A Cognitive Optimisation for Advertisement Using Particle Swarm Optimization

Joachim S. Arogundade[1], Latifat A. Odeniyi[2] and Olufunke R. Vincent[3]

[1,3]*Department of Computer Science, Federal University of Agriculture Abeokuta, Ogun State, Nigeria.*

[2]*Department of Physical Sciences, Chrisland University, Owode, Ogun State, Nigeria.*

joachimseun@gmail.com, Vincent.rebecca@gmail.com

Abstract- *In the world of marketing, online advertising is one of the most efficient ways for businesses to expand their capacities, increase their customer base and extend their income rates. However, one of the main challenges faced by advertisers, especially the newcomers in the world of internet advertising is setting up their campaigns in a structured way to enhance conversion. The main goal of structuring advertisement is to increase view ability, click through rate and conversion rate of adverts. This paper presents a cognitive optimisation for display advertisement campaigns through the behavioural targeting and retargeting campaigns techniques by using particle swarm optimization as the fundamental optimization model. We implement our methodology to optimise display advertisement scheduling and placement on mobile devices and desktop using a real-world data set. Results show that the output of our methodology is feasible to optimise the advertising campaign by selecting the set of the best features as each advert moves in the direction to its previously best position and the global best position in the advertisement networks.*

Keywords: Online Advertisement, Optimization, Particle Swarm Optimization, Behavioural targeting and Retargeting.

I. INTRODUCTION

There is much more to online advertising than simply placing an advert on the internet and hoping for the best, the most effective advertising campaign combines numerous interconnected element, all of which perform unique functions to maximise the advertisement potential [9]. Advertisements have to be set up in a structured way in order to guarantee high acceptance of users to adverts. It is also required that the number of visits that satisfy the structuring requirement is good enough to cover the advertisement goals [11].

There are two main categories of online advertising campaigns which are sponsored search engines and display advertising. In sponsored search advertising, it considers to find the business found on search engines by using related keywords, while display advertising, considers showings ads to a target audience mainly in form of banners [1]. Furthermore, optimization for display advertisement can be classified into three categories,(i) the ad-scheduling and placement optimization considers determining where to expose adverts on websites[10], (ii) revenue management and pricing optimization considers define pricing schemes and revenue management model[8],(iii) it is the approach for display advertisement effectiveness that considers the impact of content and design based on the click through rate(CTR) metric(the probability that a user generates a click on an advertisement) and the ad allocation problem [26].The importance of optimizing advert cannot be overemphasized, one of the great benefits of optimizing advertisement is that it put product and services been advertised on top for the viewability of prospective visitors [5]. Online advertisement has become one major medium which business, organization or establishment could function effectively in a competitive environment. Manufacturers and organizations use adverts as a means of reaching their intending customers, as regards the goods and services they make available [28] Furthermore, [7] stated that once the visitors see your websites and what the pages have to offer, they will be strongly motivated to click, explore and visit more of your pages and consume your content.

In the last decade, researchers have carried out some work in the area of optimization of advertising campaigns. [21] Formulated stochastic version of budget optimization problem of advertisement. [7] Proposed an optimal method for the ad allocation problem from the advertiser's point of view, presenting a nonlinear model for maximization of reach in in display advertising, the work focuses on branding scheme considering only properties of the websites. Also, [24] focuses on optimizing advertiser satisfaction on the understanding that advertisers will be more willing to make investment if they get good profit. [17] Aims to increase the performance of online publicity by selecting the right users to generate conversions. A conversion refers to a purchase, a form filling or a phone call. To this end, a set of mathematical models are generated. [1] Proposed a mixed linear programming via piecewise linear approximation of the revenue function for both branding and direct response, considering the ad location and contents in the optimization model. [15] Focuses in online bid optimization. They presented an online approach to optimize the performance metric while satisfying the smooth delivery constraint for each campaign. [3] Make an in-depth analysis of the balance that must exist between economic performance, the most profitable ad selection, and the quality that is offered to advertisers. [20] Presents a novel methodology for optimizing the micro-targeting technique in direct response display advertising campaigns by using genetic algorithms as the basis optimization model and a machine-learning based click-through rate (CTR) model.

Thus, this paper presents a cognitive optimization for display advertisement through the behavioural targeting and retargeting technique in display advertisement as an ensemble of particle swarm and an online logistic regression based CTR model. From the advertiser's point of view, our model consists on optimizing display advertisement by the selection and structuring of the parameters in a campaign, to satisfy advertisers' constraints for possible behavioural targeting and retargeting approach. In the optimization model, we proposed an ensemble of computational intelligence techniques composed of particle swarm optimization as the optimization model and the online logistic regression algorithm for CTR modelling.

Our model heuristically considers to display the most interesting adverts for customers and that the number of visits that meet the configuration requirements is sufficient to cover the advertisers' demand. In particular, we focus on a potential positive behavioural targeting and retargeting that online advertising has on the propensity and the form of user interactions with an advertiser in the future, and develop improved algorithms for the problem in this setting. We clarify these ideas on a simple scenario from the direct response display advertisement organised course of action. Thus, the proposed methodology maximizes the objective function (formally known as fitness function for PSO) considering: (i) the interest of users for the adverts measured as the average CTR of all predicted visits, and (ii) the number of visits that matches those settings, not looking for configurations that include a huge number of visits, but targeting configurations that ensure a sufficient number of visits to meet the demand of the majority of the advertisers. In addition, to calculate the average CTR of the objective function we use the online logistic regression method since its performance has been widely demonstrated by [19] and [4]. Our proposal allows suggesting advertisers' commercial successful structuring in their direct response display advertisement placement. Also from space-time complexity point of view, our proposed model is cost effective and easy to implement. Thus, the main contribution of this work considers the implementation of particle swarm optimization for optimizing direct response display advertising campaigns, since, in this research area heuristic optimization have been limited to only one previous research, especially to maximize nonlinear objective functions.

II. RELATED WORK

Previous research examined the optimization of advertising campaigns using different models and focusing on the optimization different metric and parameters in advertising. [21] In their work, Stochastic models for budget optimization in search based advertising proposes budget optimization as it arises in Internet search companies and formulate *stochastic* versions of the problem. The premise is that Internet search companies can predict *probability distributions* associated with queries in the future. They identify three natural stochastic models. Their main results are algorithmic and complexity results for the three stochastic models. their algorithmic results show that simple *prefix* strategies that bid on all cheap keywords up to some level are either optimal or good approximations for many cases it shows other cases to be NP-hard. [25] But going through the three-stochastic model proposed is too complex and it will take time to execute because of the complexity of the models.

[7] Proposed an optimal method for the ad allocation problem from the advertisers' point of view, presenting a nonlinear model for maximization of reach in in display advertising, the work focuses on branding scheme considering only properties of the websites.

[1] Formulates an advertiser's revenue maximization problem in direct response Internet display advertisement campaigns as a mixed integer program via piecewise linear approximation of the revenue function. His approach is that

ad location and content issues are explicitly incorporated in the optimization model. Computational experiments on a large-scale actual campaign indicate that adopting the optimal media schedule can significantly increase advertising revenues without any budget changes, and reasonably sized instances of the problem can be solved within short execution times.

[15] Presented an online approach to optimize the performance metrics while satisfying the smooth delivery constraint for each campaign. Their approach first applies a control feedback loop to iteratively estimate the future spending rate in order to impose smooth delivery constraints. Then, the spending rate is used to select high quality impressions and adjust the bid price based on the prior performance distribution to maximize the performance goal. The [2] model tries to select high quality impressions and adjust the bid price based on the prior performance distribution in an adaptive manner by distributing the budget optimally across time.

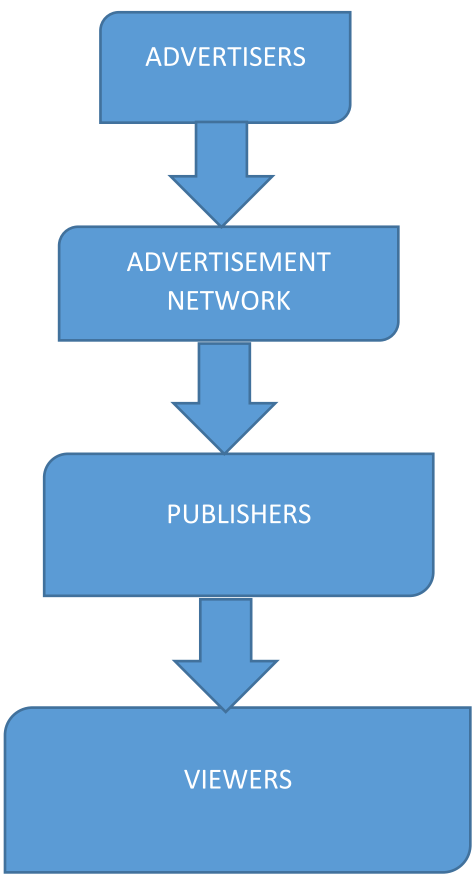

Fig 1: overview of advertisement process

[3] in their work yield optimization of display advertising with ad exchange, stated that Ad Exchanges are an emerging market for the real-time sale of online ad slots on the Internet. They presented an approach to help publishers determine when and how to access AdX to complement their contract sales of impressions. In particular, they model the publishers' problem as a stochastic control program and derive an asymptotically optimal policy with a simple structure: a bid-price control extended with a pricing function for the exchange. We show using data from real inventory that there are considerable advantages for the publishers from jointly optimization over both channels. They concluded that, Publishers may increase their revenue streams without giving away the quality of service of their reservations contracts, which still represents a significant portion of their advertising yield.

Recently, [20] present a novel methodology for optimizing the micro-targeting technique in direct response display advertising campaigns by using genetic algorithms as the basis optimization model and a machine-learning based click-through rate (CTR) model. From the advertiser's point of view, the methodology consists on optimizing display advertising campaigns by the selection and configuration of the parameters in a campaign, to satisfy advertisers' constraints for possible micro-targeting approach. The optimization model employ an ensemble of computational intelligence techniques composed of genetic algorithms as the basis of the optimization model and the online logistic regression algorithm for CTR modelling. In addition, the methodology was implemented to optimize display advertising campaigns on mobile devices, that is, adverts are displayed on mobile screens through apps as banners.

III. OPTIMISATION OF ADVERTISEMENT PROCESS

Advertisement optimization consists of a hybrid soft computing method based on particle swarm optimization and an online logistic regression, aiming to determine the best subset of features that maximizes the CTR of a given advertising campaign. In that case, the model can be adapted to any properly ad campaigns models and to set suitable parameters in the optimization process. Figure 2 shows the proposed methodology. The overall optimization procedure is carried out by particle swarm optimization. Then, each particle of the SPO correspond to campaign configurations, and at each iteration, these particles fitness values are calculated. Particularly, this methodology proposes a heuristic fitness function that optimizes both the number of visits and the average CTR of a campaign. To this end, the online logistic regression model proposed by [19] and [4] is used to model the estimated average CTR of a campaign. Details of the proposed methodology are described below.

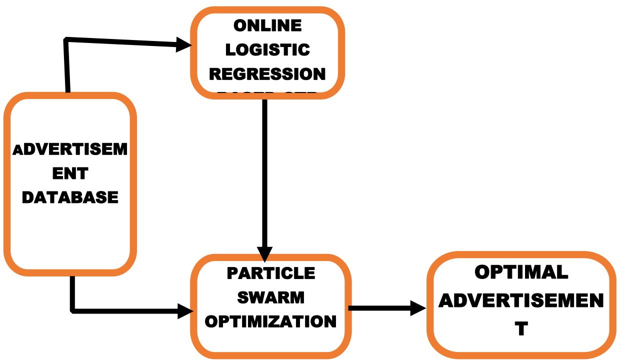

Fig 2**:** Schematic of the proposed methodology for online advertising campaigns optimization.

A. *The prediction Click through Rate model*

Consider a dataset of online advertising campaigns with D samples (rows) and N numerical and/or categorical features, fi, (columns). Each sample describes one user's exposure to an advert, represented by a set of features related to the ad, the user and the metadata. For example, these features can be the operating system, browser, age, time, and type of product, among others. An additional column, y, in the dataset represents the real value that a user clicks (1) or not (0) into the online advert. In that sense, the CTR model aims to predict the probability CTR value, p, of a visit, such that, $p = P(\frac{y}{fi})$. There are many types of research focused on accurately predicting the CTR of an advert [27] and [15]. Our goal is to implement a model as precise as possible in predicting the CTR. At the same time, it is highly recommended to build a very cost-effective model in terms of time and memory computing resources. This methodology can be easily implemented on small ad networks at a low cost and in a short period of time. For this reason, we have decided to implement the CTR estimation model proposed in [19] that is based on the online logistic regression method. The model is explained in detail in many scientific papers, e.g. [4] and [16]. Currently, this model applies the hashing trick technique of [16] and the adaptive learning rate of [6].

In a nutshell, the hashing trick is an ingenious method to model data sets with large quantities of information using a hash function. The latter transforms the categorical and numerical features of each entry into an integer value within a range between zero and D. It is highly recommended to use a large number for parameter D in order to avoid collisions. Collisions happen when different original values generate the same number after applying the hash function. Hash values are used as the array index. The benefit of applying the hashing trick is that it reduces the spatial dimensions, and therefore, the memory and the time required to create the model. The hashing trick method has become famous for its simplicity and its effectiveness [11]. The algorithm for building the online logistic regression model uses two arrays, n and; where n is an array of integers that represents the number of times each feature appears after applying the hashing trick and w is an array of real values that represent the weights associated with each feature. The values for w are updated using (1); where, $w[i]$represents the weight of the $i - th$ feature (initially set to zero), $n[i]$ represents the number of times the $i - th$ feature appears after applying the hashing trick, $y \in \{0,1\}$ is the target output value, $p \in [0,1]$is the estimated output value, and $\propto$ represents the heuristic adaptive learning rate for optimizing the online logistic regression model.

$$w[i] = w[i] - \frac{\propto (p - y)}{\sqrt{n[i]} + 1} \qquad (1)$$

Once w and n, the weight and frequency arrays, are totally updated, the output for each entry, p, is predicted using the sigmoid function of (2).

$$p = \frac{1}{1 + \exp(-\sum_{i=1}^{N} w[fi])} \qquad (2)$$

B. *Design of the particle swarm optimization for online advertisement*

Particle swarm is a swarm intelligence metaheuristic optimisation procedure that implements simple operations observed in the adaptation and evolution of species. The main concept of particle swarm optimization is that it represents the state of the algorithms by a population, which is iteratively modified until a termination criterion is satisfied [22]. In particle swarm optimization algorithm, the population $p = \{pi, \dots, pn\}$ of the feasible solutions is often called swarm. The feasible solutions $pi, \dots, pn$ are called particles. The particle swarm optimization method views the set R^d of feasible solutions as a "space" where the particles "move". For solving practical problems like the optimization of advertisement, the number of particles is usually chosen between 10 to 50 [13].

Algorithm 1*: Simple particle swarm optimization Algorithm*

(Input: particle fitness function, swarm size and total number of iterations)

1. $For\ each\ time\ step\ t\ do$
2. $For\ each\ particle\ I\ in\ the\ swarm\ do$
3. $Update\ position\ xi(t)$
4. $calculate\ particle\ fitness\ f(xi(t))$
5. $update\ pi, py$
6. end
7. $End.$

(*Output: The last best score is returned along with value of x for which we get best. best score = maximal best score*)

C. *Particle Swarm Optimization variants*

There is a plethora of different versions of PSOs, which usually modify the formula for the change of velocity (e.g., instead of $u1\ and\ u2$ they use diagonal matrices $U1\ and\ U2$, in other variants they use no inertia, but enforce an upper limit on the particle speed, there is the so-called "fully informed" PSO, and there is also a popular modification using a "constriction coefficient"). There exist versions of the PSO for constrained optimization, for discrete optimization, and for multi-objective optimization [30]

The general structure of a canonical PSO algorithm is as follows:

Procedure Particle Swarm Optimization

$begin$

$Initialize\ xi, vi\ and\ xbesti\ for\ each\ particle\ i;$

$while\ (not\ termination\ condition)\ do$

$begin$

$for\ each\ particle\ i$

$Evaluate\ objective\ function;$

$Update\ xbesti$

end

$for\ each\ i$

$Set\ g\ equal\ to\ index\ of\ neighbor\ with\ best\ xbesti;$

$Use\ g\ to\ calculate\ vi;$

$Update\ xi¼xi + vi;$

$Evaluate\ objective\ function;$

D. *Notation and Terminology of particle swarm optimization algorithm*

A population or swarm is a set of K particles located in D-dimensional space. At iteration t (where $t = 1, \dots, T$), particle i (where $i = 1, \dots, K$) has a position $Xi(t) \equiv (xi1(t), \dots, xid(t), \dots, xiD(t))$ and a velocity $Vi(t) \equiv \left(vi1(t), \dots, vid(t), \dots, viD(t)\right)$. The position and velocity components satisfy $Xmin \leq xid(t) \leq Xmax\ and\ [vid(t)] \leq Vmax$, respectively. The velocity $Vi(t = 1)$ is the rate at which particle i moves from position $Xi(t)$ to position $Xi(t = 1)$.

Each position $Xi(t)$ may directly or indirectly represent a solution of a specific problem. The objective function value of the solution of a specific problem decoded from the position $Xi(t)$ is denoted by $f(Xi(t))$. This is also known as the fitness value. In a minimization problem, Xi is 'better' than Xj if $f(Xi) < f(Xj)$. Here we only consider minimization problems since a maximization problem can be turned into a minimization problem by negating the fitness value.

The personal best position of particle i is the position at which $f(Xi)$ achieved its lowest value so far and is denoted by $Pi \equiv (pi1, \dots, pid, \dots, piD)$. The global best position, $Pg \equiv (pg1, \dots, pgd, \dots, pgD)$, is the best position found by the swarm so far. The local best position for particle i is the best position found by particles in the na-adjacent neighbourhood of particle i and is denoted by $Pli \equiv (pli1, \dots, plid, \dots, pliD)$. For odd na, the adjacent neighbourhood of particle i contains the particles$i - \frac{na-1}{2}, \dots, i, \dots, i + \frac{na-1}{2} where\ K$ is added or subtracted from the particle index if it lies outside $1, \dots, K$[29].

The near neighbour best position is denoted by $Pni = (pni1, \dots, pnid, \dots, pniD)$ in which $pnid = pjd$ where, for each d, the value of $j \neq i$ is chosen so as to maximize the value of the fitness-distance ratio,

$$\gamma(j, i, d) = \frac{f(Xi - f(pj)}{[pjd - xid]} \quad (5)$$

Hence, unlike the other best positions defined above, Pni will in general not be a position of an existing particle.

E. *Fitness Function*

The particle swarm optimization seeks for configurations that maximize two variables: the average CTR and the number of visits in an online advertising campaign. Thus, we propose a fitness function that evaluates an advertising campaign heuristically by considering: the number of visits to the campaign calculated as the number of particles in the subset, K' and the estimated average click-through rate, $CTRaverage$, as written in (3) where, T is a threshold value to avoid solutions with larger numbers of visits.

$$f(particle\ i) = CTRaverage \times \min(K', T) \quad (3)$$

Particularly, the average CTR is computed by evaluating each sample k of the subset into the machine learning model (2) to estimate the CTR value pk and obtaining the mean value of these predictions, as expressed in (4). It should be noticed that we are looking for enough number of visits to satisfy the expectations of the advertisers, taking into account that most advertisers have a limited budget. Thus, we are trying to obtain configurations which generate a high average CTR value, and with a number of visits that simply equals or exceeds, albeit slightly, the threshold value T. In this work, we propose that T should be set manually by the ad network or the advertiser.

$$CTRaverage = \frac{1}{D'} \sum_{k=1}^{K'} pk \quad (4)$$

IV. *Result and Evaluation*

A brief description of the dataset used is presented. Then, we explain the process to build and train the CTR model. Lastly, we describe the methodology of the experiments conducted in this work to optimize online advertising campaigns.

A. *Description of the Dataset*

In this paper, we use the SocialPeta Big data for decision on advertising. SocialPeta is mostly targeted to display adverts on users' mobile devices through apps, games, or mobile web pages. Adverts are oriented to a conversion such as an application download, filling out a form or product purchase. SocialPeta advertisers can target their campaign by setting several parameters, such as geolocation, time, city, Internet connection, operating system version, device type, traffic type and some others. This platform has a presence in more than 130 countries in 2016 and a volume business above 1.3 billion dollars [23]. Particularly, this dataset corresponds to 10 days of click-through data organized in more than 40 million samples, and it is composed of a set of columns that represents features from the users and the websites. Twenty-four attributes are comprised of the dataset: ad identifier, click ("0" for non-click and "1" for click), hour, banner position, site identifier, site domain, site category, application identifier, application domain, application category, device identifier, device IP address, device model, device type, device connection type, and other nine anonymized categorical variables corresponding to the contents of the adverts. Features are both numerical and categorical. Different number of instance values, minimum, maximum, four quartiles, median, mean, and class. Each column represents a feature in the dataset, except for the column value "click" which represents the target output value. As noted, this dataset is unbalanced since only 17.45% of the samples are targeted as click, and the remaining 82.55% represents non-click values.

B. *Implementation of the Prediction CTR Model*

Firstly, we built and trained the prediction CTR model, as described in Section 3.2. Thus, we randomly selected 10-million samples from the original dataset: the first 8-million samples for training and the remaining 2-million samples to evaluate the prediction CTR model. In addition, we converted all features from strings to integer values, and we applied module 2^{20} to calculate the hashing table. Lastly, we applied the online logistic regression method to build the prediction CTR model with the following settings, chosen manually: learning rate $\propto = 0.1$, and $D = 2^{20}$ as the length of arrays $n\ and\ w$. To carry out the CTR model, we use Java Studio creator 3.3.2. Different performance metrics were calculated to the resultant CTR model, in this case, the CTR-model response computes 39.67% in the logarithmic loss metric, just 1.76% worse than the best prediction CTR model obtained with the 40-million original dataset (Socialpeta, 2015); representing a good prediction CTR model that can be used for further experimentation.

C. *Configuration of the Particle Swarm Optimization*

The next step was to configure the setting parameters of the Particle Swarm Optimization. Particularly, an exploration of the probability of inertia weight and acceleration coefficients was conducted. Due to large computational time, we employed only 20-thousand samples randomly chosen from the testing dataset for the preparation of the setting parameters. To do that, we set 9 different values for w(t) and p(t) in the interval [0.1, 0.9] with steps of 0.1. Then, we performed a set of 30 experiments for each fitness value, and the average of the best fitness evaluations of that experiments was reported. Additionally, the population size was set to 500 and the percentage of elitism was 5%. For the threshold value in the fitness function, we selected T = 20. Notice that we fixed the number of iterations to 300, for comparison purposes. The best fitness evaluation was obtained when w(t) = 0.2 and p(t) = 0.3.

D. *Explanation of the Experiments*

This experiment aims to implement the proposed optimization and to determine the characteristics of the optimization procedure. The prediction CTR model depicted in Section 4.2 and the selected parameters analysed with T = 2000. The testing dataset of 2-million samples was used. The analysis of the optimization procedure considered 50 runs and the computation of the mean and standard deviation of the best fitness evaluation of each run.

Table I: shows some of the experiments carried out and the values of PSO parameters

Iterations	Fitness Evaluations	Average CTR	SPO PARAMETERS		
			K	Pi(pbest)	Pg(gbest)
162	945.91	0.3791	45	2.1304	1.0575
88	945.81	0.3791	67	1.6319	0.6239
126	945.72	0.3789	60	2.3255	1.0056
64	945.68	0.3781	96	0.6485	2.6475
184	944.93	0.3781	116	0.5287	3.1913
224	944.85	0.3781	290	1.6319	0.6239
70	944.76	0.3779	50	0.0513	4.9087
25	943.98	0.3777	44	0.1136	3.9789
55	943.91	0.3769	78	0.1564	3.8876
201	943.87	0.3765	89	0.2699	3.3953

E. *Discussion*

Firstly, the proposed online advertising campaigns optimization method was implemented and We ran the same algorithm 10 times in order to determine the characteristics of the optimization procedure. In which the strong line represents the mean,μ, of the best fitness evaluation at each iteration, and the thin line the standard deviation,σ, of the best fitness evaluation at each position. Particularly, the interval between $[\mu - \sigma, \mu + \sigma]$. It can be seen that 90% of the mean best fitness evaluation, i.e. 0.9μ = 945, is reached in 20 position. Also, the mean best fitness evaluation at the last iteration (300) reaches a value of 1300 that represents 35% relatively less than the theoretical best fitness evaluation of 2000 (assuming that T=2000 and the average CTR is 1.0). This behaviour reflects that the proposed method based on particle swarm optimization is well configured and implemented. On the other hand, the gbest of each repetition was stored. The best particle at each repetition sorted by the fitness function value.

This table shows: some of the experiment, the iteration of the SPO when the best fitness value was found, the number of features in the configuration, the average CTR of the subset selected, the real size of the whole subset before using the threshold value, the fitness function value, and the features selected. The number of features is 9, the average CTR value is 0.72, and the fitness function value is 1447.94. The mean frequency of the features is 19.90, then features above it, i.e. fi for all i ={5, 6, 7, 8, 10, 11, 12, 13, 16, 19, 20, 23}, were the most frequent ones considered by the P= 50. These features are important since these configurations provide high evaluation which implies that the average CTR is larger as well as the size of the campaign. In addition, it is evident that features i = {14, 15} were not considered since those have high variance in their values generating too many small groups of ad campaigns. The latter impacts on the fitness evaluation, resulting in their lack of selection. To this end, the selected features suggest that the optimal advertising campaign should be configured with these requirements.

V. *Conclusion*

In this article, we present a cognitive optimization for display advertisement campaigns through the behavioural-targeting and retargeting technique in direct response display advertising campaigns as an ensemble of Particle Swarm Optimization and an online logistic regression CTR model. We consider the methodology to be very useful for small advertising networks because they can optimize advertiser's campaigns effectively using few resources. In this way, small networks may be more competitive and may have more satisfied advertisers. Our methodology can be applied very easily to real-time bidding that is a huge auction involving many ad networks. This improvement can make online advertisement a much more attractive system to advertisers. Supervised ML models enable predicting CTR, in such a way that it is not any longer necessary that advertisers invest in campaigns before getting an optimal configuration. Since it is possible to simulate users' behaviour artificially, better suggestions could be made to advertisers in less time and at little expense, which may increase their satisfaction degree.

REFERENCES

[1]. Aksakalli, V., 2012. Optimizing direct response in internet display advertising. *Electronic Commerce Research and Applications* 11 (3), 229-240.

[2]. Aronowich, M., Benis, A. J., Yanai, R., Vind, G., Jun. 25 2014. *Budget distribution in online advertising. US Patent App.* 14/314,151.

[3]. Balseiro, S. R., Feldman, J., Mirrokni, V., Muthukrishnan, S., 2014. Yield optimization of display advertising with ad exchange. *Management* Science (12), 2886-2907.

[4]. Chapelle, O., Manavoglu, E., Rosales, R., 2015. Simple and scalable response prediction for display advertising. *ACM Transactions on Intelligent Systems and Technology (TIST)* 5 (4),

[5]. Chen, G., Cox, J. H., Uluagac, A. S., Copeland, J. A., 2016. In-depth survey of digital advertising technologies. *IEEE Communications Surveys & Tutorials* 18 (3), 2124-2148.

[6]. Chin, W.-S., Zhuang, Y., Juan, Y.-C., Lin, C.-J., 2015. A learning-rate schedule for stochastic gradient methods to matrix factorization. In*: Pacific-Asia Conference on Knowledge Discovery and Data Mining. Springer*, pp. 442-455

[7]. Danaher, P., Lee, J., Kerbache, L., 2010. Optimal internet media selection. *Marketing Science* 29 (4), 336 – 347.

[8]. Edelman, B., Ostrovsky, M., 2007. Strategic bidder behavior in sponsored search auctions. *Decision support systems* 43 (1), 192-198

[9]. Evans, D. S., 2008. The economics of the online advertising industry. *Review of network economics* 7 (3). Evans, D. S., 2009. The online advertising industry: Economics, evolution, and privacy. *The journal of economic perspectives* 23 (3), 37-60.

[10]. Evans, D. S., 2008. The economics of the online advertising industry. *Review of network economics* 7 (3).

[11]. Evans, D. S., 2009. The online advertising industry: Economics, evolution, and privacy. *The journal of economic perspectives* 23 (3), 37-60.

[12]. Goldfarb, A., 2014. What is different about online advertising? *Review of Industrial Organization* 44 (2), 115-129.

[13]. Goldfarb, A., Tucker, C., 2011. Online display advertising: Targeting and obtrusiveness. Marketing Science 30 (3), 389-404. IAB, Jun. 25 2016. Interactive advertising bureau. Inform of IAB.

[14]. Jin Qin, Yi-xin Yin and Xiao-juan Ban, 2011.Hybrid Discrete Particle Swarm Algorithm for Graph Coloring Problem, *journal of computers*, vol 6, doi:10.4304/jcp.6.6.1175-1182.

[15]. Lee, Kuang-chih, Orten, B., Dasdan, A., Wentong, L., 2012. Estimating conversion rate in display advertising from past performance data. In: *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, pp. 768-776.

[16]. Lee, K.-C., Jalali, A., Dasdan, A., 2013. Real time bid optimization with smooth budget delivery in online advertising. In: *Proceedings of the Seventh International Workshop on Data Mining for Online Advertising*. ACM, p. 1.

[17]. Li, P., Shrivastava, A., Moore, J. L., Konig, A. C., 2011. Hashing algorithms for large-scale learning. In: *Advances in neural information processing systems*. pp. 2672-2680.

[18]. Liu, Y., Pandey, S., Agarwal, D., Josifovski, V., 2012. Finding the right consumer: optimizing for conversion in display advertising campaigns. In: *Proceedings of the fifth ACM international conference on Web search and data mining*. ACM, pp. 473-482.

[19]. Lucier, B., Paes Leme, R., Tardos, E., 2012. On revenue in the generalized second price auction. In: *Proceedings of the 21st international conference on World Wide Web*. ACM, pp. 361-370.

[20]. McMahan, H. B., Holt, G., Sculley, D., Young, M., Ebner, D., Grady, J., Nie, L., Phillips, T., Davydov, E., Golovin, D., et al., 2013. Ad click prediction: a view from the trenches. In: *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining.* ACM, pp. 1222-1230.

[21]. Miralles-Pechuan, L., Rosso, D., Jimenez, F., Garcia, J. M., 2017. A methodology based on deep learning for advert value calculation in cpm, cpc and CPA networks. *Soft Computing* 21 (3), 651-665.

[22]. Muthukrishnan, S., Pfial, M., Svitkina, Z., 2007. Stochastic models for budget optimization in search-based advertising. In: *International Workshop on Web and Internet Economics. Springer*, pp. 131-142.

[23]. Omar Ilaya, Cees Bil (2009). A Particle Swarm Optimisation Approach to Grap Permutations, Particle Swarm Optimization, *Aleksandar Lazinica (Ed.),* ISBN: 978-953-7619-48-0

[24]. Pandey, S., Dutta, G., Joshi, H., 2017. Survey on revenue management in media and broadcasting. *Interfaces* 47 (3), 195 - 213.

[25]. Perlich, C., Dalessandro, B., Hook, R., Stitelman, O., Raeder, T., Provost, F., 2012. Bid optimizing and inventory scoring in targeted online advertising. In: Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, pp. 804-812.

[26]. Provost, F., Dalessandro, B., Hook, R., Zhang, X., Murray, A., 2009. Audience selection for on-line brand advertising: privacy-friendly social network targeting. In: *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining.* ACM, pp. 707-716.

[27]. Richardson, M., Dominowska, E., Ragno, R., 2007. Predicting clicks: estimating the click-through rate for new ads. In: *Proceedings of the 16th international conference on World Wide Web*. ACM, pp. 521-530.

[28]. Shan, L., Lin, L., Sun, C., Wang, X., 2016. Predicting ad click-through rates via feature-based fully coupled interaction tensor factorization. *Electronic Commerce Research and Applications* 16 (2016), 30-42.

[29]. Vincent, O.R, O. Folorunso, A. D. Akinde, 2009. Agent-Based Advert Placement System for Broadcasting Stations. *Issues in Informing Science and Information Technology Volume* 6 DOI: 10.28945/1067

[30]. Xinchao. X, 2010. A perturbed particle swarm algorithm for numerical optimization. *Applied Soft Computing*, 10:119-124, 2010.

[31]. Yudong Zhang, Shuihua Wang, and Genlin Ji, 2015. A Comprehensive Survey on Particle Swarm Optimization Algorithm and Its Applications, *Mathematical Problems in Engineering* Volume 2015, Article ID 931256, doi.org/10.1155/2015/931256