before callback is called.

arguments (any)When callback is called, optional arguments to move.

4.2. 2. setInterval(callback, delay[, arguments]):

callback(function) The call function when a timer is running out.

Delay (digit) The digit of milliseconds to wait before callback is called.

recommender system that is based on the surveys when the user visits the publisher's website.

For the recommendation system Alternating least squares

Algorithm(ALS) is used arguments (any) When callback is called, optional arg move

ALS Pseudo code is as follows:

1. ratings<- spark.createDataFrame(sd)
2. ratingsRdd <- ratings.rdd.map(list)
3. rd2 <- ratingsRdd.map(lambda l: Rating(arguments...)
4. collect rd2
5. model <- ALS.train(rd2, rank, numIterations)
6. prediction <- model.predict(arg1,arg2)
7. TopUsers<-model.recommendProductsForUsers(arg)
8. collect TopUsers
9. Display predictions

4.3. About Surveys

The Surveys are conducted in the publisher's website to gain the end-users' interest so that based on their interests the Ad will be shown of their interest.

4.4. How are we identifying the user without using cookies ?

Using the local storage of the browser, we have implemented user identification. Each browser has its own local storage. When a publisher's website is loaded in the browser, we check for the value stored in the localstorage. If the value is present in the localstorage, then we use that value as a unique user. If value is not present, we create a new user in our database and store the ID of that created user in the browser's localstorage. In this way we keep track of a user.

4.5. Pseudocode for User Identification through Local Storage

token ← onLoad(check value in localstorage of browser)
if token is null:
    id← createUser(return userId)
else:
    id← value of token
    findById(id)

Now inside the database, we are storing UserId,

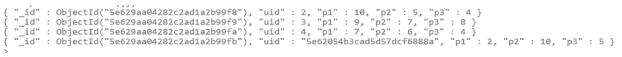

Figure 9: Userid with top 3 products stored in mongodb.

```
>>>
== RESTART: C:\Users\heena\AppData\Local\Programs\Python\Python36\rec21.py ==
(1, (Rating(user=1, product=2, rating=2.9563389903112096), Rating(user=1, produc
t=10, rating=1.9926428899695798), Rating(user=1, product=5, rating=1.00719650643
68422)))
(2, (Rating(user=2, product=10, rating=1.9838917623003418), Rating(user=2, produ
ct=5, rating=1.975106512319801), Rating(user=2, product=4, rating=0.983160644823
541)))
(3, (Rating(user=3, product=9, rating=1.9512649630695273), Rating(user=3, produc
t=7, rating=0.9896028623890158), Rating(user=3, product=8, rating=0.982617680930
6857)))
(4, (Rating(user=4, product=7, rating=1.9746526271978055), Rating(user=4, produc
t=6, rating=1.9659090561938974), Rating(user=4, product=8, rating=0.991698123382
1058)))
2.9563389903112096
```

Figure 10: Output of prediction

2010-2019 S-JPSET : Vol. 12, Issue Supplementary-1, ISSN : 2229-7111 (Print) and ISSN : 2454-5767 (Online)

Figure 4: Final Result after many computations shown to the End user

Advertiser's Campaign Number and Rating which the users had given to a particular Campaign of Advertiser.

From the above ALS algorithm, the following predictions were shown in Figure -9. From fig 9, It is shown that the user with id : 5e62054b3cad5d57dcf6888a had rated product 2 the highest, Figure-10 explains the output of predictions.

Figure 4 shows the final result; after many filtrations

The winner's Ad is displayed on the publisher's website.

5. LACUNA OF THE EXISTING SYSTEMS AND ITS COMPARISON WITH THE PROPOSED WORK

5.1. Privacy Issues

Cookies can store a wide range of information, including personally identifiable information which includes Name

- Home address
- E-mail address
- Telephone nmber

A third-party cookie is one that is placed on a user's hard disk by a Web site from a domain other than the one a user is visiting.

5.2. Existing Systems and its comparison table

Table 1: Comparison with existing system

Parameters	*Existing System*	*Our System*
Tracking Method	Cookie-based tracking	Survey based tracking
Rewards	Normally publishers are not rewarded for ad companies using their platform to understand user (ad viewer).	Publisher is rewarded for each survey displayed to their users in the form of cryptocurrency. With this cryptocurrency, publisher can start their own Ad campaign without actually investing in real money.
Privacy	Many ad companies try to take user's (ad viewer) Personal information.	Our system does not store any info. like ad viewer's name, phone no. , actual address.

5.3. Privacy concerned recommendation system

According to a report by data privacy consultants,Levels of concern over online privacy are growing.The report found Eightynine percent of British consumers are worried about data privacy, with 60% of those people saying that they are more concerned than they were a year ago. So, our system will use surveys for collecting user data, thereby creating transparency between system and user.

In our project we will use a Recommendation system, these systems are software tools used to generate and provide suggestions for items and other entities to the users by exploiting various strategies. Hybrid recommender systems combine two or more recommendation strategies in different ways to benefit from their complementary advantages.

Surveys will be fed into the recommendation system which will increase the targeting accuracy.

6. CONCLUSION

In existing systems cookies are used for collecting user data and cookies can store personal information like name,email address, phone number etc. Crypt'Ad will use surveys for collecting user information, thereby creating transparency between the Ad System and Users. This paper discusses the design and implementation of Crypt'Ad, which is a focused Privacy Reward based Ad System. The system contains an Ad Server that incorporates bidding between the selected Advertisers and will serve the publisher's site's highest bidder. The work can be extended to include Mobile Ads and rewarding Publishers for displaying the surveys to their website for user data collection, and publishers can use these rewards and can start their own Ad campaign using the same platform.

7. ACKNOWLEDGMENT

We would like to express our deep regards towards Principal madam Dr. (Mrs), J.M. Nair, HOD madam Dr. (Mrs) Nupur Giri and Project Mentor Mrs. Lifna C.S for their guidance and appreciation. It is due to their patience and enthusiasm, which has given a sense of direction and purposefulness to this paper.

8. REFERENCES

[1] Yuan, Y., Wang, F., Li, J., & Qin, R. (2014, October). A survey on real time bidding advertising. In Proceedings of 2014 IEEE International Conference on Service Operations and Logistics, and Informatics (pp. 418-423). IEEE.

[2] Adikari, S., & Dutta, K. (2015, May). Real time bidding in online digital advertisement. In International Conference on Design Science Research in Information Systems (pp. 19-38). Springer, Cham.

[3] F.Keusch, "Why do people participate in Web surveys? Applying survey participation theory to Internet survey data collection," Management review quarterly, vol. 65, pp. 183-216, Jun 2015

[4] Wang, Y., Burgener, D., Kuzmanovic, A., & Maci, G. (2011, June). Understanding the network and user-targeting properties of web advertising networks. In 2011 31st International Conference on Distributed Computing Systems (pp. 613-622). IEEE.

[5] Isinkaye, F. O., Folajimi, Y. O., & Ojokoh, B. A. (2015). Recommendation systems: Principles, methods and evaluation. Egyptian Informatics Journal, 16(3), 261-273.

[6] Guha, S., Cheng, B., & Francis, P. (2011, March). Privad: Practical privacy in online advertising. In USENIX conference on Networked systems design and implementation (pp. 169-182).

[7] Radhika, V. M., Thottungal, A., & Nizar, M. A. (2016, August). An enhanced model for behavioral targeting in online advertising. In 2016 International Conference on Data Science and Engineering (ICDSE) (pp. 1-5). IEEE.

[8] Cantrell, C. (2002). U.S. Patent Application No. 09/726,422.

[9] Mayer, J. R., & Mitchell, J. C. (2012, May). Third-party web tracking: Policy and technology. In 2012 IEEE symposium on security and privacy (pp. 413-427). IEEE.

[10] Pilászy, I., Tikk, D., Takács, G., Bottyán, A. N., & Zibriczky, D. (2014). U.S. Patent No. 8,676,736. Washington, DC: U.S. Patent and Trademark Office.

[11] West, W., & Pulimood, S. M. (2012). Analysis of privacy and security in HTML5 web storage. Journal of Computing Sciences in Colleges, 27(3), 80-87.