An Adaptive Sponsored Search Mechanism δ-Gain Truthful in Valuation, Time, and Budget

Rica Gonen*and Elan Pavlov†

Abstract

This paper presents an online sponsored search auction that motivates advertisers to report their true budget, arrival time, departure time, and value per click. The auction is based on a modified *Multi-Armed Bandit (MAB)* mechanism that allows for advertisers who arrive and depart in an online fashion, have a value per click, and are budget constrained.

In tackling the problem of truthful budget, arrival and departure times, it turns out that it is not possible to achieve truthfulness in the classical sense (which we show in a companion paper). As such, we define a new concept called *δ-gain*. δ-gain bounds the utility a player can gain by lying as opposed to his utility when telling the truth. Building on the δ-gain concept we define another new concept called *relative ϵ-gain*, which bounds the relative ratio of the gain a player can achieve by lying with respect to his true utility. We argue that for many practical applications if the δ-gain and or the relative ϵ-gain are small, then players will not invest time and effort in making strategic choices but will truthtell as a default strategy. These concepts capture the essence of dominant strategy mechanisms as they lead the advertiser to choose truthtelling over other strategies.

In order to achieve δ-gain truthful mechanism this paper also presents a new payment scheme, Time series Truthful Payment Scheme (TTPS), for an online budget-constrained auction mechanism. The payment scheme is a generalization of the VCG principles for an online scheduling environment with budgeted players.

Using the concepts of δ-gain truthful we present the only known budget-constrained sponsored search auction with truthful guarantees on budget, arrivals, departures, and valuations. Previous works that deal with advertiser budgets only deal with the non-strategic case.

1 Introduction

With the advent of advertising as a pillar [9] of Internet commerce, there is an acute need for improved means of increasing the value achieved by advertising agencies. In the increasingly competitive and high stakes duel between the main advertising search engines (Google, Microsoft and Yahoo!) every bit of advantage is important.

In this competition mechanism design is an important part of optimizing the monetization of search advertising. Mechanism design allows us to define allocations and payments that maximize the welfare of participants. In doing so search engines can attract advertisers who have a strong interest (high valuation) in users interacting with their ad placements.

1.1 Problem Setting Considerations

The main tool that a mechanism designer can bring to the table is *preference elicitation* which essentially means finding incentives (via payment rules) that motivate the participants to honestly report their valuations for any possible allocation.

Indeed, in assuming that advertisers have a known valuation per click as well as a bounded budget, many authors have suggested algorithms that increase welfare for the search engine e.g., [1]. Some authors have even suggested mechanisms which do not assume the knowledge of CTRs but learn them while running the algorithm [19].

However we argue that the assumption of known valuations is unrealistic. In practice advertisers' values are *private information* and hence advertisers might be motivated to act strategically to increase their utility. In [13] we suggest a truthful multi-armed bandit (MAB) mechanism for the case where advertisers have no budget and are always available to show an ad. Further-

*Yahoo! Research Labs, 701 First Street, Sunnyvale, CA 94089. Email: gonenr@yahoo-inc.com We are submitting this paper for confidential review,

†E15-313 Media Lab, MIT, Cambridge MA, 02149. Email: elan@mit.edu

more, [13] allows the different slots (possible places to display an ad) to be of different quality (although the slot quality ratio is unrelated to the advertiser).

In this paper we make the restricting assumption that the slots are of equal quality. This assumption is not necessary other than to manage the complexity of the algorithm's presentation and allows us to express the essential elements of the model where advertisers have budget constraints as well as time constraints. We believe that this scenario captures the core nature of advertising. For example, advertisers commonly value a click through more highly in the pre-Christmas gift season than during the rest of the year.

Our follow-on paper [14] creates an auction that is truthful in budget, arrival, departure, and valuation while recognizing that slots are not of equal quality.

The budget constraint is harder to justify theoretically, inasmuch as what is important is the marginal utility from additional clicks. However, budgets decrease advertiser risk and are a standard assumption in the theory as well as a standard assumption in practice. Advertiser budgets cause a theoretical difficulty in that it is well known [16, 8] that it is impossible to maximize welfare given the existence of budget constraints. Even when advertisers are time constrained in addition to being budget constrained [12]. Hence, we define our approximation relative to the optimal allocation under budget constraints.

Four parameters are assigned to each participant; arrival time, departure time, value per click, and budget. These parameters are private information that must be reported to the mechanism. Our results take another step toward capturing reality by allowing advertisers multiple arrivals and departures to and from the system.

Some of these parameters pose special challenges. For instance the arrival and departure times pose a challenge as [18] showed in a different context that it is impossible to achieve a truthful scheduling mechanism. [18]'s scheduling problem can hint that in our setting, there exists an impossibility of scheduling advertisers who desire a single impression with a value per click that is identical to their budget.

Fortunately, in our setting, it is quite sensible to assume, as in [17], that the budget is much larger than the value for a single click. In [12] we extended [18]'s impossibility to apply to budget-constrained sponsored search auctions and to apply even when the above assumption is made. Nevertheless assuming that the budget is much larger than the value per click allows us, together with the new payment scheme (TTPS), to bound the number of "free" clicks a player can receive by lying.

In order to formalize this intuition we define two new concepts which we call *δ-gain* and *relative ϵ-gain*. δ-gain is a bound on the utility *gain* that a player can achieve by lying independent of the lie's size. In contrast to prior work which bounded the size of a lie, δ-gain allows lies to be arbitrarily large. We argue that if the maximum utility gain from a lie is small (in our case $O(value)$) then players will forgo this risky gain in favor of the simplicity of truthtelling. Our mechanism has the additional property that this gain can only be achieved at the risk of hazarding the entire budget. To express this we define the relative ϵ-gain concept which not only ensures a small additive gain in utility but also ensures that the relative gain is small with respect to the total utility achieved when acting truthfully. We believe that the above definitions are of independent interest and can be applied to a range of other mechanism design problems.

We assume throughout this paper that an advertiser has zero value if his ad is not clicked on[1]. Under this assumption the value to the advertiser is the *value per click* $\times$ CTR.

1.2 Selection of Underlying Bandit Algorithm

Since the advertising search engine is assumed to be interested in maximizing welfare, which depends on the CTR as well as on the private values of the advertisers, it is natural to cast the problem as a multi armed bandit.

The multi-armed bandit is a well-studied problem [20, 3, 4] which deals with the balancing of exploration and exploitation in online problems with multiple possible solutions. In the simplest version of the MAB problem a user must choose at each stage (the number of stages is known in advance) a single bandit/arm. This bandit will yield a reward which depends on some *hidden* distribution. The user must then choose whether to exploit the currently best known distribution or to attempt to gather more information on a distribution that currently appears suboptimal. The MAB is known to be solvable via the Gittins [11] index and there are solutions which approximate the optimal expected payoff. We choose to generalize the MAB solution in [10] dues to its simplicity and optimal sampling complexity. Our solution retains the sample complexity of [10] (what we call the suboptimal exposure complexity) and

[1]In practice some advertisers are interested in rasing visibility.

hence is sample complexity optimal.

Since we want to elicit the private information of the advertisers we must design a MAB which is truthful. Models of imperfect and symmetric information for prices have been extensively studied recently, e.g., [5, 6]. The MAB has been recently studied in a more general setting by [7] but using a weaker notion of truthfulness.

Although the MAB has been extensively studied it has generally been studied in the context of a single user choosing from non-strategic arms [15] even when studied in the context of slot auctions [19]. In [13] we constructed a truthful MAB under the assumption that the only private valuation was the value of an advertiser (although in contrast to the current work we allowed slots of different quality). The mechanism was called the **M***ulti-***A***rmed truth***F***ul band***I***t* **A***uction (MAFIA)* and hence the current mechanism is called **B***udgeted* **M***ulti-***A***rmed truth***F***ul band***I***t* **A***uction (B-MAFIA)*.

In the context of an online keyword auction the arms/advertisers will act as strategic utility maximizing agents who will be the slots for the MAB. Our mechanism will also achieve a good approximation of the optimal welfare (under budget constraints) and hence improves over current methods of sampling and/or heuristic-based modeling.

Organization: The paper is organized as follows: In section 2 we present our model, the bandit problem, and define our assumptions. We provide intuition by looking at the case of a single slot in section 3. Section 4 presents the B-MAFIA algorithm which its properties of δ-gain truthfulness, welfare approximation, and sub-optimal exposure complexity with minimal sampling loss; are analyzed in section 5. We conclude and discuss future extensions in section 6.

2 The Model

In our model a set N of risk neutral, utility-maximizing advertisers bid for advertising slots based on a keyword ($|N| = n$). This paper focuses on the bidding process for a single keyword, as multiple keywords are analogous in current mechanisms. It is therefore supposed w.l.o.g. that the keyword appears at every time t. Whenever that keyword appears in the search at time t, K_t[2] slots of advertisements appear in the results.

The advertisers arrive and depart the system in an online manner and may arrive and depart several times. S_t denotes the set of advertisers present in the system at time t. Each advertiser i has a private value for each click through (independent of the slot the ad originally appeared in) which is denoted by v_i. For every arrival and departure each advertiser i also has a privately known arrival and departure times, denoted a_i and l_i respectively, and a privately known budget denoted b_i.

The algorithm runs from time starting at $t = 1$ and ending at $t = T$. Each time period is called a round. During each round, the algorithm allocates advertisers to the K_t slots (or if there are too many slots to some portion of the slots). When advertiser i appears in a slot during some time period t we say that i received an impression. We denote the number of impressions (plus 1) that advertiser i receives from i's first arrival by e_i. We also denote the number of clicks that advertiser i received during his current stay in the system by ω_i.

In this paper it is assumed that the "quality" of each slot (which is essentially the probability of a click though if an advertisement appears in that slot) is identical in all K_t slots and is independent of the advertisers. In our paper [13] it is assumed that the "quality" of each slot is monotonically decreasing and is independent of the advertisers. In our working paper [14] we show that the common assumption of that the quality of a slot is independent of the advertiser can be relaxed.

Each advertiser i has a *click through rate* α_i which is the probability of a click on the advertisement given that the advertiser was allotted an impression. The value α_i is unknown to i as well as to the mechanism. Since α_i is unknown to i as well as to the mechanism, we must estimate it at each time t and denote the observed probability at time t by α_i^t. We denote the payoff of advertiser i by $x_i^t = v_i \cdot \alpha_i^t$

Finally, by $\overline{v}_i$ we denote the bid for each click-through stated by advertiser i to the mechanism. $\overline{a}_i, \overline{l}_i$, and $\overline{b}_i$ respectively denote the arrival time, the departure time, and the budget stated by advertiser i to the mechanism when advertiser i enters the system. (a_i and l_i may be reported multiple times with multiple entries).

To achieve this paper's main claim of a truthful report of budget, time, and value, we make the natural assumption that advertiser i's reported budget, b_i is significantly larger than his value, i.e., $\overline{b}_i >> v_i$ for every advertiser. In practice this is indeed the case for the keywords auctions currently in use[3]. This assumption

[2]We assume for the ease of exposition that $K_t = K_{t+1} = K$ for all time period t. We also assume without the loss of generality that $K \leq |N| = n$, since superfluous slots can remain blank.

[3]Typical valuations for click through are several cents while the budgets for those click throughs are on the order of hundreds of dollars.

is commonly made even in non-strategic settings (e.g., [17]).

For ease of exposition $\overline{X}_i$ denotes the vector of parameters stated by advertiser i in a single arrival, i.e., $(\overline{a}_i, \overline{l}_i, \overline{v}_i, \overline{b}_i)$. The mechanism charges advertiser i a price denoted p^{l_i} every time i departs. At every round we charge advertiser i an "interim price' (This price can decrease as time goes on as well as increase.) $p_i^t \leq \omega_i \cdot \overline{v}_i$, where ω_i is the number of clicks i received during his current sojourn in the system, i.e., from $\overline{a}_i$ to time t. Since our advertisers are budget constrained we denote the budget i has remaining at time t as $B_i^t \geq 0$. It is assumed that advertisers have quasi-linear utility functions[4] and consequently at each departure time advertiser i obtains utility of $\alpha_i^{l_i} \cdot (\omega_i \cdot v_i - p^{l_i})$.

2.1 The Bandit Problem

A summery on the bandit problem model can be found in Appendix A

3 Illustration of the Protocol for the Single-Slot Case

We illustrate the main idea behind our protocol for the simple case where there is only a single slot available at any given time. In this case for each time period t we can look at the set of available advertisers S_t (note that since advertisers enter and depart the system this set might increase or decrease). For each advertiser $i \in S_t$ we have an estimation of i's click through rate α_i^t as well as an estimate of how accurate our estimation is, i.e., a bound on $|\alpha_i^t - \alpha_i|$ which depends on the number of impressions e_i that advertiser i received. We will denote this bound by γ_{e_i} (the definition of γ_{e_i} is elaborated on below).

Consider the set S_t. Naturally, this set has an i such that $v_i * \alpha_i^t$ is maximal.(In practice we have to ensure that there is sufficient remaining budget. Details appear in the technical part of the paper.) Suppose w.l.o.g. that the maximal element is the first element that our bandit algorithm explores (i.e., allocates a slot to). If the algorithm merely chooses to exploit then it could just allocate the slot to the first advertiser. However, there are other possible advertisers that are worthy of consideration. These are the advertisers j s.t. $v_i\alpha_i^t - \gamma_{e_i} < v_j\alpha_j^t + \gamma_{e_j}$ since the errors of i and j overlap. Therefore the algorithm allocates the slot to a random advertiser whose slot overlaps with the maximal element.

[4] As long as their budget constrained is maintained

This generalization of [10] works (in a PAC sense) if the advertisers are non-strategic (but arrive in an online fashion). Of course, if the advertisers are strategic we have to motivate them to give the correct values. If advertisers' arrival and departure times are public knowledge and advertisers are not budget constrained then one could set prices of allocated advertisers to be defined as the critical values at each time to receive the slot, (i.e., the minimum value advertiser i can report and still be allocated to the slot) and extract true reported values from the advertisers. However, since we do not assume that arrival and departure times are public knowledge and our advertisers are budgeted the incentive solution has to take a more subtle approach.

3.1 Incentive Problem Example

The following example illustrates the problem:

Example 1. *Assume a single ad slot in every page presented and 100 time increments to auction the one slot. Also assume two advertisers called i and j. Advertiser i is interested in all of the time rounds with a CTR of 1 and a value of 2. However, i's budget is limited to 25. During the first 50 rounds advertiser j arrives with value 0.5 and CTR 1.*

If advertiser i reported his true arrival time he will be allocated the first 50 rounds as his value is 2>0.5 and will be charged 0.5, the critical value. i will then exhaust his budget and be unable to compete over the next 50 rounds which he is interested in. If advertiser i lies about his arrival time and claims he is interested in only the last 50 rounds advertiser j will be given the first 50 time rounds for the price of 0[5] *and advertiser i will be given the last 50 time rounds for the price of 0. This will motivate i not to bid in the first 50 rounds and to lie about his arrival time.*

Since the mechanism is online, it is also possible that advertiser j does not appear in the system. If advertiser i anticipated js arrival and lied about his arrival time no one is given the slot in the first 50 time rounds and hence i suffers a diminished gain caused by not having bid on the first 50 slots, and the auctioneer suffers a diminished welfare.

This problem is quite general and can cause all of the advertisers to attempt to strategically report their

[5] or in practice for some small reserve price

arrival and departure times. Such advertisers' strategic behavior can result in a large loss of welfare and revenue.

Although we would like to ensure truthfulness despite the advertisers' ability to strategically manipulate arrival and departure times, it turns out that this is impossible ([18] and [12]). Our approach is to circumvent the impossibility of solving the problem by allowing the advertisers to gain from lying but bound the amount that advertisers gain as a small fraction of their utility; while the amount that they can potentially lose is their entire budget. We argue that advertisers will not pursue such small gains, especially as the risk of loss is great, and hence will act truthfully.

3.2 Pricing Scheme

Our pricing scheme works in two layers;the critical value per round and the critical values series per all rounds in a single duration.

First, for every advertiser the algorithm computes the critical value of each time round. The critical value is the minimum value that allows the advertiser to win in the computed round. If the auction were not an online auction the critical value of a round is sufficient for the pricing scheme to maintain truthfulness. However, our auction is an online auction and so the algorithm needs to consider all possible series of clicks that the advertiser could have potentially enjoyed in his current arrival. For instance if advertiser i gained 5 clicks during the time he was present in the system, the algorithm needs to consider all series of 5 time rounds where advertiser i could have potentially received a click.

The rounds that the algorithm needs to consider where i *could have* received a click are; the ones that i actually got a click (and obviously also got an impression) as well as the times that i did not receive an impression (but possibly could have).

In the rounds he got an impression but not a click we already know he couldn't get a click (since the user did not click on the ad). Therefore the second layer of the algorithm finds the critical series of clicks, meaning the minimum sum of values that could have resulted in the same number of clicks as the one the advertiser actually gained in his current arrival.

More formally the way that the critical series is computed follows:

For every advertiser i and any given time t there are three possibilities:

1. i did not gain an impression at time t. Denote the set of critical values at times such that i did not get an impression in duration $[\overline{a}_i, t]$ by W^i_{1t}
2. i had an impression but not a click through at time t. Denote the set of critical values at times such that i got an impression but not a click in duration $[\overline{a}_i, t]$ by W^i_{2t}
3. i had an impression and a click through at time t. Denote the set of critical values at times such that i got an impression and a click in duration $[\overline{a}_i, t]$ by W^i_{3t}

If an advertiser i had w_i click-throughs in duration $[\overline{a}_i, t]$ the algorithm sums the w_i lowest critical values from the sets $W^i_{1t} \cup W^i_{3t}$. We show that this payment scheme (TTPS) will yield the desired prices and show that these prices allow an advertiser i to gain at most $O(v_i)$ regardless of the lie's size and that this gain can only be achieved if i's budget is depleted.

Finally, the truthfulness of the reported budget and value using the above payment scheme (TTPS) is similarly argued. We now proceed to formally define our algorithm for the general case.

4 Budgeted Multi-Armed truthFul bandIt Auction(B-MAFIA)

This section presents our B-MAFIA algorithm. The main algorithm appears in table 1. Variable t states how long the protocol has been running. Since advertisers can arrive and depart from the protocol at any time, t is used only as a synchronization device. Each arriving advertiser's parameters are initialized in table 2. The possible candidates for sampling are chosen based on their distance from other advertisers as well as our level of confidence in the accuracy of their current CTR. The selection process is described in table 3. The sampled advertisers' parameters are updated in table 4. Finally, the algorithm charges departing advertisers in 5 and updates the time.

The next section proceeds by analyzing the algorithm's properties.

5 The B-MAFIA Analysis

This section analyzes the properties of the B-MAFIA algorithm. The properties we focus on are truthfulness and welfare approximation. We analyze the truthfulness for all of the reported parameters: arrival time, departure time, value per click through and budget.

The Main Algorithm
1. Call "Initialization Sub Procedure" from table 2 for any arriving advertisers 2. Call "Choosing the Advertisers to Sample" in table 3 to determine which subset of advertisers to consider sampling from. 3. Randomly pick K advertisers from S'_t to sample such that if advertiser $i \in S'_t$ is picked then if $i \in S'_\tau$ where $\tau \leq t$ there does not exist an advertiser $j \in S'_\tau$ and $j \in S'_t$ that was not sampled from time τ to time t. 4. Call "Update Parameters for Chosen Advertisers" in table 4 5. For every advertiser $i \in S_t$ (a) Compute the critical values c_i^t. (b) Compute interim price according to Compute-Price(input:for $\bar{a}_i \leq \tau \leq t$ c_i^τ, ω_i output:p_i^t) in table 5 and update remaining budget $B_i^t = \bar{b}_i - p_i^t$ (c) Charge departing advertisers their current interim price: if $\bar{l}_i = t$, i is charged $p_i^t = p^{l_i}$ 6. Update the time t.

Table 1: The main algorithm used for our budgeted multi-armed bandit sampling problem

Initialization Sub Procedure
1. The arriving advertiser reports his value per click $\bar{v}_i$, arrival time $\bar{a}_i$, departure time $\bar{l}_i$, and budget $\bar{b}_i$ that can be spent between $\bar{a}_i \leq t \leq \bar{l}_i$. Note that these reported values can be incorrect. 2. For each arriving advertiser, set the variables' values as follows: • Initial click through rate $x_i^t = 0$ • Total clicks $\omega_i = 0$ • Price charged for current clicks $p_i^t = 0$ • If this is i's first visit set $e_i = 1$.

Table 2: Initialization for each new advertiser

Due to lack of space we defer the B-MAFIA algorithm's analysis to the appendix. The truthfulness analysis can be found in appendix B.1 and the welfare approximation analysis can be found in appendix B.2.

6 Conclusions and Future Work

In this paper we presented a dominant δ-gain truthful multi-armed bandit mechanism approximating the maximum welfare that can be achieved under budget constraints. We achieve this approximation via elicitation of advertisers private information which consists of their multiple arrivals and departures in which they wish to advertise as well as their value per click and budget.

Our solution utilizes a novel definition of truthfulness we call δ-gain truthfulness and which allows us to bound the possible gain from *any possible* lie an advertiser might submit to the mechanism.

This paper's result also hold when advertisers bid for multiple keywords under the same budget constraints (even with different time constraints) as long as they have the same value for all keywords. In the full version we extend the treatment to deal with this case.

We leave open the important practical question of optimization, although we used [10] as our departure point, starting with a more efficient bandit mechanism should yield higher welfare. We also leave open the question of how to deal with players who bid for multiple keywords with different values for each keyword but with a total budget constraint.

Acknowledgements

The authors would like to thank Uri Feige, Chris Meek, and Seffi Naor for helpful comments.

References

[1] G. Aggarwal, A. Goel and R. Motwani. Truthful Auctions for Pricing Search Keywords. *Proceding of EC'06*

[2] A. Archer, C. Papadimitriou, K. Talwar, and E. Tardos. An approximate truthful mechanism for combinatorial auctions with single parameter agents. *In Proc. of the 14th SODA, 2003.*

Choosing the Advertisers to Sample
1. Set S_t to be the set of advertisers currently in the system (by their reported arrival and departure times).   2. Set $S'_t = S_t$ to be the set of potential advertisers to sample   3. For every advertiser $i \in S_t$, define confidence parameter: $\gamma_{e_i} = \sqrt{\frac{\log(cne_i{}^2/\chi)}{e_i} \cdot \frac{1}{K}}$ where c is a constant and χ is a probability parameter.   4. Remove from the set of advertisers S'_t all of the suboptimal advertisers: for every advertiser $i \in S_t$ if there exist K other advertisers z such that $\min\{x_z^t, B_z^t\} - \min\{x_i^t, B_i^t\} > \gamma_{e_i} + \gamma_{e_z}$ (i.e., z has remaining budget and payoff which is larger than i's even if the errors are in i's favor), then i is suboptimal and hence we can update $S'_t = S'_t \backslash i$.

Table 3: Choosing the Advertisers to Sample

Update Parameters for Chosen Advertisers
For every sampled advertiser $i \in S'_t$ update parameters: 1. $e_i = e_i + 1$ and update γ_{e_i} accordingly.   2. If i got a click update $\omega_i = \omega_i + 1$,   3. Update α_i^t (and x_i^t) accordingly if i was click/not clicked

Table 4: Updating Advertisers Who Were Sampled

COMPUTE-PRICE   (INPUT:FOR $a_i \leq \tau \leq t$ c_i^τ, ω_i OUTPUT:p_i^t)
$p_i^t = \arg\min\{\sum_{\omega_i} c_i^\tau\}$ such that $a_i \leq \tau \leq t$ and $c_i^\tau \in W_{1t}^i \cup W_{3t}^i$ (Choose the ω_i smallest critical values in $W_{1t}^i \cup W_{3t}^i$).

Table 5: Compute Prices (TTPS)

[3] D. A. Berry and B. Fristedt. Bandit problems. Sequential allocation of experiments. *Chapman and Hall 1985*

[4] D. Bergemann and J.Valimaki. Bandit Problems. *Social Science Research Network Electronic Paper Collection.*

[5] Bergemann, Dirk and Valimaki, Juuso "Learning and Strategic Pricing," emph Econometrica, Econometric Society, vol. 64(5), pages 1125-49, September. 1996.

[6] Dynamic price competition Dirk Bergemann and Juuso Valimaki *Journal of Economic Theory, Volume 127, issue 1,pp. 232-263* 2006

[7] Bergemann and Valimaki Efficient Auctions, Available at: $http://papers.ssrn.com/sol3/papers.cfm?abstract_id = 936633$

[8] C. Borgs, J. Chayes, N. Immorlica, M. Mahdian, and A. Saberi. Multi-unit auctions with budget-constrained bidders. *ACMConference on Electronic Commerce (EC 2005).*

[9] B. Edelman, M. Ostrovsky and M. Schwarz. Internet Advertising and the Generalized Second Price Auction: Selling Billions of Dollars Worth of Keywords. *Working paper 2005*

[10] E. Even-Dar, S. Manor, and Y. Mansour. PAC Bounds for Multi-Armed Bandit and Markov Decision Processes. *The Fifthteenth Annual Conference on Computational Learning Theory 2002*

[11] J.C. Gittins. Multi-armed Bandit Allocation Indices. *Wiley, New York. Mathematical Reviews: MR90e:62113 (1989)*

[12] R. Gonen. On the Hardness of Truthful Online Auctions with Multidimensional Constraints *Submitted 2007*

[13] R. Gonen and E. Pavlov. An Incentive-Compatible Multi Armed Bandit Mechanism. *Third Workshop on Sponsored Search Auctions WWW2007, PODC 2007*

[14] R. Gonen and E. Pavlov. An Incentive Compatible Budgete constraind sponsered search auction with monotonicaly decreasing slots. *Working paper 2007*

[15] R. Kleinberg. Anytime Algorithms for Multi-Armed Bandit Problems. *Proceedings of the 17th*

ACM-SIAM Symposium on Discrete Algorithms (SODA 2006).

[16] Jean-Jacques Laffont and Jacques Robert. Optimal auction with financially constrained buyers. *Economic Letters 52(1996) 181-186*

[17] A. Mehta, A. Saberi, U. Vazirani, and V. Vazirani. Adwords and Generalized Online Matching. *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science (FOCS): 264-273, IEEE Computer Society 2005..*

[18] R. Lavi and N. Nisan Online Ascending Auctions for Gradually Expiring Items *Proceedings of SODA'05.*

[19] S. Pandey, and C. Olston. Handling Advertisements of Unknown Quality in Search Advertising. *To appear in the proceedings of NIPS 06*

[20] H. Robbins. Some Aspects of the Sequential Design of Experiments. *In Bulletin of the American Mathematical Society, volume 55, pages 527535, 1952.*

A The Bandit Problem

The multi-armed bandit problem, originally described by Robbins [20], is a statistical decision model of an agent trying to optimize his decisions while improving his information at the same time. In the multi-arm bandit problem, the gambler has to decide which arm of K different slot machines to play in a sequence of trials so as to maximize his reward.

The bandit problem is best formulated as an infinite horizon Markov decision problem in discrete time with time index $t = 0, 1, ...$ At each time t the decision maker chooses amongst N arms and we denote this choice by $a_t \in \{1, ..., N\}$. If $a_t = i$, a random payoff x_t^i is realized and we denote the associated random variable by X_t^i. In our slot auction $x_t^i = \alpha_i^t \cdot v_i$ where the click through rate α_i^t is the random payoff element of the problem while the value v_i is a constant, hence the total payoff for arm i is $v_i \times \alpha_t^i$. The state variable of the Markovian decision problem is given by s_t where in our slot auction a vector of all allocated advertisers click-through-rate at time t, α_i^t and 0 if i is not allocated a slot in time t. The distribution of x_t^i is $F^i(\cdot; s_t)$. The state transition function ϕ depends on the choice of the arm and the realized payoff: $s_{t+1} = \phi(x_t^i; s_t)$. Let S_t denote the set of all possible states in period t. A feasible Markov policy $a = \{a_t\}_{t=0}^{\infty}$ selects an available alternative for each conceivable state s_t, i.e., $a_t : S_t \to \{1, ..., N\}$. Payoffs are evaluated according to the discounted expected payoff criterion where the discount factor χ satisfies $0 \le \chi < 1$. The motivation for assuming a discount factor is that the seller of the slot auction prefers payment sooner rather than later. The payoff from each i depends only on outcomes of periods with $a_t = i$. In other words, we can decompose the state variable s_t into N components $(s_t^1, ..., s_t^N)$ such that for all i: $s_{t+1}^i = s_t^i$ if $a_t \neq i$, $s_{t+1}^i = \phi(s_t^i, x_t)$ if $a_t = i$, and $F^i(\cdot, s_t) = F^i(\cdot; s_t^i)$.

Let $X^i(s_t^i)$ denote the random variable with distribution $F^i(\cdot; s_t^i)$. Then the problem of finding an optimal allocation policy is the solution to the following intertemporal optimization problem:

$$V(s_0) = \sup_a \{E \sum_{t=0}^{\infty} \chi^t X^{a_t}(s_t^{a_t})\}$$

where $V(s_0)$ is the value function of the bandit problem.

B The B-MAFIA Analysis

B.1 Truthfulness

Although the algorithm handles advertisers with multiple visits to the system, the truthfulness is proven assuming that every advertiser visits the system only once. It will be shown below that every visit of advertiser i can allow him a bounded gain of $O(v_i)$. (We also show that the gain relative to the true utility is $O(\frac{v_i}{b_i})$ which remains independent of the number of arrivals). Therefore in theory advertisers can increase their gain by artificially entering and exiting the system multiple times. One approach is to slightly modify the algorithm such that advertiser i reports his budget b_i once, for all visits to the system. As this is not the way the budget is handled in practice by search companies, we choose to have advertisers submit their budget every visit. Moreover in practice if an advertiser increases his number of visits he can lose utility both from the loss of possible advertising slots when his sub budget runs out as well as from increased costs for slots that the advertiser receives when the minimum critical values are taken independently for each visit.

Truthfulness is shown by analyzing a single visit of an advertiser to the system and by showing that *if* the algorithm correctly finds the best advertisers, i.e., that the multi-armed-bandit problem is correctly solved when advertisers are non-strategic, every advertiser i gains his maximum expected utility $\alpha_i^{l_i} \cdot (\omega_i \cdot v_i - p^{l_i})$

for every visit to the system if i reports his true values, $\overline{v}_i = v_i, \overline{a}_i = a_i, \overline{l}_i = l_i$, and $\overline{b}_i = b_i$. Since truthfulness holds for every advertiser i and for every visit, it follows that truthfulness holds over B-MAFIA as a whole.[6]

Since we show that for finite time the algorithm can only hope to succeed with some probability which is less than 1, we can not hope to show that the mechanism is *always* truthful. However, following [2] we show that the algorithm is truthful w.h.p. which depends on the parameters γ_{e_i}. Since γ_{e_i} has a constant c which we can use to fine-tune the tradeoff between expected success and expected loss of welfare we can ensure that the probability of success which we denote by $1-\lambda$ is arbitrarily high. This will then ensure that the probability of an advertiser to gain by lying is bounded by some arbitrary low constant we denote by θ.

To illustrate why the algorithm requires truthfulness with high probability and in expectation, consider the following scenario:

Example 2. *Suppose there are two advertisers with CTRs of $\alpha_1 > \alpha_2$ and $v_1 = v_2$. Furthermore, we will even assume that arrival and departure times are public information. Since the algorithm approximates the optimal welfare there exists some finite time t after which only one of the advertisers will be sampled. However, for any t there is a non-zero possibility (albeit arbitrarily small) that $\alpha_1^t < \alpha_2^t = \alpha_2$. In this case advertiser 1 has an incentive to lie such that $\bar{v}_1 \neq v_1$ and $\bar{v}_1\alpha_1^t > v_2\alpha_2^t$.*

However, the truthfulness proof insures that the kind of gain illustrated in the example above is bounded for the advertisers (i..e, the probability θ of the lie illustrated in the example helping the advertiser) if λ (which is a bound on θ) is sufficiently small. Since we can make λ arbitrarily small and if the algorithm succeeds in finding the optimal allocation, lying will be harmful and even reduce gains for the advertisers, the algorithm achieves truthfulness in expectation.

In addition to our mechanism only being truthful with high probability, in the B-MAFIA algorithm there are cases in which the advertiser can gain utility by lying about some of his parameters. This in one respect weakens the result as compared to our original MAFIA algorithm [13]. However, in [12] we show that achieving a truthful mechanism is impossible given budget and time constraints.

[6]It is important to recall that our model studies a one-shot incomplete information game. This means that advertisers do not change their valuations in the different time periods of the algorithm.

A similar and illustrative impossibility results was shown in a different context [18].

[18] setting when translated to our context has unit demand and no budget constraints. In [12] the impossibility result is extended to handle the case where demand is determined by the budget as well as multi-item demand. Even more so it is shown that even the assumption made in [17] that the budget is significantly larger than the value is insufficient to completely eliminate the possibility of gain in many settings.

As a result of the impossibility of optimizing welfare [18] choose to look at what they call *set-Nash-equilibria* whereas in our setting we choose to bound the gain of the players. We argue that if the utility gained by the advertiser is bounded, with respect to the total utility and does not depend on the magnitude of the lie, the lie strategy will not be used when a high risk of loss is tied to a low expectation of gain.

To capture the above notion we define the concepts of δ-gain and relative ϵ-gain. The following definitions allow for a general setting definition of the δ-gain and relative ϵ-gain concepts.

Let $\mathcal{M}$ be a mechanism such that given n vectors of parameters of n players $\overline{vec}_i$ where $1 \leq i \leq n$, it outputs an allocation denoted $\mathcal{A}(\overline{vec}_1, \ldots, \overline{vec}_n)$ and a set of payment $\mathcal{P}(\overline{vec}_1, \ldots, \overline{vec}_n)$ (a payment for every player). Let vec_i denote the true parameter vector of player i and $\overline{vec}_i$ the one reported to the mechanism. Also let $U_i(\mathcal{A}(\overline{vec}_1, \ldots, \overline{vec}_n), \mathcal{P}(\overline{vec}_1, \ldots, \overline{vec}_n))$ the utility of player i given input parameters vectors $\overline{vec}_1, \ldots, \overline{vec}_n$ and allocation $\mathcal{A}$. For simplicity of presentation we will denote the utility of player i by $U_i(\overline{vec}_1, \ldots, \overline{vec}_n)$. For the ease of the following definitions' presentation we denote $U_i(\overline{vec}_1, \ldots, \overline{vec}_n)$ by $\overline{u}$,
$U_i(\overline{vec}_1, \ldots, \overline{vec}_{i-1}, vec_i, \overline{vec}_{i+1} \ldots, \overline{vec}_n)$ by $\overline{u}'$,
$U_i(vec_1, \ldots, vec_n)$ by u and
$U_i(vec_1, \ldots, vec_{i-1}, \overline{vec}_i, vec_{i+1}, \ldots, vec_n)$ by u'

Definition 1. *We say that $\mathcal{M}$ is a δ-gain mechanism if for any player i it holds that*

$$\max_{\overline{vec}_i} u' - u < \delta$$

Definition 2. *We say that $\mathcal{M}$ is a dominant δ-gain mechanism if for any player i and any $\overline{vec}_j$ of the other players j it holds that*

$$\max_{\overline{vec}_i} \overline{u} - \overline{u}' < \delta$$

We believe that the notion of δ-gain (and dominant δ-gain) captures the essence of incentive compatibility. Although truthfulness might not be the optimal strategy it is nearly optimal for every player and given a setting where manipulating values incurs some cost (and/or some risk), the cost of strategizing will exceed the minute benefit of lying.

Furthermore in the full version we show that the B-MAFIA mechanism holds a stronger property than dominant δ-gain. We show that the B-MAFIA mechanism is also *dominant relative ϵ-gain* which intuitively means that the gain is bounded as a small fraction of the total value which would be achieved by a truthful player.

Denote by $\upsilon_i(\mathcal{A}(\overline{vec}_1, \ldots, \overline{vec}_n))$ the value of player i for allocation $\mathcal{A}(\overline{vec}_1, \ldots, \overline{vec}_n)$. For simplicity of presentation we will denote it $\upsilon_i(\overline{vec}_1, \ldots, \overline{vec}_n)$. For the ease of the following definitions' presentation we denote $\upsilon_i(\overline{vec}_1, \ldots, \overline{vec}_n)$ by $\overline{\eta}$,
$\upsilon_i(\overline{vec}_1, \ldots, \overline{vec}_{i-1}, vec_i, \overline{vec}_{i+1} \ldots, \overline{vec}_n)$ by $\overline{\eta}'$,
$\upsilon_i(vec_1, \ldots, vec_n)$ by η and
$\upsilon_i(vec_1, \ldots, vec_{i-1}, \overline{vec}_i, vec_{i+1}, \ldots, vec_n)$ by η'

Definition 3. *We say that $\mathcal{M}$ is a relative ϵ-gain mechanism if for any player i it holds that*

$$\max_{\overline{vec}_i} \frac{\eta' - \eta}{\eta} < \epsilon$$

Relative ϵ-gain means that not only is the gain bounded in absolute terms but it is also bounded in relative terms to the total gain. Similarly to the concept of δ-gain we define the dominant relative ϵ-gain.

Definition 4. *We say that $\mathcal{M}$ is a dominant relative ϵ-gain mechanism if for any player i and any $\overline{vec}_j$ of the other players j, it holds that*

$$\max_{\overline{vec}_i} \frac{\overline{\eta} - \overline{\eta}'}{\overline{\eta}'} < \epsilon$$

Before we proceed to state our main lemma we need some notation:

Let α_i^t and $\hat{\alpha}_i^t$ be i's click-through-rate found by the algorithm at time t with probability $1-\lambda$ and probability λ respectively when advertiser i reports $X_i = (a_i, l_i, v_i, b_i)$ and let $\overline{\alpha}_i^t$ and $\hat{\overline{\alpha}}_i^t$ be i's click-through-rate found by the algorithm at time t with probability $1-\lambda$ and probability λ respectively when advertiser i reports $\overline{X}_i = (\overline{a}_i, \overline{l}_i, \overline{v}_i, \overline{b}_i)$.

Similarly let ω_i and $\hat{\omega}_i$ be the number of clicks advertiser i received by time t with probability $1-\lambda$ and probability λ respectively when advertiser i reports $X_i = (a_i, l_i, v_i, b_i)$. Let $\overline{\omega}_i$ and $\hat{\overline{\omega}}_i$ be the number of clicks advertiser i received by time t with probability $1-\lambda$ and probability λ respectively when advertiser i reports $\overline{X}_i = (\overline{a}_i, \overline{l}_i, \overline{v}_i, \overline{b}_i)$.

Also let p_i^t and $\hat{p}_i^t$ be i's computed price by the algorithm at time t, with probability $1-\lambda$ and probability λ respectively when advertiser i reports $X_i = (a_i, l_i, v_i, b_i)$ and let $\overline{p}_i^t$ and $\hat{\overline{p}}_i^t$ be i's computed price by the algorithm at time t with probability $1-\lambda$ and probability λ respectively when advertiser i reports $\overline{X}_i = (\overline{a}_i, \overline{l}_i, \overline{v}_i, \overline{b}_i)$.

The following theorem states the main truthfulness claim:

Theorem 1. *Given advertiser i, for all $\overline{v}_i \neq v_i$, $\overline{a}_i \neq a_i$, $\overline{l}_i \neq l_i$, $\overline{b}_i \neq b_i$, reported by advertiser i and for all time periods t such that $a_i \leq t \leq l_i$ that results in price $\overline{p}_i^t$ it holds that,* $(1-\lambda)(\alpha_i^t \cdot (\omega_i \cdot v_i - p_i^t) + 3v_i) + \lambda \cdot (\hat{\alpha}_i^t \cdot (\hat{\omega}_i \cdot v_i - \hat{p}_i^t)) \geq (1-\lambda)(\overline{\alpha}_i^t \cdot (\overline{\omega}_i \cdot v_i - \overline{p}_i^t)) + \lambda(\hat{\overline{\alpha}}_i^t \cdot (\hat{\overline{\omega}}_i \cdot v_i - \hat{\overline{p}}_i^t))$

Proof. The proof is composed of several different claims.

- The algorithm is truthful in high probability for sufficiently small c (large γ_{e_i} for all i).
- The mechanism is truthful in expectation.
- An advertiser can not gain by lying about arrival time regardless of the other advertisers' reports.
- An advertiser i can gain at most v_i by lying about departure time regardless of the other advertisers reports.
- An advertiser i can gain at most v_i by lying about value regardless of the other advertisers' reports.
- An advertiser i can gain at most v_i by lying about budget regardless of the other advertisers' reports

By combining the above claims it follows that the algorithm is dominant δ-gain when $\delta = 3 * v_i$[7]. Furthermore, the analysis will show that the gain can only occur when i depletes his budget. Since the value of an advertiser for an allocation is the number of clicks he receives from arrival time to departure time (multiplied by his value per click) and since an advertiser is not charged more than his value per click on any of the clicks he receives, his value over an allocation is at least his budget b_i. This almost shows that the relative gain is bounded by $O(\frac{v_i}{b_i})$. However, since advertisers can try to simply "game" the mechanism is order to increase their relative gain at the cost of decreased utility, the analysis needs to be a little more careful. In the full version we

[7] We believe that the constant is an artifact of the details of our proof.

prove that for every advertiser i the expected dominant relative ϵ-gain is bound by $O(\frac{v_i}{b_i})$.

Note that the last three claims hold for advertiser i for every stay in the system. Therefore, the algorithm is relative ϵ-gain with $\epsilon = \frac{3*v_i}{b_i}$ regardless of the number of arrivals. Denote by: $u_i^1 = (\alpha_i^{l_i} \cdot (\omega_i \cdot v_i - p^{l_i}) + 3v_i)$,$u_i^2 = (\overline{\alpha}_i^{\overline{l}_i} \cdot (\overline{\omega}_i \cdot v_i - \overline{p}^{l_i}))$,$u_i^3 = (\hat{\overline{\alpha}}_i^{\hat{\overline{l}}_i} \cdot (\hat{\overline{\omega}}_i \cdot v_i - \hat{\overline{p}}^{l_i}))$,$u_1^4 = (\hat{\alpha}_i^{l_i} \cdot (\hat{\omega}_i \cdot v_i - \hat{p}^{l_i}))$

Claim 1. *If* $\lambda \leq \min_{i \in N} \left\{ \frac{u_i^1 - u_i^2}{u_i^1 - u_i^2 + \max_{i \in N}\{u_i^3 - u_i^4\}} \right\}$ *then for all advertisers their optimal strategy assumes that the optimal allocation is found*

Proof. If we set λ to be small enough, meaning $\lambda \leq \min_{i \in N} \left\{ \frac{u_i^1 - u_i^2}{u_i^1 - u_i^2 + \max_{i \in N}\{u_i^3 - u_i^4\}} \right\}$ then the probability of the algorithm not correctly finding the best advertisers will result in an arbitrarily low probability of an advertiser being able to gain by lying and hence with high probability the advertiser will tell the truth. Formally, since $\theta \leq \lambda$ then as $\lambda_{c \to \infty} \to 0$ then $\theta \to 0$ So when setting λ as above all that remains is to show that for every advertiser i at every departure time l_i, it holds that $\alpha_i^{l_i} \cdot (\omega_i \cdot v_i - p^{l_i}) + 3v_i \geq \overline{\alpha}_i^{\overline{l}_i} \cdot (\overline{\omega}_i \cdot v_i - \overline{p}^{l_i})$. □

So when setting λ as above all that remains is to show that for every advertiser i at every departure time l_i, it holds that $\alpha_i^{l_i} \cdot (\omega_i \cdot v_i - p^{l_i}) + 3v_i \geq \overline{\alpha}_i^{\overline{l}_i} \cdot (\overline{\omega}_i \cdot v_i - \overline{p}^{l_i})$.

The following lemma shows that the mechanism is truthful in expectation.

Lemma 1. *The B-MAFIA mechanism is truthful in expectation.*

The proof is omitted and appears in the full paper.

We use the following two technical lemmas to prove the main lemma's claims regarding the bounded gain advertiser i can achieve when lying.

Lemma 2. *For any set of non-negative numbers Q of cardinality $|Q|$, for any value $d < |Q|$ and for any non-negative number q:*

- *The value of the d'th smallest number in Q is always less or equal than the $d+1$ smallest number of $Q \cup q$ regardless of the value of q.*
- *The sum of the d'th smallest values in Q is no greater than the sum of the $d+1$ smallest values of $Q \cup q$*

Proof. .

- If $q \leq$ to the d'th smallest number in Q then the d'th smallest number in Q is equal to the $d+1$'th smallest number in $Q \cup q$ as the d'th smallest in Q becomes the $d+1$'th smallest in $Q \cup q$.

  If $q >$ than the d'th smallest number in Q then the d'th smallest number in Q is less than the $d+1$'th smallest number in $Q \cup q$ as q becomes at least the $d+1$'th smallest in $Q \cup q$.

- As the d'th smallest value in Q is always less or equal than the $d+1$ smallest number of $Q \cup q$ it follows that at least d of the values is the sum of the $d+1$ smallest values of $Q \cup q$ must be taken from Q.

□

The following key lemma bounds the amount that the temporary price can *decrease*:

Lemma 3. *For any vector $\overline{X}_i$ and for any time $t_1 < t_2$ it holds that $p_i^{t_1} - p_i^{t_2} \leq \overline{v}_i$. Furthermore for price $p_i^{t_1}$ computed by $\overline{X}_i = (\overline{a}_i, \overline{l}_i, \overline{b}_i, v_i)$, $p_i^{t_1} - p_i^{t_2} \leq v_i$.*

Proof. First it is shown that for any time period t, $t_1 < t \leq t_2$ it holds that $p_i^t - p_i^{t+1} \leq \overline{v}_i$. In other words, the decrease in advertiser i's temporary price from one period of time to the next is bounded by $\overline{v}_i$. The last claim is true as the highest critical value that can be consider in every time period for the price computation is i's reported value, i.e. $\overline{v}_i$, and the lowest critical value that could have been added to the price computation by time period $t+1$ is zero.

Next it is shown that for any time period $t_1 < t_2$ the decrease in advertiser i's temporary price from one period of time to a different period of time (not necessarily the next) is bounded by $\overline{v}_i$. Let t, t' be time periods such that $t < t'$. Assume to the contrary that: $p_i^t - p_i^{t'} = \overline{v}_i$ and $p_i^t - p_i^{t'+1} > \overline{v}_i$. There are three cases to consider:

- $c_i^{t'+1} \in W_{1(t'+1)}^i$: Since $p_i^t - p_i^{t'} = \overline{v}_i$ advertiser i has enough budget(accumulated at time t') to be given an allocation at time $t'+1$. As advertiser i did not get an impression at time $t'+1$ it follows that the critical value at time $t'+1$, i.e. $c_i^{t'+1}$ is greater than $\overline{v}_i$. As the additional critical value for the price computation at time $t'+1$ is over $\overline{v}_i$ the price could not have decreased from time t' to time $t'+1$ in contradiction to our contrary assumption that $p_i^t - p_i^{t'+1} > \overline{v}_i$.

- $c_i^{t'+1} \in W^i_{2(t'+1)}$: If advertiser i got an impression but not a click, $c_i^{t'+1}$ is not considered for the price $p_i^{t'+1}$ computation. Thus $p_it' + 1 = p_i^{t'}$ which contradicts our contrary assumption.

- $c_i^{t'+1} \in W^i_{3(t'+1)}$: If advertiser i got a click at time $t'+1$ then there is an additional critical value to consider for the price computation $p_i^{t'+1}$ but there is also an additional critical value to sum into the price. It follows from lemma 2 that no price decrease can occur from time t' to time $t'+1$ in contradiction to our contrary assumption.

Finally it needs to be argued that i's price difference of time t_1 and t_2 is actually bounded by v_i and not just $\overline{v}_i$ when $p_i^{t_1}$ is computed by the parameter vector $\overline{X}_i = (\overline{a}_i, \overline{l}_i, \overline{b}_i, v_i)$ where v_i is reported truthfully. Recall that it was claimed that the reason advertiser i's temporary price can change by $\overline{v}_i$ follows from the fact that the highest critical value that can be considered in every time period for the price computation is i's reported value, i.e. $\overline{v}_i$ and the lowest critical value that could have been added to the price computation by time period t_2 is zero. As advertiser i is either over or under paying for a click where his reported value $\overline{v}_i$ is taken into consideration, his price for the click when reporting his true value would have been v_i and therefore the difference between $p_i^{t_1} - p_i^{t_2}$ is bounded by v_i. □

Following claim 1 we assume in the next claims that the algorithm correctly finds the best advertisers. So fixing advertiser i, we first want to show that the advertiser does not lie about his budget regardless of his stated value and times. There are essentially two possible lies that i can make regarding budget. i can either increase his budget or decrease his budget. We will show that both of these possible lies yield at most a bounded gain of v_i.

Claim 2. *Given advertiser i with stated parameters $\overline{X}_i = (\overline{a}_i, \overline{l}_i, \overline{b}_i, \overline{v}_i)$, and for all other advertisers j with stated parameters $\overline{X}_i = (\overline{a}_i, \overline{l}_i, \overline{b}_i, \overline{v}_i)$, if advertiser i increases his stated budget by reporting $\overline{b}_i > b_i$ he can increase his utility by at most v_i for all time periods $a_i \leq t \leq l_i$ accumulatively.*

Proof. If advertiser i increases his reported budget then he may get exposures that he would not have if he reported his true budget. Let time period t' be such that i received an impression when bidding $\overline{b}_i$ and not when bidding b_i. The last scenario may happen only if i's value is higher than the critical value at time t', i.e., $c_i^{t'} < v_i$, but the true remaining budget is lower then the critical value, i.e., $B_i^{t'} < c_i^{t'}$ and the remaining budget resulting from the inflated lie budget is greater then the critical value $\overline{B}_i^{t'} \geq c_i^{t'}$.

There are two cases to consider:

- $t' \in \overline{W}^i_{3t'}$ - i got a click at time t' when increasing his stated budget. In this case one may think that an additional critical value appears in $\overline{W}^i_{3t'} \cup \overline{W}^i_{1t'}$ as $|\overline{W}^i_{3t'}| = |W^i_{3t'}| + 1$. However that is not the case as when i reports his true budget i does not receive an impression at time t'. Therefore, $|\overline{W}^i_{1t'}| = |W^i_{1t'}| - 1$. Since the number of clicks increased by one but the critical values considered did not change the price can only increase.

- $t' \in \overline{W}^i_{2t'}$ - i did not get a click at time t' when increasing his stated budget. In this case when i is lying he removes the critical value of time t' from consideration. This means $|W^i_{1t'}| - 1 = |\overline{W}^i_{1t}|$ but i does not add any other critical values to $\overline{W}^i_{3t}$. Therefore $|\overline{W}^i_{3t'} \cup \overline{W}^i_{1t'}| < |W^i_{3t'} \cup W^i_{1t'}|$ and $|W^i_{3t'}| = |\overline{W}^i_{3t'}|$. Thus the price computed can only increase.

Now consider advertiser i who, by increasing his stated budget, benefited from exposures that he would not have if he had reported his true budget. Such a situation leads i to run out of temporary remaining budget in some time period t where he would still have enough budget remaining if he stated the true budget.

There are two cases to consider:

- $c_i^t \in W^i_{3t}$ - i got a click at time t when reporting his true budget. In this case i does not change the critical values for consideration as c_i^t is removed from $\overline{W}^i_{3t}$ and added to $\overline{W}^i_{1t}$ meaning $|\overline{W}^i_{3t} \cup \overline{W}^i_{1t}| = |W^i_{3t} \cup W^i_{1t}|$. Nevertheless the number of clicks is reduced. Similar argument to lemma 3 hold here as the most "expensive" critical value that can be reduced out of the price computation is $\overline{v}_i$ and since i was over/under paying for at lease one click by the critical value $\overline{v}_i$ the most i can gain in this case is v_i.

- $c_i^t \in W^i_{2t'}$ - i did not get a click at time t when reporting his true budget. In this case when i is lying he adds the critical value of time t to consideration meaning $|W^i_{1t}| < |\overline{W}^i_{1t}|$, but does not reduce any other critical values to $\overline{W}^i_{3t}$. Therefore $|\overline{W}^i_{3t'} \cup \overline{W}^i_{1t'}| > |W^i_{3t'} \cup W^i_{1t'}|$ and the number of clicks to consider does not change meaning

$|W^i_{3t'}| = |\overline{W}^i_{3t'}|$. Again similar argument to lemma 3 holds here and the most i can gain in this case is v_i.

□

The following claim shows that i can also gain at most v_i when decreasing his stated budget. Note that the events of raising and lowering budget are mutually exclusive.

Claim 3. *Given advertiser i with stated parameters $\overline{X}_i = (\overline{a}_i, \overline{l}_i, \overline{b}_i, \overline{v}_i)$, and for all other advertisers j with stated parameters $\overline{X}_i = (\overline{a}_i, \overline{l}_i, \overline{b}_i, \overline{v}_i)$, if advertiser i decreases his stated budget by reporting $\overline{b}_i < b_i$ he can only increase his utility by at most v_i for all time periods $a_i \leq t \leq l_i$ cumulatively.*

Proof. At first glance, this appears quite obvious. Decreasing budget can only cause i to compete in fewer time slots and hence can only decrease i's utility. In fact, the problem is more subtle as reducing the stated budget can cause i not to compete in some time slot t, meaning i gets no exposure $c^t_i \in \overline{W}^i_{1t}$ and hence reduces the price that i pays at time t and on. There are two cases to consider in which the price can be reduced:

- If $c^t_i \in W^i_{3t}$ then there is one less click to sum in the price computation as $|\overline{W}^i_{3t}| = |W^i_{3t}| - 1$. Since the critical value c^t_i reduced of W^i_{3t} is added to $\overline{W}^i_{1t}$ the set of critical values to consider for the price computation remains unchanged meaning $|W^i_{3t} \cup W^i_{1t}| = |\overline{W}^i_{3t} \cup \overline{W}^i_{1t}|$. Therefore the price can be reduced.
- If $c^t_i \in W^i_{2t}$ then set $\overline{W}^i_{1t}$ is greater than W^i_{1t} while set $\overline{W}^i_{3t}$ remains unchanged with respect to W^i_{3t}. As $|\overline{W}^i_{3t}| = |W^i_{3t}|$ the number of clicks to sum for the price computation remains the same while the total set is greater meaning $W^i_{3t} \cup W^i_{1t} \subset \overline{W}^i_{3t} \cup \overline{W}^i_{1t}$. Therefore the price can be reduced.

For both cases following lemma 3 we conclude that i can not reduce the price more then v_i and thus his utility does not improve more than v_i. □

The proof of dominant δ-gain truthfulness continues by bounding the gain in utility that can be achieved by lying about the value. It is shown that for any stated parameters $\overline{X}_i = (\overline{a}_i, \overline{l}_i, b_i, \overline{v}_i)$ the utility i can gain by reporting $\overline{v}_i \neq v_i$ is bounded by v_i. We commence with the report of a decreased value.

Claim 4. *Given advertiser i with stated parameters $\overline{X}_i = (\overline{a}_i, \overline{l}_i, b_i, \overline{v}_i)$, and for all other advertisers j with stated parameters $\overline{X}_i = (\overline{a}_i, \overline{l}_i, \overline{b}_i, \overline{v}_i)$, if advertiser i decreases his stated value by reporting $\overline{v}_i < v_i$ he can not increase his utility at any time period $a_i \leq t \leq l_i$.*

Proof. First note that if there is no budget constraint, by monotonicity i clearly will be allocated fewer time periods when declaring $\overline{v}_i$ as opposed to declaring v_i. Since i has a positive marginal utility from each time period allocation. This shows that if there is no budget constraint i can not gain by reducing his value.

Alas, we have budget constraints and therefore i might be forced to pay more for the time periods i receives when declaring his true higher valuation. Therefore, a slightly more delicate analysis is required.

Let t be a time period in which i received an impression when declaring $\overline{v}_i$. If i also receives an impression at time t when declaring v_i then it follows that i is no worse off at time t when declaring v_i than when declaring $\overline{v}_i$, since the same critical value is taken into account for the price computation c^t_i and the number of clicks is identical in both cases.

Let t' be a time period in which i did not receive an impression when declaring $\overline{v}_i$. If i does not receive an impression also at time t' when declaring v_i then it follows that i is no worse off at time t' when declaring v_i than when declaring $\overline{v}_i$, since the same critical value is taken into account for the price computation c^t_i and the number of clicks is identical in both cases.

Therefore, there are two cases to consider:

- time t in which i receives an impression when declaring $\overline{v}_i$ but not when declaring v_i. As $\overline{v}_i < v_i$ the above situation can only happen if i has insufficient temporary budget at time t. Also note that i has the same budget for declaring v_i and $\overline{v}_i$ as we assumed that i declares his true budget following claims 3 and 2. Since i has no temporary budget left when declaring v_i there must be some time period in which i received a click through when declaring v_i and not when declaring $\overline{v}_i$. Therefore the number of click throughs is no less when declaring v_i than when declaring $\overline{v}_i$. It remains to show that the payments are no less when declaring $\overline{v}_i$. There are two cases to consider:
  - i got a click at time t - $c^t_i \in \overline{W}^i_{3t}$. In this case when i is lying he increases the number of clicks to consider for the price computation

while the critical values to consider remain unchanged meaning $|\overline{W}^i_{1t} \cup \overline{W}^i_{3t}| = |W^i_{1t} \cup W^i_{3t}|$. The last fact is easy to see as c^t_i is add to $\overline{W}^i_{3t}$ while being removed from W^i_{1t}. Thus the price computed can only increase.

  - i did not get a click at time t - $c^t_i \in \overline{W}^i_{2t}$. In this case when i is lying he removes the critical value of time t, c^t_i, from consideration for the price computation but does not add any other critical values that need to be considered as his ad is exposed but not clicked. Formally $|\overline{W}^i_{1t} \cup \overline{W}^i_{3t}| < |W^i_{1t} \cup W^i_{3t}|$ but note that the number of clicks to consider for the price computation is identical. Thus the price computed can only increase.

- time t' in which i does not receive an impression when declaring $\overline{v}_i$ but receives an impression when declaring v_i. As $\overline{v}_i < v_i$ the above situation can only happen if $\overline{v}_i < c^t_i$. There are two cases to consider:

  - i got a click at time t' when reporting v_i - In this case the critical value $c^{t'}_i$ is removed from $\overline{W}^i_{3t'}$ and added to $\overline{W}^i_{1t'}$ and therefore $|\overline{W}^i_{1t'} \cup \overline{W}^i_{3t'}| = |W^i_{1t'} \cup W^i_{3t'}|$, while the number of clicks to consider is reduced and thus the price can be reduced. However as advertiser i lost a click in order to reduce the price and the price can be reduced at most by v_i according to lemma 3 it follows that i did not gain utility by lying.

  - i did not get a click at time t' when reporting v_i - In this case the critical value $c^{t'}_i$ is removed from $\overline{W}^i_{2t'}$ and added to $\overline{W}^i_{1t'}$ and therefore $|\overline{W}^i_{1t'} \cup \overline{W}^i_{3t'}| > |W^i_{1t'} \cup W^i_{3t'}|$ while the number of clicks to consider is identical both when i stats v_i and when i stats $\overline{v}_i$. At first glance it looks like the price can be reduced in this case but that is not the case as $c^{t'}_i > \overline{v}_i$ and therefore there must be other points in time where i got a click when stating $\overline{v}_i$ in which the critical value is lower. Thus i does not reduce his price.

□

Claim 5. *Given advertiser i with stated parameters $\overline{X}_i = (\overline{a}_i, \overline{l}_i, \overline{b}_i, \overline{v}_i)$, and for all other advertisers j with stated parameters $\overline{X}_i = (\overline{a}_i, \overline{l}_i, \overline{b}_i, \overline{v}_i)$, if advertiser i increases his stated value by reporting $\overline{v}_i > v_i$ he can only increase his utility by at most v_i for all time periods $a_i \leq t \leq l_i$ accumulatively.*

Proof. Similarly to claim 4's proof, let t be a time period in which i received an impression when declaring $\overline{v}_i$. If i also receives an impression at time t when declaring v_i then it follows that time t is no worse off for i when declaring v_i then when declaring $\overline{v}_i$ as the same critical value is taken into account for the price computation c^t_i and the number of clicks is identical in both cases.

Let t' be a time period in which i did not receive an impression when declaring $\overline{v}_i$. If i does not receive an impression also at time t' when declaring v_i then it follows that time t' is no worse off for i when declaring v_i than when declaring $\overline{v}_i$ as the same critical value is taken into account for the price computation c^t_i and the number of clicks is identical in both cases.

Therefore, we focus on two cases:

- Time t in which i receives an impression when declaring $\overline{v}_i$ but not when declaring v_i. As $\overline{v}_i > v_i$ the above situation can only happen if $v_i < c^t_i$. There are two cases to consider:

  - i got a click at time t when reporting $\overline{v}_i$. In this case the critical value c^t_i is added to $\overline{W}^i_{3t}$ and removed from $\overline{W}^i_{1t}$ and therefore $|\overline{W}^i_{1t} \cup \overline{W}^i_{3t}| = |W^i_{1t} \cup W^i_{3t}|$ while the number of clicks to consider increases and thus the price can only increase.

  - i did not get a click at time t when reporting $\overline{v}_i$. In this case the critical value c^t_i is added to $\overline{W}^i_{2t}$ and removed from $\overline{W}^i_{1t}$ and therefore $|\overline{W}^i_{1t} \cup \overline{W}^i_{3t}| < |W^i_{1t} \cup W^i_{3t}|$. The number of clicks to consider is identical both when i stats v_i and when i stats $\overline{v}_i$ and thus the price can not decrease.

- Time t' in which i does not receive an impression when declaring $\overline{v}_i$ but receives an impression when declaring v_i. As $\overline{v}_i > v_i$ the above situation can only happen if i has insufficient temporary budget at time t'. Although i loses his impression when lying he might be able to reduce the price. It remains to show that the payments are no less when declaring $\overline{v}_i$. There are two cases to consider:

- i got a click at time t' - $c_i^{t'} \in W_{3t'}^i$. In this case when i is lying he decreases the number of clicks to consider for the price computation while the critical values to consider remain unchanged. This means $|\overline{W}_{1t'}^i \cup \overline{W}_{3t'}^i| = |W_{1t'}^i \cup W_{3t'}^i|$. The last fact is easy to see as $c_i^{t'}$ is removed from $\overline{W}_{3t'}^i$ while being added to $W_{1t'}^i$. Thus the price computed can decrease by at most $\overline{v}_i$. Though the price can decrease by at most $\overline{v}_i$, i's gain is at most v_i as the critical value reduce from the price computation was an over charge for i.
- i did not get a click at time t' - $c_i^{t'} \in W_{2t'}^i$. In this case when i is lying he adds the critical value of time t', $c_i^{t'}$, to consideration for the price computation. Formally $|\overline{W}_{1t'}^i \cup \overline{W}_{3t'}^i| > |W_{1t'}^i \cup W_{3t'}^i|$ but the number of clicks to consider for the price computation is identical. Thus the price computed can decrease by at most $\overline{v}_i$. Similar to the previous case although the price can decrease by at most $\overline{v}_i$, i's gain is at most v_i as the critical value remove from the price computation was an over charge for i.

□

We conclude our proof of dominant δ-gain truthfulness by bounding the gain in utility that can be achieved by lying about the duration of stay in the system. We show that for any stated parameters $\overline{X}_i = (\overline{a}_i, \overline{l}_i, b_i, \overline{v}_i)$ the utility i can gain by reporting a different duration in the system is bounded by v_i. First it is shown that the stated duration of a stay includes (in a set-theoretic sense) the true duration:

Claim 6. *Given advertiser i with stated parameters $\overline{X}_i = (\overline{a}_i, \overline{l}_i, \overline{b}_i, \overline{v}_i)$, and for all other advertisers j with stated parameters $\overline{X}_i = (\overline{a}_i, \overline{l}_i, \overline{b}_i, \overline{v}_i)$, if advertiser i decreased his stated duration in the system by reporting $\overline{a}_i > a_i$ or $\overline{l}_i < l_i$ or both then he can not increase his utility at any time period $a_i \leq t \leq l_i$.*

Proof. First assume that the advertiser i only lies about his departure time such that i departs at earlier time than he should have, i.e., $\overline{l}_i < l_i$. In this case i gets identical exposure until time $\overline{l}_i$ in his current stay in the system (assuming w.l.o.g no lies in previous stays, e_i is identical in both cases until time $\overline{l}_i$) and hence γ_{e_i} is the same in both cases. It follows that the algorithm learns i's click-through rate $\alpha_i^{l_i}$ in an identical manner at time $\overline{l}_i$ regardless of i's reported departure time.

Therefore, the only way i can increase his utility is by decreasing his price. However, the price that i pays is computed as the minimum critical prices during time in which i either received a click through or did not have an exposure. Decreasing the reported time can only result in the mechanism calculating the minimum over a smaller set if i received no exposures from time $\overline{l}_i$ to l_i as the number of click-throughs remains the same.

In the case of clicks from time $\overline{l}_i$ to l_i as i has positive marginal utility in every click in time t where $a_i \leq t \leq l_i$ i can not increase utility by losing a desirable click.

Now consider an advertiser i who lies about his arrival time. By lying about the arrival time, $\overline{a}_i > a_i$, i can affect γ_{e_i} at time $\overline{a}_i$ (as e_i is decreased at time $\overline{a}_i$ when lying). In this case i will receive more frequent exposures when lying than when telling the truth, but that will happen only until e_i is identical in both cases. At that point in time e_i is equal in both cases and therefore γ_{e_i} will be equal too and i can not expect a greater number of exposures when reporting $\overline{a}_i$.

Since, i can gain no more exposures when reporting $\overline{a}_i$ it remains to show that i can not reduce his price by reporting $\overline{a}_i$. Denote by t the time period in which e_i is identical both when i reported $\overline{a}_i$ and when reporting a_i.

For any time point t' such that $\overline{a}_i \leq t' < t$, if i receives an exposure both when reporting a_i and $\overline{a}_i$ then i can not change the price as the number of critical values and clicks to consider are identical in both cases.

The only times where i can change the price is at times t' such that i gets an exposure when reporting $\overline{a}_i$ and no exposures when reporting a_i. Such time periods can occur for two reasons; i required more exploitation by the bandit algorithm when lying as no exploitation occurred from time a_i to $\overline{a}_i$, or i's budget ran out when reporting the truth due to clicks from time a_i to $\overline{a}_i$ but i had budget left when lying.

There are two cases to consider:

- i received a click at time t' when lying, meaning $c_i^{t'} \in \overline{W}_{3t'}^i$. In this case the critical value $c_i^{t'}$ is removed from $\overline{W}_{1t'}^i$ but is added to $\overline{W}_{3t'}^i$ and so $|\overline{W}_{1t'}^i \cup \overline{W}_{3t'}^i| = |W_{1t'}^i \cup W_{3t'}^i|$. Although the number of clicks maybe less when i is lying, the number of clicks from time $\overline{a}_i$ to t' increased and therefore the price can only increase on that period of time. For any time period before $\overline{a}_i$ we already know that i has a positive marginal utility for every click and therefore i can not gain from lying in this case.

- i received exposure but no click when lying, meaning $c_i^{t'} \in \overline{W}^i_{2t'}$. In this case the critical value $c_i^{t'}$ is removed from $\overline{W}^i_{1t'}$ and therefore $|\overline{W}^i_{1t'} \cup \overline{W}^i_{3t'}| < |W^i_{1t'} \cup W^i_{3t'}|$ while the number of clicks does not change. Thus the price can only increase in this case when i is lying.

□

It now remains to show that i can not gain anything by decreasing his arrival time, i.e. $\overline{a}_i < a_i$.

Claim 7. *Given advertiser i with stated parameters $\overline{X}_i = (\overline{a}_i, \overline{l}_i, b_i, \overline{v}_i)$, and for all other advertisers j with stated parameters $\overline{X}_i = (\overline{a}_i, \overline{l}_i, \overline{b}_i, \overline{v}_i)$, if advertiser i decreased his stated arrival time by reporting $\overline{a}_i < a_i$ he can not increase his utility at any time period $a_i \leq t \leq l_i$.*

Proof. Naturally, i receives no value from any clicks before time a_i. It remains to show that i can not receive a lower price at any time t where $a_i \leq t \leq l_i$ as a result of lying about his arrival time and receive no exposures or clicks at time $\overline{a}_i \leq t < a_i$.

There are three cases to consider:

1. i did not receive exposure at time t, $\overline{a}_i \leq t < a_i$. There are two possible reasons advertiser i does not get exposure at time period t.
   (a) His critical value at time t is greater than his value, $c_i^t > v_i$. In this case, the critical value of any time period where i got a click would be less than c_i^t and hence will not effect i's price.
   (b) His critical value at time t is less (or equal) then his value, i.e., $c_i^t \leq v_i$ but i has insufficient budget, in particular, $B_i^t < v_i$. If c_i^t is very low and there is some other very high critical value $c_i^{t'}$ in time t' where $a_i \leq t' \leq l_i$ that was taken into the price computation, i can gain on time t' at most v_i. Recall that such gain can occur only when i ran out of remaining budget and that it is assumed $b_i >> v_i$. Since $B_i^t = b_i - p_i^t < v_i$ it means that i payed at least $b_i - v_i$ on undesired times to enjoy the gain of v_i. If $b_i \geq 2 \cdot v_i$ there is no gain in i's utility at any time t'.
2. i received an impression but no click at time t, $\overline{a}_i \leq t < a_i$, in this case i can not affect the price as the critical value in those cases is not taken into account for the price computation.
3. i received a click at time t, $\overline{a}_i \leq t < a_i$ if that click increased the total number of clicks that i received when lying with comparison to the total number of clicks when being truthful meaning $|W^i_{3t''}| > |W^i_{3t'}|$, where $a_i \leq t' \leq l_i$ and $\overline{a}_i \leq t'' \leq l_i$ then it follows from lemma 2 that i can not improve his computed price at time t''. The last argument holds only if the clicks at time t did not reduce the remaining budget at any time t' such that i could not get an exposure due to lack of budget. If such a situation occurs there are two cases to consider:
   - i received an exposure but no click at time t' when telling his true arrival. In this case the critical value $c_i^{t'}$ is added to the critical values set in which the price is calculated from, meaning $|\overline{W}^i_{1t'} \cup \overline{W}^i_{3t'}| > |W^i_{1t'} \cup W^i_{3t'}|$. If $c_i^{t'}$ is very low and it is added to the price computation instead of some high critical value, i may gain at most v_i for every such time t'. Though it appears that the gain is unbounded, i must lose budget to enjoy such gain. Similar to the case where i runs out of budget before time a_i when lying, recall that the gain only occurs when i expended his remaining budget and also recall that we assumed $b_i >> v_i$. Since $B_i^{t'} = b_i - p_i^{t'} < v_i$ then i payed at least $b_i - v_i$ in undesired rounds to enjoy the gain of v_i. If $b_i \geq 2 \cdot v_i$ then there is no gain in i's utility at any time t'.
   - i received a click at time t' when telling the true arrival. In this case as the critical value $c_i^{t'}$ is not added or reduced from the critical values set in which the price is calculated from, meaning $\overline{W}^i_{1t'} \cup \overline{W}^i_{3t'} = W^i_{1t'} \cup W^i_{3t'}$. Since i has a positive marginal utility for every click in time t', by lying i can only lose.

□

In order to complete the proof that i can not benefit from lying about arrival and departure times, it only remains to show that i can not increase his utility by lying about departure times. Unfortunately this is not true; however, we can bound the benefit that i can achieve from such a lie.

Claim 8. *Given advertiser i with stated parameters $\overline{X}_i = (\overline{a}_i, \overline{l}_i, b_i, \overline{v}_i)$, and for all other advertisers j with stated parameters $\overline{X}_i = (\overline{a}_i, \overline{l}_i, \overline{b}_i, \overline{v}_i)$. If advertiser i increased his stated departure time by reporting $\overline{l}_i > l_i$ he can improve his utility by at most v_i for all time periods $a_i \leq t \leq l_i$ accumulatively.*

The proof follows:

Proof. Similar to claim 7, i receives no value from any clicks after time l_i. It remains to show that i's gain, as a result of lying about his departure time where he receives no exposures or clicks at time $l_i \leq t < \overline{l}_i$, is bounded, at any time t where $a_i \leq t \leq l_i$.

For any time t s.t. $l_i < t \leq \overline{l}_i$ there are exactly three possibilities:

1. $c_i^t \in W_{1t}^i$ (no impression)- There are two possible reasons advertiser i does not get an exposure at time period t.
   (a) His critical value at time t is greater than his value, $c_i^t > v_i$. We show that i achieves no advantage from this case.
   (b) His critical value at time t is less (or equal) then his value, i.e., $c_i^t \leq v_i$ but i has insufficient budget, in particular, $B_i^t < v_i$. We bound the total advantage i can receive from all such cases by v_i.

   We show the proof of the above points in lemma 4 below
2. $c_i^t \in W_{2t}^i$ (impression but no click through). We show that i receives no advantage from such t's (lemma 5 below
3. $c_i^t \in W_{3t}^i$ (click through). We show that this case also does not help i (lemma 6 below

□

The following lemmas prove each of these cases separately.

Lemma 4. *The cumulative benefit to advertiser i during all time periods $a_i \leq t' \leq l_i$ of lying in time periods t, $l_i < t \leq \overline{l}_i$ such that $c_i^t \in W_{1t}^i$ (no impression), is bounded by v_i.*

Proof. The proof is divided into two cases such that i does not receive an impression in time t:

1. i's value was too low, $c_i^t > v_i$.
2. i has insufficient budget left, $B_i^t < v_i$.

In the first case i receives no benefit from time t. In this case by definition $c_i^t > v_i$. For any time t' s.t. i received a click through in time t^*, it has to hold that i also received an impression and hence $c_i^{t^*} \leq v_i$. Therefore, $c_i^t > c_i^{t^*}$. Since there are $|W_{3t}^i|$ click throughs and hence i must only pay the $|W_{3t}^i|$ lowest values in $W_{1t}^i \cup W_{3t}^i$ and there are at least $|W_{3t}^i|$ lower values than c_i^t (in W_{3t}^i) it follows that c_i^t does not affect the payment.

The second case is harder. Note that during every time t s.t. player i does not have enough budget but otherwise would have received an impression $0 \leq B_i^t < c_i^t < v_i$. In particular for the first time t^* as well as the last time t^{**} that this situation happens to advertiser i. Since our other claims show that i can not gain by any other way of lying about his arrival and departure times it follows that advertiser i's remaining budget at time t^{**}, $B_i^{t^{**}}$ means that i's temporary price at time t^{**} is actually $p_i^{t^{**}} = b_i - B_i^{t^{**}}$. However, if i would tell the truth about his departure time he would have to pay at most $p_i^{t^*} = b_i - B_i^{t^*}$ (it is possible less) and hence according to lemma 3 can gain at most $b_i - B_i^{t^{**}} - (b_i - B_i^{t^*}) \leq v_i$. □

Lemma 5. *Advertiser i can not improve his utility at any time period $a_i \leq t' \leq l_i$ by lying in time periods t, $l_i < t \leq \overline{l}_i$ such that $c_i^t \in W_{2t}^i$, i.e., i receives an impression but no click through.*

Proof. Since $c_i^t \in W_{2t}^i$ i can not affect the price as the critical values in W_{2t}^i are not taken into account for the price computation. □

Lemma 6. *Advertiser i can not improve his utility at any time period $a_i \leq t' \leq l_i$ by lying in time periods t, $l_i < t \leq \overline{l}_i$ such that $c_i^t \in W_{3t}^i$, i.e., i gets a click.*

Proof. It will be shown that i can not decrease his payment in any time period t', $a_i \leq t' \leq l_i$. For any time period t, $l_i < t \leq \overline{l}_i$ there are two effects on i's payment. The first is that an additional critical value appears in $W_{1t}^i \cup W_{3t}^i$ and hence might reduce the payment. However, the second effect is that the number of click throughs increases by one and hence by lemma 2 i can not reduce his payment. □

Assuming that B-MAFIA correctly finds the best advertisers with probability $(1 - \lambda)$ bounded as above, then the B-MAFIA algorithm is a dominant δ-gain truthful mechanism with probability $(1 - \theta)$. □

B.2 Welfare Maximization Approximation

This section shows that our algorithm approximates the optimal welfare of a budgeted sponsored search auction where the CTRs of the advertisers are known. However, before we can explore the welfare approximation proof we have to deal with the truthfulness properties. If advertisers incorrectly report their values then it is obviously impossible to maximize welfare. The main problem, is that our proof of truthfulness assumes approximation of the welfare. Since the proof of approximation requires truthfulness we are in somewhat of a bind.

In order to resolve this problem we show that we can in fact decouple the two proofs. This will follow from the fact that the $\theta - truthfulness$ property and the $\lambda - welfare$ property are positively correlated. The more truthful B-MAFIA is the better welfare it can achieve and vice-versa. In other words one can set λ to be small such that dominant δ-gain truthfulness is reached with probability $1-\theta \to 1$. Using lemma 7 and assuming truthfulness with probability 1, it is shown that the algorithm correctly finds the best advertisers with probability $1 - \lambda$.

Lemma 7. *Given a B-MAFIA algorithm that is dominant δ-gain truthful with probability $(1-\theta)$ that correctly finds the best advertisers with probability $1 - \lambda$ then if θ increases then λ increases and if θ decreases then λ decreases.*

Proof. We will show that if θ decreases then λ decreases. The other case is similar. If θ decreases it follows that the algorithm is dominant δ-gain truthful with higher probability. In this case the advertisers will report their true value with higher probability. As every advertiser i's observed payoff $x_i^t = \overline{v}_i \cdot \alpha_i^t$, left budget is $B_i^t = \overline{b}_i - p_i^t$, and $\overline{a}_i \leq t \leq \overline{l}_i$, if $\overline{v}_i = v_i$, $\overline{b}_i = b_i$, $\overline{a}_i = a_i$, and $\overline{l}_i = l_i$, with high probability then $\min\{x_z^t, B_z^t\} - \min\{x_i^t, B_i^t\} = \min\{\overline{x}_z^t, \overline{B}_z^t\} - \min\{\overline{x}_i^t, \overline{B}_i^t\}$ with higher probability and therefore for every advertiser $i \in S_t$ if there exist K other advertisers z such that $\min\{x_z^t, B_z^t\} > \min\{x_i^t, B_i^t\}$ and $\min\{x_z^t, B_z^t\} - \min\{x_i^t, B_i^t\} > \gamma_{e_i} + \gamma_{e_z} \Rightarrow \min\{\overline{x}_z^t, \overline{B}_z^t\} > \min\{\overline{x}_i^t, \overline{B}_i^t\}$ and $\min\{\overline{x}_z^t, \overline{B}_z^t\} - \min\{\overline{x}_i^t, \overline{B}_i^t\} > \gamma_{e_i} + \gamma_{e_z}$ with higher probability. Thus the lack of consideration/consideration of the desired/undesired advertiser at time t is with lower probability λ. □

Denote by OPT the welfare achieved by an offline optimal truthful algorithm that allocates budgeted advertisers with known click through rates to arriving advertisers. Also denote by ALG the welfare achieved by our B-MAFIA algorithm while learning the click through rates of the advertisers.

Denote by $\widehat{x}_i$ the real payoff of advertiser i and by x_t^i the observed payoff of advertiser i at time t. Let $\widehat{x}_{\max_K}$ be the advertiser with the K'th highest real payoff and let $\Delta_i = \widehat{x}_{\max_K} - \widehat{x}_i$.

The following theorem bounds the loss of welfare (in a PAC sense) of our B-MAFIA algorithm due to the sampling process.

Theorem 2. *With probability $1-\lambda$ B-MAFIA achieves a $O(\sum_{i=2}^{n} \Delta_i)$ approximation of the optimal welfare.*

To prove our theorem we first claim that the B-MAFIA algorithm achieves its worst expected approximation of welfare when all advertisers are present in the system simultaneously and there are no budget constraints. The intuition behind having all advertisers present in the system simultaneously is simple. Consider a single-slot setting and an optimal advertiser i who is present in all time periods. If there exist another non optimal advertiser j such that j is not present in the system some of the times, the B-MAFIA algorithm will allocate the slot to i with probability 1 at times where j is not around and thus improve approximation. The intuition behind the budget constraint improving the B-MAFIA approximation is that a reduction in B-MAFIA approximation can occur as a result of over sampling and particularly over sampling of a non optimal advertiser. When the budget is unlimited, over sampling is unbounded and therefore leads to a worse approximation.

A formal proof follows:

Lemma 8. *B-MAFIA achieves the worst (expected) approximation of the welfare when $a_i = a_j, l_i = l_j$ and $b_i = \infty$ for all advertisers i, j*

Proof. The proof is divided into two parts; infinite budget impact and simultaneous duration impact. We start by showing the infinite budget impact.

Assume that for advertiser i $b_i < \infty$. We look at several cases:

- Neither B-MAFIA nor OPT deplete the budget b_i. In this case setting $b_i = \infty$ does not change either mechanism and so the approximation remains identical whether the budget is $b_i = \infty$ or $b_i < \infty$.

- OPT depletes the budget but B-MAFIA does not. In this case if B-MAFIA had more budget i.e.,

$b_i = \infty$ to use it would not improve its approximation as B-MAFIA did not even deplete the bounded budget.

- B-MAFIA depletes the budget but OPT does not. In this case clearly, B-MAFIA over samples i. If B-MAFIA had more budget i.e., $b_i = \infty$ to use it might have over sampled even more and achieved a worse approximation.

- Both OPT and B-MAFIA deplete the budget b_i. There are two cases:

  - OPT allocates more slots to i than B-MAFIA. This can not happen since B-MAFIA charges the *cheapest* possible slots to i at every time period and therefore left with the highest possible budget for the future slots.

  - OPT allocates fewer slots to i than B-MAFIA. In this case the budget prevents B-MAFIA from allocating even more slots to i and hence prevents B-MAFIA from reducing the approximation ratio further.

The proof continues by showing the simultaneous duration impact. Assume that there exist advertiser i such that $[a_i, l_i] \neq [a_j, l_j]$ for advertisers j. If $[a_i, l_i] \cap [a_j, l_j] = \emptyset$ B-MAFIA, just like OPT, will allocate i during his time periods regardless of whether he is optimal or not as there is no other advertiser to allocate and therefore B-MAFIA will perform the same as OPT. Similarly if $[a_i, l_i] \cap [a_j, l_j] \neq \emptyset$ then for any time period where i is present and the other advertisers are not, B-MAFIA just like OPT, will allocate i regardless of whether he is optimal or not. In both cases B-MAFIA's approximation will compare better with OPT than if i had arrived and left the system with the other advertisers j. □

Now we can proceed to bound the loss in welfare assuming that there is no budget constraint and that all advertisers arrive and depart the system at the same time, i.e., the following lemma assumes that $a_i = a_j, l_i = l_j$ and $b_i = \infty$ for all advertisers i, j.

First it will be argued that the B-MAFIA algorithm considers the best advertisers with probability $1 - \lambda$ in every round. As the largest loss of welfare occurs when there is no budget constraint and all advertisers arrive and depart the system at the same time it suffices to show that with probability $1-\lambda$ there exist a time round where B-MAFIA allocates the K slots to the optimal advertisers.

Lemma 9. *If $b_i = \infty$ for all i and $a_i = a_j, l_i = l_j$ for all i, j then there exist τ such that the B-MAFIA algorithm finds the optimal welfare $\sum_{i \in N} \alpha_i \cdot v_i$ with probability $1 - \lambda$ for every $\tau \leq t \leq T$.*

Proof. Since we assume that $b_i = \infty$ for all i and $a_i = a_j, l_i = l_j$ for all i, j then the set S_t never changes and every advertiser that is removed from the set S'_t does not appear in a future set S'_τ, $\tau > t$. Therefore for this proof one can assume that advertisers are removed from S'_t until the optimal ones are left.

The main argument of the proof is that the observed payoff x_i^t of advertiser i at time t that was not removed is within γ^{e_i} of the true payoff x_i. As γ^{e_i} goes to zero as e_i increases (which is to say as t increases due to the fact that i's number of exposures increase with time) then after long enough time we are left with the advertiser with the best payoff, i.e. maximized social welfare.

As our B-MAFIA algorithm allocates K slots and not just one we also need to verify that we get the best K payoffs.

Let S'_t be the set of advertisers left in the auction at time t. For any time t and advertiser $i \in S'_t$ we have that,

$$\Pr[|x_i^t - x_i| \geq \gamma^{e_i}] \leq e^{-(\gamma^{e_i})^2 e_i} \leq \frac{\chi}{cn(e_i)^2 K} \qquad (1)$$

The first inequality follows from the Chernoff bound and the second by substituting γ^{e_i} in the bound. By union bound over the K slots and by union bound over all exposures e_i which is essentially union bound over times from $t = 1$ to T it follows that with probability at least $1 - \chi/n$ for any time t and any advertiser $i \in S'_t$, $|x_i^t - x_i| \leq \gamma^{e_i}$. Therefore with probability $1 - \chi$, the K'th best advertisers are never eliminated.

As the algorithm's sample complexity is bounded (see lemma 11), there exist time period τ where the algorithm finds the optimal welfare $\sum_{i \in N} \alpha_i \cdot v_i$ with probability $1 - \chi$ there after $\tau \leq t \leq T$. □

Lemma 10. *If $b_i = \infty$ for all i and $a_i = a_j, l_i = l_j$ for all i, j then the loss of welfare is tightly bound by $O(\sum_{i=2}^{n} \Delta_i)$ from OPT.*

Proof. Since we assume that $b_i = \infty$ for all i and $a_i = a_j, l_i = l_j$ for all i, j then the set S_t never changes and every advertiser that is removed from the set S'_t does not appear in a future set S'_τ, $\tau > t$. Therefore for this proof one can assume that advertisers are removed from S'_t until the optimal ones are left.

The welfare loss per advertiser i which is removed at some time period must be at least Δ_i since each advertiser is sampled at least once.

With arbitrary high probability the advertiser is sampled at most until e_i exposures were given s.t. $\Delta_i \leq 2\gamma^{e_i}$. Therefore, $\Delta_i < O(\sqrt{\frac{\log e_i^2}{e_i}})$ or $\Delta_i^2 = O(\frac{\log e_i}{e_i})$. It can be seen, by substitution, that the solution for e_i is $e_i = O(\frac{1}{\Delta_i^2} \log \frac{1}{\Delta_i^2})$. Since the welfare lost is $\Delta_i * e_i$ the loss is $O(\Delta_i * \frac{1}{\Delta_i^2} \log \frac{1}{\Delta_i^2}) = O(\frac{1}{\Delta_i} \log \frac{1}{\Delta_i^2}) = O(\frac{\log \Delta_i}{\Delta_i}) \rightarrow_{\Delta_i \rightarrow \infty} 0$ so the loss is bounded by the loss for small Δ_i and hence the welfare loss is optimal (up to constants). □

B.3 Suboptimal Exposure Complexity

This subsection bounds the number of suboptimal exposures that the B-MAFIA algorithm gives to advertisers for time 1 to T. A suboptimal exposure is an exposure given by the B-MAFIA algorithm for an advertiser i whose real payoff $\widehat{x}_i$ is not among the best K real payoffs at the time that he receives the exposure.

Following from lemma **??**'s proof the largest number of suboptimal exposures occurs when $b_i = \infty$ for all i and $a_i = a_j, l_i = l_j$ for all i, j and so the remainder of this subsection assumes such setting.

Lemma 11. *The B-MAFIA algorithm suboptimal exposure complexity is bounded by* $O\left(\sum_{i=2}^{n} \frac{\log(\frac{n}{\chi \cdot \Delta_i})}{\Delta_i^2 \cdot K}\right)$

Proof. Similar to lemma 10's proof, since we assume that $b_i = \infty$ for all i and $a_i = a_j, l_i = l_j$ for all i, j then the set S_t never changes and every advertiser who is removed from the set S'_t does not appear in a future set S'_τ, $\tau > t$. Therefore for this proof one can assume that advertisers are removed from S'_t until the optimal advertisers remain.

To prove the algorithm's suboptimal exposure complexity one needs to bound the number of exposures it will take to remove an advertiser from the set S'_t.

Let $x_t^{\max_K}$ be the K'th highest payoff advertiser in S'_t and denote that advertiser k. Consider advertiser i who should be removed then $x_t^{\max_K} - x_t^i \geq \gamma_{e_i} + \gamma_{e_k}$.

The assumption that advertiser i's real payoff and the observed one are different by at most γ_{e_i}, i.e., $|x_t^i - \widehat{x}_i| \leq \gamma_{e_i}$ yields that $\Delta_i - (\gamma_{e_i} + \gamma_{e_k}) = (\widehat{x}_{\max_K} - \gamma_{e_k}) - (\widehat{x}_i + \gamma_{e_i}) \geq x_t^{\max_K} - x_t^i \geq \gamma_{e_i} + \gamma_{e_k}$. The last inequality follows from the fact that i is considered for removal. Since Δ_i is a constant, $\gamma_{e_i} \rightarrow 0$, $\gamma_{e_k} \rightarrow 0$ there exists an e_i and e_k s.t. $\Delta_i \geq 2 \cdot \gamma_{e_i} + 2 \cdot \gamma_{e_k}$. By substituting $e_i = O\left(\frac{\log(\frac{n}{\chi \cdot \Delta_i})}{\Delta_i^2 \cdot K}\right)$ and $e_k = O\left(\frac{\log(\frac{n}{\chi \cdot \Delta_i})}{\Delta_i^2 \cdot K}\right)$ it holds that $\Delta_i > 2 \cdot \sqrt{\frac{\log(cne_i{}^2/\chi)}{e_i} \cdot \frac{1}{K}} + 2 \cdot \sqrt{\frac{\log(cne_k{}^2/\chi)}{e_k} \cdot \frac{1}{K}}$. As the algorithm removes all but K advertisers, the suboptimal exposures complexity of the algorithm is then (approximately) $\sum_{i=2}^{n} e_k$. □