OptiMine®
BID OPTIMIZATION SOFTWARE

MORE MATH.
BETTER MATH.
BEST ROI.

ACHIEVING THE GOLD STANDARD IN PAID-SEARCH BID OPTIMIZATION

UNDERSTANDING THE BENEFITS OF GLOBAL KEYWORD-LEVEL MODELING

WHITE PAPER: THE GOLD STANDARD

INFO@OPTIMINE.COM • WWW.OPTIMINE.COM

"Achieving the Gold Standard in Paid-Search Bid Optimization" is a detailed look at the four primary methods of optimizing keyword bidding: rules-based optimization and three different approaches to model-based optimization (global keyword-level modeling, global cluster-level modeling, and local optimization). This page provides a brief summary of each method presented in order from highest to lowest performance. Read on for a more detailed look at each of these different optimization approaches.

MODEL-BASED OPTIMIZATION

Global optimization with keyword-level modeling: The customer selects the financial goal to maximize, e.g., revenue, profit or return on ad spend. Every keyword is individually modeled using the customer's cost and conversion history, an array of models, and a large number of variables to identify the specific models and variables that best predict performance for each keyword. The bid price for a given keyword is set to best maximize the desired global financial outcome. The models are re-trained frequently with the newest cost and conversion data, and the models and variables best predicting keyword performance might change accordingly. Even those keywords with low conversion rates are individually modeled and each keyword is treated as its own unique market, all aimed at maximizing the overall paid-search financial performance. This approach must be fully automated, given the millions of computations to determine the best treatment for each keyword. The result is a fully transparent process, individually modeled keywords, and excellent global financial performance.

Global cluster-level modeling: Global optimization is the overall goal, but keywords with lower conversion rates are clustered to create sufficient data for modeling. Net, a "hybrid" approach is typically used with the keyword-level global approach used for head terms and clustering used for tail terms. Global cluster-level bidding tends to be stable, but what is gained in stability is lost in performance and automation. Performance suffers as clusters ignore the fact that every keyword is unique and the value of the aggregated data is outweighed by the loss of that uniqueness. A variety of factors such as seasonal changes, expanded keyword lists, and changes in product offerings can render clusters obsolete. When obsolescence occurs, statisticians are typically needed to tune the models manually, so cluster-based solutions are rarely pure software applications with the loss of automation and transparency.

Local optimization: Local optimization simply sets bids so that each keyword meets the selected financial goal. For example, if the target ROAS is 200 percent, then each keyword is bid to obtain an individual ROAS of 200 percent. There is no trade-off between keywords. In another application, the bid price of each keyword is set based on the historical average value. Again, no trade-off or distinction is made between keywords. Either approach is simple but performance is lower as keywords are not assessed based on their relative impact on maximizing overall performance.

RULES-BASED OPTIMIZATION

Rules-based optimization is touted for being simple and easy to understand. For example, a rule might state, "If ROAS is less than 200 percent, then lower bids by 10 percent." However, if rules are layered upon rules, simplicity and understandability are quickly negated, making it difficult to understand what will happen to the bids. The significant issue with rules-based optimization is performance. Rules-based systems are reactive, with pre-defined responses to certain situations. In a rules-based system, historical data is not considered because the situation drives the reaction. Because of their reactive nature, rules-based optimization systems can be very good at protecting position, but playing defense rarely leads to optimal results. In general, models-based systems are predictive, while rules-based systems are reactive.

INTRODUCTION

Paid-search advertisers have a range of choices when selecting a keyword-bidding approach: rules-based bidding and three different versions of model-based solutions. This paper will analyze the differences in these bidding techniques and explain why global keyword-level bid optimization delivers performance gains of 25 percent to 200 percent more than what can be achieved with the other bidding approaches.

PAID SEARCH OVERVIEW

To maximize paid-search success, it's necessary to balance several elements. The most critical for large paid-search campaigns is optimizing and placing bids on thousands or millions of keywords every day. Advancing technology has resulted in the development of tools to help advertisers optimize daily keyword bids, but the methods employed and performance improvements vary.

Google's advertising model is both simple and fiendishly complex. In most cases, paid-search advertisers bid on keywords daily to locate their text ads in the most advantageous positions on Google's search engine. The bid, combined with Google's Quality Score—a formula Google has developed that factors in an ad's quality, relevance and historical click-through performance—determines an ad's position.[1]

In effect, advertisers have to predict the future cost and value of every single keyword in their search engine marketing campaigns. It's not a trivial task. Large paid-search advertisers may need to optimize tens of thousands to millions of keywords to meet a variety of different campaign objectives. Not all keywords have the same weight, of course. Head terms, or keywords that trigger a large number of clicks and conversions, are usually the focus of most SEM efforts. Tail terms, or keywords that might receive just a few clicks per day or week, receive less attention. But regardless of where a keyword lies on the distribution, if its bid is too low, sales are lost. If the bid is too high, more money is spent on that keyword than its potential return or value merits. Achieving the best performance takes automation, science and domain knowledge.

Over the past few years, search engine management consultants and digital marketing platform vendors have developed a variety of products and services to help advertisers optimize their paid-search spend. In most cases, the efforts involve some combination of human analysis and model-based bid optimization software, sometimes called an auto-bidder. Technology is essential because the scale of the challenge easily outstrips human analysis alone, especially when it comes to tail terms, which have little click history and which make up the vast majority of keywords.

A common perception among paid-search advertisers is that model-based bid optimization software is something of a black-box technology that is only appropriate for tail terms and that advertisers should "manually optimize the important head terms of an account."[2] This perception is based on some of the early solutions in the marketplace, which did indeed keep their inner workings secret and typically delivered only modest short-term performance gains that tended to taper off over time.

Today, however, bid optimization technology has developed to the point that it can maximize financial performance across all keywords while being

[1] Microsoft adCenter uses a similar system.

[2] Josh Dreller, searchengineland.com, September 30, 2011. The Technology Behind Autobidding: Q&A With OptiMine's Dr. Rob Cooley

transparent and meeting critical business constraints – for example, maximizing profit while providing some minimum order volume, or maximizing revenue while maintaining a minimum return on ad spend (ROAS). For advertisers managing a large number of keywords, some form of model-based bid optimization technique is virtually essential. But success in using bid optimization software requires understanding the advantages and limitations of the different model-based techniques and the recognition that there are no set-and-forget solutions. And bid optimization still requires some degree of human intervention from analysts with domain knowledge and an understanding of the business.

BID OPTIMIZATION BASICS

For large, complex campaigns, model-based optimization is superior to rules-based in every way, but it's not a simple choice between A and B. Of the three model-based methods, only one – global keyword-level modeling – delivers consistently superior performance improvement.

Rules vs. Models

The benefits of modeling exceed those of rules-based optimization, which reacts to situations rather than predicting and adjusting by learning historical conversion data.

As shown in Figure 1, there are two main approaches to bid optimization: rule-based versus model-based. A model-based system uses historical performance data to train statistical models to predict future performance. For example, a model-based system could predict the bids necessary to achieve a 200 percent ROAS tomorrow. They contrast with rules-based systems, which use a pre-defined set of reactions to certain situations. For example "if ROAS is less than 200 percent, then lower bids by 10 percent." In general, models-based systems are predictive while rules-based systems are reactive.

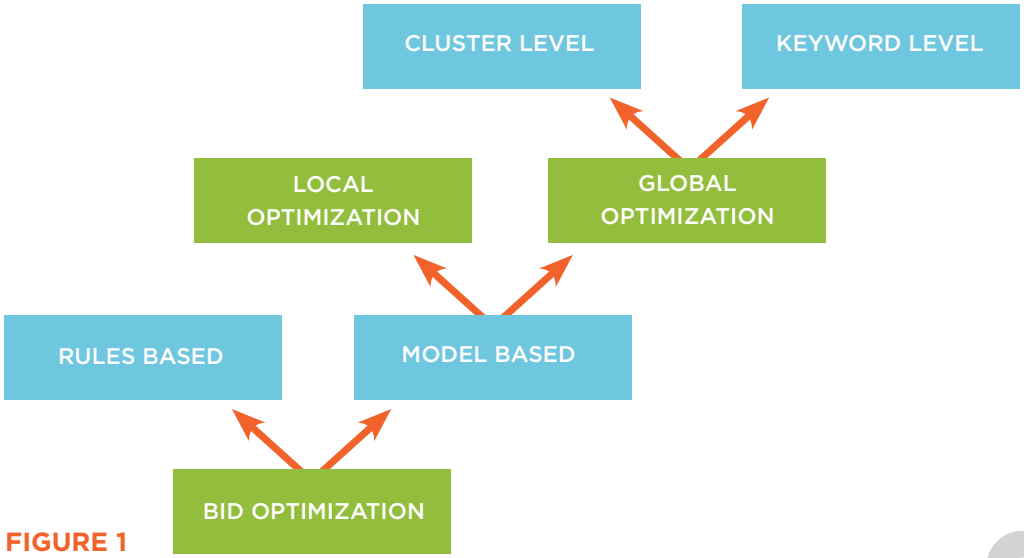

FIGURE 1

Global Optimization vs. Local Optimization

Global optimization strives for the success of the whole campaign, while local optimization makes sure every keyword achieves the same goal, regardless of the overall impact.

Within model-based approaches, we need to make a distinction between local and global optimization. (See Figure 2). Local optimization simply bids each keyword separately from the other keywords in an SEM program. For example, if your target ROAS is 200 percent, then every keyword is bid to obtain an individual ROAS of 200 percent. A local solution won't trade off a low ROAS on one keyword with a high ROAS from another.

Global optimization (referred to as a portfolio approach by some vendors) considers all of the keywords at once, assigning bids so that, on average, the group as a whole maximizes a goal while meeting certain constraints, such as a specific ROAS or minimizing cost per action (CPA). The advantage of global optimization is that it treats each keyword appropriately with respect to the whole campaign. For example, it might turn out that one keyword can drive significant revenue at a ROAS of 180 percent, while another drives the same amount of revenue at a ROAS of 220 percent. As long as the average ROAS is 200 percent, the global solution will declare success. This approach generally provides higher value from a set of keywords than local optimization.

		KEYWORD B	
		Low Bid Cost: $120 Orders: 2 Cost per Order: $60	**High Bid Cost:** $350 Orders: 5 Cost per Order: $70
KEYWORD A	**Low Bid Cost:** $300 Orders: 10 Cost per Order: $30	Total Cost: $420 Orders: 12 Cost per Order: $35	Total Cost: $650 Orders: 15 Cost per Order: $43.33
	High Bid Cost: $660 Orders: 11 Cost per Order: $60	Total Cost: $780 Orders: 13 Cost per Order: $60	Total Cost: $1010 Orders: 16 Cost per Order: $63.13

Goal:
Maximizing orders

Orders Constraint:
$60 maximum cost per order

Local Optimization Best Solution

Global Optimization Best Solution

FIGURE 2: In this simple example there are **two keywords (A & B)** that only have **two possible bids each (High & Low)**, for a total of four possible bidding scenarios. **The goal** is maximizing orders; **the constraint** is a $60 maximum cost per order. With **local optimization**, all keywords must meet the constraint. With **global optimization**, all keywords are considered at once and bid individually, leading to a substantially lower cost per order and a dramatically higher ROAS.

Global Keyword vs. Global Cluster Modeling

Global keyword optimization results in improvements that are significantly better than cluster modeling because each keyword is treated individually. Cluster modeling creates groups of keywords, treating each member the same as the others – a hazardous shortcut for advertisers with large, complex paid-search campaigns.

Within model-based global optimization approaches there are two distinct levels of modeling: keyword-level modeling and cluster-based modeling.

Cluster model-based systems were developed to cope with the "sparse data" problem of tail terms for which there is very little historical data available. As the name implies, clustering aggregates data from several keywords to hundreds and even thousands of keywords, which are assumed to have the same general performance characteristics. Simply put, clustering was developed to assemble data sets large enough to permit human analysts to apply more traditional statistical techniques to determine bids.

While clustering may seem like a necessary way to manage keywords that receive very few clicks or impressions, the technique is in many ways sub-optimal. On the plus side, using clusters leads to model stability, meaning the results are repeatable, which is something statisticians love. But the negative effects of using clustering are substantial. For one, keywords in a cluster might seem similar, but each keyword is actually unique at some level. While statisticians might value the extra data, individual keywords lose their uniqueness in a cluster. The result is a loss of performance compared to modeling each keyword as a unique entity (see Figure 3), as the advertiser spends too much on some keywords in a cluster and too little on others.

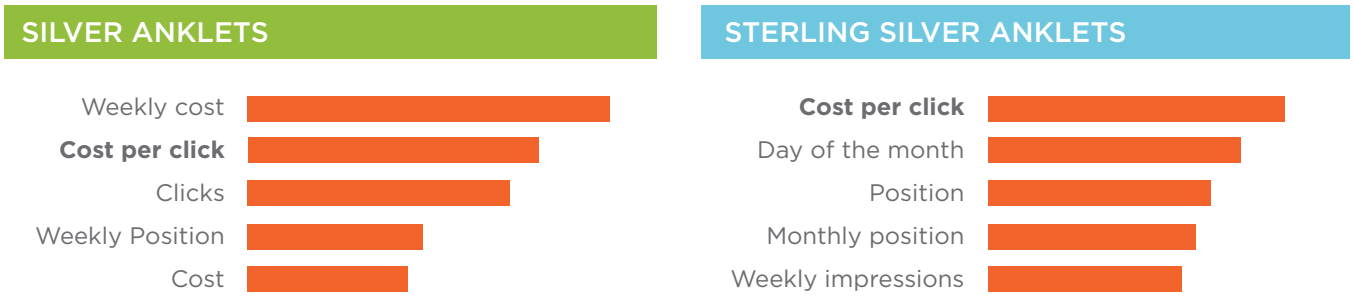

FIGURE 3: Clustering is inherently sub-optimal since each keyword loses its uniqueness. The figure above shows two similar, "live" customer keywords and the variables affecting performance as determined by keyword-level modeling. Note that only one of the top variables is common between the two keywords. If these two keywords had been clustered and forced to use a common set of variables, performance for one or both keywords would be reduced. A second problem with clustering is that it cannot be automated and requires regular human attention to maintain performance. Often customers will achieve an initial improvement in bidding performance followed by declines as the clustering goes out of date. In a "complex" bidding situation – e.g., significant numbers of new keywords, changing campaigns, significant seasonality, etc. – clustering is even more difficult to keep current.

A second issue is clustering typically requires statisticians to tune the models manually. As a result, cluster-based solutions are rarely pure software applications and thus can't offer the degree of automated bid optimization needed by advertisers managing thousands or millions of keywords in a dynamic bid-based marketplace. A third issue is the time-intensive task of creating clusters, as all statistical models decay over time, making it necessary to periodically retrain the statistical models involved. For these two reasons, cluster models tend to be used far beyond their useful life cycle with a resulting negative impact on performance.

Clustering survives due to the belief that keyword-level modeling, while far superior in performance terms, can't be done on tail terms for which there simply isn't enough data available to build accurate models. That's why most paid-search bid optimization vendors extensively rely on clustering, even though they might not reveal this fact to customers. Why? Clustering has yet to be automated and requires human analysis driving up cost and response times, even as it leads to sup-optimal performance on the vast majority of keywords in a campaign.

Fortunately for advertisers, this belief is false. Keyword-level modeling can be done on keywords with as few as 10 conversions a year. What's required is the right balance of the appropriate math, software automation, and transparency into the specific variables that drive individual keyword performance. **Global keyword-level modeling** is the gold standard of bid management. It regularly improves performance (profit, revenue, ROAS, etc.) by 25 percent or more in controlled tests against global cluster-level, local keyword-level and rules-based technologies. The basic elements driving success for global keyword-level modeling can be summarized as follows.

Appropriate math: The mathematical approach to global keyword-level modeling rejects the "requirement" to cluster in order to create sufficient click or impression data. Clustering might produce an acceptable "average" for all keywords in the cluster but has little relevance to the future performance of the individual keywords the cluster contains. This accounts for the sub-optimal performance of clustering versus modeling the specific performance of the individual keywords within the cluster, which is achievable for even keywords with small amounts of data. Achievable, that is, using the appropriate mathematical modeling techniques – such as structural risk minimization, a technique that trains models to become simpler as data sets become sparser – as opposed to complex clustering techniques, which deal with sparse data by building models that are often far more elaborate than the data will support.

In one controlled test, using the right math to individually model each keyword drove 216 percent more account sign-ups than a competing cluster-level solution. One issue in this test was the age of clusters used, which hadn't been refreshed and contained keyword groupings that were simply obsolete. Keyword-level modeling techniques identified a handful of good keywords hidden in clusters of bad keywords. Separating those out and bidding them up led to the volume increase.

Since keyword-level modeling eliminates the need to cluster, it can be achieved through software automation. For advertisers dealing with large numbers of keywords who want to avoid the performance sacrifice in long-tail keywords that is inevitable with clustering, automation is usually an advantage in a dynamic advertising marketplace, both in terms of lower cost and superior daily bid optimization of all keywords in an SEM program. The same type of software automation techniques can be applied to retraining the models that predict keyword performance. The models can be "taught" to automatically learn based on the most recent inputs that make them much more responsive to changing marketplace conditions.

Transparency: Transparency is the ability to make visible to the advertiser the individual variables that drive keyword performance to build understanding and trust. In this sense, clusters are about as transparent as mud, since the complexity of clusters makes them almost impossible to describe. Modeling keywords individually allows advertisers to make intelligent determinations on each and every keyword in an SEM campaign. They can see the decisions the software made and why. Figure 3 compared five variables that drove performance in two similar but distinct keywords. Using clustering, marketers would be forced to conclude that only one variable – cost per click – was the common and determinant performance factor, whereas in reality, the two keywords have very dissimilar performance profiles and require very different optimization strategies.

BOTTOM LINE

Model-based keyword bid optimization is by far superior to rules-based, but selecting the right model-based system is critical to maximizing paid-search campaign performance. Where local optimization is incrementally better than rules-based, the performance improvement does not approach that of global cluster or global keyword. But only global keyword-level modeling, the gold standard of bid optimization, will truly maximize paid-search performance.

Keyword-level modeling for all keywords in an SEM program, not just the low-hanging fruit of head terms, is a demonstrably superior approach and delivers overall performance gains of 25 percent or more over clustering and other techniques. Software-driven bid optimization techniques that use keyword-level modeling also provide advertisers with greater flexibility and control. Clusters take time to create and rebuild in response to changing business goals and marketplace conditions. With automated keyword-level bid optimization techniques, advertisers can continually change campaign goals and constraints and rebid keywords at any frequency they need.

Keyword-level, global bid optimization is the ultimate solution for deriving maximum profit from SEM campaigns that involve a large number of keywords. It's the right choice for optimizing head terms, tail terms and everything in between.

Robert Cooley, PhD, is chief technology officer for OptiMine Software. OptiMine's sophisticated bid optimization software forecasts the performance of each paid-search ad placement each day and automatically sets optimal bids. OptiMine's unmatched predictive analytics software trains up to 24 mathematical models per keyword each week and examines more than 100 variables to automatically maximize keyword performance daily.

For more information, visit www.optimine.com.