The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-20)

VCG under Sybil (False-Name) Attacks — A Bayesian Analysis

Yotam Gafni, Ron Lavi, Moshe Tennenholtz
Technion - Israel Institute of Technology
Haifa 32000 Israel
yotam.gafni@campus.technion.ac.il, {ronlavi, moshet}@ie.technion.ac.il

Abstract

VCG is a classical combinatorial auction that maximizes social welfare. However, while the standard single-item Vickrey auction is false-name-proof, a major failure of multi-item VCG is its vulnerability to false-name attacks. This occurs already in the natural bare minimum model in which there are two identical items and bidders are single-minded. Previous solutions to this challenge focused on developing alternative mechanisms that compromise social welfare. We re-visit the VCG auction vulnerability and consider the bidder behavior in Bayesian settings. In service of that we introduce a novel notion, termed the *granularity threshold*, that characterizes VCG Bayesian resilience to false-name attacks as a function of the bidder type distribution. Using this notion we show a large class of cases in which VCG indeed obtains Bayesian resilience for the two-item single-minded setting.

1 Introduction

In recent years, the scale of web auctions has grown significantly — whether for ads, flight tickets or cloud computing resources — among many other commodities. Contrary to the settings that apply to government leases or items sold in an auction house, in such web auctions it is much easier to forge false identities and submit multiple bids under false names. This expands the action space of an auction participant and may turn mechanisms that proved formally truthful assuming only a single bid to not be strategy-proof. Such is the case of the most classic combinatorial auction mechanism, VCG. VCG is truthful and efficient assuming each bidder submits one bid under her name (Nisan et al. 2007). However, assuming false-name attacks (a.k.a. sybil attacks) are possible, that is no longer the case, as demonstrated in (Yokoo, Sakurai, and Matsubara 2002). The main effort to address this rising issue was done by Yokoo et al. They suggested various mechanisms that are *false name attack proof* (Yokoo, Sakurai, and Matsubara 2001; Yokoo 2003). The main disadvantage of these mechanisms is that they in a way incorporate the false-name attacks to be built into the mechanism, and this significantly reduces their efficiency. In (Iwasaki et al. 2010) the main result shows that under reasonable conditions the worst case efficiency of any false-name proof combinatorial auction with m items is as bad as $\frac{2}{m+1}$.

As for VCG, in (Alkalay-Houlihan and Vetta 2014) the authors show that under some assumptions on the bidders' valuations, and in a setting with complete information, if a pure Nash equilibrium exists when false-name bids are considered, VCG still has a reasonable welfare guarantee. Naturally, Nash equilibrium might not exist. Nevertheless, the above may hint that VCG may actually be a good false-name proof mechanism in situations in which a pure equilibrium exists.

Notice that when the behavior of the other bidders is uncertain, in some scenarios a false-name attack may benefit the attacker, while on others the attack can harm her. So while truthfulness is certainly no longer a dominant strategy, it is worth asking under what circumstances it might be a *truthful Bayesian Nash Equilibrium*. That is, when the auction participants only have partial probabilistic information regarding other bidders' valuations, when is it the best strategy for a bidder to bid truthfully only her valuation (in a single bid), assuming all others are truthful as well? Notice that social welfare is optimized if such an equilibrium exists. For a seller that wishes to maximize social welfare, under the conditions in which a truthful Bayesian equilibrium exists, VCG is the best choice of an auction mechanism.

In this work, we introduce — for the first time, to the best of our knowledge — a Bayesian equilibrium analysis of VCG under false-name bids. In service of that we use a bare-minimum model: an auction of two identical items, and single-minded bidders. Namely, each bidder is interested in a single item or in the pair of items. The Bayesian setting is given by a per-item valuation distribution, and a probability q for a bidder to have a single-item demand. Notice that for $q = 1$ we get the standard Vickrey auction which is false-name proof. Hence, our model captures the minimal step forward from the standard Vickrey auction setting to the general multi-item case. As we will see, the analysis of such a model already brings out intricate techniques and conclusions, which are essential to address the problem.

To obtain our results, we present a criterion we name "Granularity threshold" to measure the effect of the gran-

ularity of bidders' demands on the mechanism's Bayesian truthfulness. In the framework of our model, this measure is associated with the value of q. The lower the granularity threshold is, then even with a lower q value the mechanism remains truthful. We show results that prove that in our model, the more "granular" bidders' demands are, the more VCG is Bayesian resilient to false-name attacks.

Our main results state that for two bidders ($n = 2$) and two items ($m = 2$) there is a truthful Bayesian equilibrium whenever the probability q of a bidder to demand one item rather than two is at least $\frac{2}{3}$, with *any* valuation distribution (Section 3). We then show that for any larger number of bidders ($n > 2$), such a global granularity threshold can no longer be derived and valuation distributions do matter (Section 4). Given the above, we consider general beta distributions over per-item valuations. This family includes many natural distributions, and is used widely in statistics and economic theory (Gupta and Nadarajah 2004; Krishnamoorthy 2016). With parameters $\alpha = \beta = 1$, it is the uniform distribution, and with large $\alpha = \beta$ parameter values, it resembles a Normal distribution. Interestingly, it admits useful connections to computer algebra techniques, which allows us to provably present good granularity thresholds, i.e. low q values, in tested cases (Section 5). Lastly, we focus on an important attack form we call the *split attack*. In that attack a bidder who has a 2-item demand with valuation θ submits two 1-item demand bids with value θ. Interestingly, we observe that in many cases this is the best false-name attack in the sense that it determines the granularity threshold (Section 6).

2 Our Model and the Bayesian Resilience Criterion

We study a multi-item auction with two identical items and n single-minded bidders. The type of a bidder i is denoted by $\hat{\theta_i} = (g_i, \theta_i)$ where $g_i \in \{1, 2\}$ is the number of items desired by the bidder, and θ_i is the per-item value of the bidder. We assume w.l.o.g. that the per-item values are normalized to be in $[0, 1]$. The utility of a bidder is quasi-linear, i.e. if the bidder receives j items and pays a price p_i, her utility is

$$\begin{cases} -p_i & j < g_i \\ g_i \cdot \theta_i - p_i & otherwise. \end{cases}$$

In this paper, we focus on the analysis of the VCG mechanism in a setting where each bidder may submit multiple single-minded bids in an anonymous fashion so that VCG treats each bid as if it comes from a separate bidder. Formally, VCG is the mechanism that given bids $B = (\hat{\theta_1}, ..., \hat{\theta_N})$ (where $N \geq n$) allocates the items to maximize the social welfare:

$SW(B) = max_{I \subseteq \{1,...,N\} s.t. \sum_{i \in I} g_i \leq 2} \sum_{i \in I} g_i \cdot \theta_i$.

Let $W(B)$ be an index set I of bidders that attains the maximal social welfare, and let $L(B)$ be its complement. Every bid b makes the following payment:

$$\begin{cases} SW(B \setminus \hat{\theta_b}) - (SW(B) - g_b \cdot \theta_b) & b \in W(B) \\ 0 & b \in L(B). \end{cases}$$

Every real bidder $i \in \{1, ..., n\}$ receives her value from the union of items won by her submitted bids, and pays the respective sum of prices.

Throughout our analysis, given the symmetric situation, we fix one bidder, bidder n, and analyze her utility. This is done when she declares her true type and when she submits multiple false-name bids, assuming all other bidders reveal their true types. We term the case where the bidder submits false-name bids – an "attack". In our analysis, we refer to the other bidders as the *"adversary"* bidders. We introduce the notation $\tilde{n} = n-1$, the number of adversary bidders. We call the vector of adversary true types and the corresponding truthful bids *"adversary setup"*.

We refer to bidder i with $g_i = 1$ as a "1-type", and denote her value by v_i. Similarly, we refer to a bidder i with $g_i = 2$ as a "2-type", and denote her value w_i. We assume two separate indices, for 1-types and for 2-types. Formally, let $k = |\{1 \leq i \leq t, g_i = 1\}|$, the number of 1-type bids. The 1-type by $v_1, ..., v_k$ and the 2-type values are denoted by $w_1, ..., w_{t-k}$. Notice that the VCG mechanism implies that for any adversary setup, a bidder's utility is uniquely determined by

$$\begin{aligned} \tilde{v_1} &= \max\{v_i\}_{1 \leq i \leq k}, \\ \tilde{v_2} &= second - \max\{v_i\}_{1 \leq i \leq k}, \\ \tilde{w_1} &= \max\{w_i\}_{1 \leq i \leq t-k}, \end{aligned}$$

with $\max \emptyset = 0$. We therefore write for bidder n bidding $(\hat{\theta'_1}, ..., \hat{\theta'_m})$ and adversary setup $\hat{\theta}_1, \ldots, \hat{\theta}_{\tilde{n}}$,

$$u_{\hat{\theta'_1}, ..., \hat{\theta'_m}}(\hat{\theta}_1, \ldots, \hat{\theta}_{\tilde{n}}) = \tilde{u}_{\hat{\theta'_1}, ..., \hat{\theta'_m}}(\tilde{w_1}, \tilde{v_1}, \tilde{v_2})$$

. Also notice that if biddern n submits more than one 2-type bid, all but the top one never enters the winning set. The same holds for all but the top two 1-type bids submitted. Since in VCG adding more losing bids can only increase the price paid, we conclude that a bidder that wishes to maximize utility never submits more than one 2-type bid and two 1-type bids. In fact, we prove a stronger claim:

Lemma 1. *Under VCG, in our setting, for any attack of bidder n, $S = \hat{\theta_1}, ..., \hat{\theta_m}$, one of the two must be true:*

- *There exists an attack $(1, x), (1, y)$ such that for any adversary setup, n's utility from $(1, x), (1, y)$ is not lower than her utility from S, or,*
- *For any adversary setup, n's utility from the truthful bid is not lower than her utility from S.*

Proof. We already observed that a bidder prefers to bid $\tilde{w_1}, \tilde{v_1}, \tilde{v_2}$, some of which may be zero bids or equivalently omitted over bidding S. Now, if $2\tilde{w_1} \leq \tilde{v_1} + \tilde{v_2}$, regardless of the adversary bids, $(2, \tilde{w_1})$ is never in the winning set and thus could only increase the price for bidder n. This constitutes the first case. Otherwise, for any adversary bids either the higher 1-type bid enters the winning set, the 2-type bid enters it or no bids enter it, i.e., at most one of the bids $\tilde{w_1}, \tilde{v_1}, \tilde{v_2}$ win. In this case, it would have been better to bid truthfully by VCG truthfulness for single bids (Nisan et al. 2007). □

We subsequently regard a false-name bid attack as a bid vector $(1,x),(1,y)$, w.l.o.g. assuming $x \geq y$. Since the utility function u depends on the bidder's true type and her bids, we introduce the conditional utility notation $\tilde{u}_{\hat{\theta},truth}(\tilde{w_1},\tilde{v_1},\tilde{v_2})$ and $\tilde{u}_{\hat{\theta},attack(x,y)}(\tilde{w_1},\tilde{v_1},\tilde{v_2})$.

Examples

We adjust an example from (Yokoo, Sakurai, and Matsubara 2002). Assume the following two bidders' true types:
Bid 1 - (2, 0.5), Bid 2 - (2, 0.4)

When bidder 1 bids truthfully, she wins the two items and pays 0.8. Consider what happens if Bidder 1 splits her bid:

Bid 1 - (1, 0.5), Bid 1* - (1, 0.5), Bid 2 - (2, 0.4)

Bidder 1 still wins both items, but with a lower payment of 0.6.

The example is highly dependent on whether Bidder 1 knows her adversary bids. We can now see an example where the same false-name attack causes a significant loss to the attacker (double than her previous gain). Consider the same behavior for bidder 1 as in example 1, but with a different adversary:

Bid 1 - (2, 0.5), Bid 2 - (1, 0.4)

In this case bidder 1 wins the two items and pays 0.4. If bidder 1 splits her bid the same way as before —

Bid 1 - (1, 0.5), Bid 1* - (1, 0.5), Bid 2 - (1, 0.4)

Bidder 1 still wins both items, but now her payment is 0.8.

The above hints at the potential benefit of a Bayesian approach: one might not submit false-name bids as her potential losses might be higher than her potential gains.

A Bayesian approach

We assume each bidder knows her own type and a distribution over the i.i.d. parameterized $\tilde{n}$ adversary true types. We assume that the distribution of true types is given by $0 \leq q,(1-q) \leq 1$ which is the probability of a bidder i to have $g_i = 1,2$ respectively, and by a per-item valuation distribution $\theta_i \sim F$ which is independent of the bidder's g_i.

A bidder's expected utility given her true type $\hat{\theta}$, her bids vector S and q, F is:

$$E^{q,F,n}_{\hat{\theta},S}[u] = \sum_{k=0}^{\tilde{n}} \binom{\tilde{n}}{k} q^k (1-q)^{\tilde{n}-k} E^{k,F,n}_{\hat{\theta},S}[u],$$

where

$$E^{k,F,n}_{\hat{\theta},S}[u] = \int_{v_1=0}^{1} f(v_1) ... \int_{v_k=0}^{1} f(v_k)$$
$$\int_{w_1=0}^{1} f(w_1) ... \int_{w_{\tilde{n}-k}=0}^{1} f(w_{\tilde{n}-k}) u(v_1,...,w_{\tilde{n}-k}).$$

Notice that since Lemma 1 holds for any adversary bids, it immediately extends to expected utilities as well.

Definition 2.1. *For a set of parameters q, F, n, we say that truthfulness is a Bayesian Nash Equilibrium (BNE) if for any bidder's type $\hat{\theta}$, $\forall 0 \leq y \leq x \leq 1, \hat{\theta} \in \{1,2\} \times [0,1]$,*

$$E^{q,F,n}_{\hat{\theta},truth}[u] \geq E^{q,F,n}_{\hat{\theta},attack(x,y)}[u].$$

We say that an attack is "beneficial" for a given adversary setup if there exists a true type such that the attack increases a bidder's utility over her truthful bid. If we do not specify a concrete adversary setup, an attack is "beneficial" if it increases the bidder's expected utility.

The following technical lemma holds:

Lemma 2. *To prove that a truthful BNE exists under some settings, it suffices to analyze 1-type attackers that have $\theta = 1 \geq x, y$, and 2-type attackers that have $x = 1 \geq \theta, y$.*

A Criterion for Bayesian Resilience

We already noted that for $q = 1$, under any per-item valuation distribution F, and number of bidders n, we expect a truthful BNE, as it is basically equivalent to the single-item case where VCG is truthful in dominant strategies. The following technical lemma holds:

Lemma 3. *When all adversaries are 1-type an attack is never beneficial. Thus, $q = 1$ always induces a truthful BNE.*

Our main question is what is the minimal q value that guarantees Bayesian resilience, i.e. that truth-telling is in equilibrium.

Definition 2.2. *Define $q^*_{n,g,\theta,x,y,F} \in [0,1]$, the "granularity threshold", as the minimal q such that no true type (g,θ) bidder prefers to attack with $(1,x),(1,y)$ given that her adversaries' types are chosen with $q' > q^*_{n,g,x,y,F}$ and valuation distribution F. If any of the parameters are omitted, we assume the supremum of all different q^* values with different instantiations of the parameters. We call*

$$q^*_n = sup_{g\in\{1,2\},0\leq y\leq x\leq 1,F} q^*_{n,g,x,y,F}$$

the "global granularity threshold" for n.

The fact that VCG is truthful when there is only a single item, which is equivalent to $q = 1$, gives us the intuition that if we have a high enough q, then Bayesian resilience would be obtained. As q^* is lower, the distribution is more resilient to Sybil attacks.

3 A Full Characterization of the Global Granularity Threshold for $n = 2$

Theorem 3.1. $q^*_{n=2} = \frac{2}{3}$.

The remainder of this section proves this theorem.

Lemma 4. $q^*_{n=2,g=1} = \frac{1}{2}$.

Proof. By Lemma 2 we assume $y, x \leq \theta = 1$. For $n = 2$ we have $\tilde{n} = 1$ and so we have one adversary, either of type

1 or 2. We separate the utility function u to u^1, u^2, based on the adversary type. We then have

$$u^2_{(1,1),attack(x,y)}(w_1) =$$
$$\begin{cases} 1 & 0 \le w_1 \le \frac{y}{2} \\ 1-2w_1+y & \frac{y}{2} < w_1 \le \frac{x}{2} \\ 1-4w_1+y+x & \frac{x}{2} < w_1 \le \frac{x+y}{2} \\ 0 & otherwise, \end{cases}$$

$$u^1_{(1,1),attack(x,y)}(v_1) = \begin{cases} 1-2v_1 & 0 \le v_1 \le y \\ 1-y & y < v_1 \le 1, \end{cases}$$

$$u^2_{(1,1),truth}(w_1) = \begin{cases} 1-2w_1 & 0 \le w_1 \le \frac{1}{2} \\ 0 & otherwise, \end{cases}$$

$$u^1_{(1,1),truth}(v_1) = 1.$$

Define $\Delta(\theta_1) = u^2_{truth}(\theta_1) - u^2_{attack}(\theta_1) + u^1_{truth}(\theta_1) - u^1_{attack}(\theta_1)$. We have

$$\Delta(\theta_1) =$$
$$\begin{cases} 1-u^2_{attack}(\theta_1) \ge 0 & 0 \le \theta_1 \le y, \theta_1 \le \frac{1}{2} \\ 6\theta_1 - 1 - y - x \ge 0 & \frac{1}{2} < \theta_1 \le y \\ y - u^2_{attack}(\theta_1) + 1 - 2\theta_1 \ge 0 & y < \theta_1 \le \frac{1}{2} \\ 4\theta_1 - 1 - x \ge 0 & y, \frac{1}{2} < \theta_1 \le \frac{x+y}{2} \\ y \ge 0 & \frac{x+y}{2} < \theta_1, \frac{1}{2} < \theta_1. \end{cases}$$

This covers all the non-trivial cases. This yields for $q = \frac{1}{2}$,

$$E^{q=\frac{1}{2},F,n=2}_{\hat{\theta},truth} - E^{q=\frac{1}{2},F,n=2}_{\hat{\theta},attack(x,y)} = \frac{1}{2}E[\Delta(\theta_1)] \ge 0.$$

By Lemma 3,

$$E[u^1_{truth} - u^1_{attack(x,y)}] \ge 0.$$

Hence, for an attack to be beneficial it must hold that

$$E[u^2_{truth} - u^2_{attack(x,y)}] < 0.$$

So, the expression

$$E[(1-q)(u^2_{attack} - u^2_{truth}) + q(u^1_{attack} - u^1_{truth})]$$

is monotone increasing in q, and for all $q' > \frac{1}{2}$ it is non-positive. In addition, it is straightforward to construct an example that shows a beneficial attack for any $q < \frac{1}{2}$, based on the above expressions. We conclude that $q^*_{n=2,g=1} = \frac{1}{2}$. □

Lemma 5. $q^*_{n=2,g=2} = \frac{2}{3}$.

Proof. Similar to the one given in Lemma 4. □

4 Global Granularity Thresholds for $n > 2$: $q^*_n = 1$

Based on the result of $q^*_{n=2} = \frac{2}{3}$ for two bidders, we may now ask what are the q^* values for different values of n.

Theorem 4.1. *For any $n > 2, q < 1$, there is an attack x, y and distribution F such that the attack is beneficial.*

Proof. We construct F as a discrete distribution for the sake of a more concise argument. A continuous distribution close enough to F satisfies the same argument.

We examine an attacker of type 1. By Lemma 2 we write $\theta = 1$. We choose $x = 1, y = \frac{1}{2}$. We show how we choose $\epsilon(q,n)$ later, and we define $F_{\epsilon(q,n)}$ with $Pr(0.6) = 1 - 2\epsilon$, $Pr(0.5) = Pr(1) = \epsilon$.

We prove that the attack is beneficial for $q < 1, n > 2$. Consider the following possible adversary configurations. First we ignore the possibility that the adversary value is 0.5 (the probability for this is $(1-\epsilon)^{\tilde{n}}$):

There exists at least one adversary of type 2 with value 1: the utility for both attack and truthful bidding is 0.

There exist at least two adversaries of type 1: then y is not in the winning set and also doesn't affect the price. Thus, the utility for attack and truthful bidding is the same.

There exists exactly one adversary of type 1: Note that all other adversaries of type 2 have a value of 0.6. If the type 1 adversary has value 0.6, then attack and truthful bidding utility are both 0.4. If the type 1 adversary has value 1, then attack utility is 0.5 while truthful utility is 0.8.

No adversaries of type 1: we are left with only type 2 adversaries with value 0.6. In this case, the truthful utility is 0 and the attack utility is 0.1.

Define $\delta = \frac{(1-q)^{\tilde{n}}}{1000}$ and choose

$$\epsilon(q,n) = \min\left\{\frac{1-q}{3000q}, \frac{1}{10^3}, \frac{1-\sqrt[\tilde{n}-1]{\frac{1}{10}}}{2}, 1-\sqrt[n-1]{\frac{3}{3+\frac{(1-q)^{\tilde{n}}}{1000}}}\right\},$$

which satisfies

$$0.1(1-2\epsilon)^{\tilde{n}}(1-q)^{\tilde{n}} + (0.3)\tilde{n}(1-2\epsilon)^{\tilde{n}-1}\epsilon(1-q)^{\tilde{n}-1}q > \delta, \quad (1)$$

$$(1-\epsilon)^{\tilde{n}}\delta - 3(1-(1-\epsilon)^{\tilde{n}}) > 0. \quad (2)$$

Equation 1 implies that the expected utility of the attack, given that there is no 0.5 value adversary, is more than δ. The first summand refers to the fourth case we described and the second summand to the third case. Equation 1 implies that the overall expected utility is positive, where the constant 3 in the equation was chosen since an attack can decrease the utility relative to truthful bidding by at most 3: since the attacker is a 1-type, truthful bidding yields at most a utility of one, and an attack can lose at most 2.

This puts a positive lower bound on the difference between attack utility and truthful bid utility, and thus the attack is beneficial as defined. □

5 Beta Distributions Granularity Thresholds: a Computer Algebra Approach

Given the result of the previous section, we know that it is impossible to guarantee a truthful Bayesian equilibrium for arbitrary distributions when $q < 1$. Hence, in this section we instead consider a general family of distributions, namely the Beta distributions. This family of distributions

includes many natural distributions that are parameterized by two parameters α and β. We consider these parameter values to be integers. The probability density function of $Beta(\alpha,\beta)$ is then $f_{\alpha,\beta}(\theta_i) = (\alpha+\beta-1)!\frac{\theta_i^{\alpha-1}(1-\theta_i)^{\beta-1}}{(\alpha-1)!(\beta-1)!}$. Our aim is to find the corresponding granularity thresholds $q^*_{n,F=Beta(\alpha,\beta)}$.

We address this challenge as follows. We show that the difference between the expected utility of a 1-type truthful bidder and the expected utility of an attacker is a polynomial P (in x,y,q). Similarly, we show that the difference between the expected utility of a 2-type truthful bidder and the expected utility of an attacker is a polynomial Q (in θ,y,q). The minimal q for which in the domain $[0,1]\times[0,1]\times[q,1]$ these polynomials are positive is in fact $q^*_{n,F=Beta(\alpha,\beta)}$. This reduces the problem into a computer algebra problem, given parameters n,α,β which can be solved using state-of-the-art techniques. Our results show that indeed the desired granularity thresholds for the tested Beta distributions' parameter values are spread around 0.5.

A reduction to a polynomials positivity decision problem

Let $I_{\theta_i}(\alpha,\beta)$ be the cumulative distribution function of $Beta(\alpha,\beta)$ at θ_i. We develop the Bayesian expectation utility expressions.

$$E^{k,Beta(\alpha,\beta),n}_{\hat{\theta},*}[u] =$$

$$\int_{\theta_1=0}^{1}\dots\int_{\theta_n=0}^{1}\prod_{t=1}^{\tilde{n}} f_{\alpha,\beta}(\theta_t)u(\theta_1,\dots,\theta_{\tilde{n}})d\theta_{\tilde{n}}\dots d\theta_1 =$$

$$\sum_{1\le i\ne j\le k}\ \sum_{1\le m\le\tilde{n}-k}\int_{\theta_1=0}^{1}\dots\int_{\theta_{\tilde{n}}=0}^{1}\prod_{t=1}^{\tilde{n}} f_{\alpha,\beta}(\theta_t)\tilde{u}(\theta_m,\theta_i,\theta_j)$$

$$\mathbb{1}_{g_m=1,\forall p,g_p=2\to\theta_m\ge\theta_p,g_i=g_j=1,\forall p,g_p=1\to\theta_i\ge\theta_j\ge\theta_p}d\theta_{\tilde{n}}\dots d\theta_1$$

$$= \frac{k!}{(k-2)^+!}\frac{(\tilde{n}-k)!}{(\tilde{n}-k-1)^+!}\underbrace{\int_{w_1=0}^{1}\int_{w_2=0}^{w_1}\dots\int_{w_{n-k}=0}^{w_1}}_{n-k}$$

$$\underbrace{\int_{v_1=0}^{1}\int_{v_2=0}^{v_1}\int_{v_3=0}^{v_2}\dots\int_{v_k=0}^{v_2}}_{k}\prod_{t=1}^{n} f_{\alpha,\beta}(\theta_t)$$

$$\tilde{u}(w_1,v_1,v_2)dv_k\dots dv_1 dw_{\tilde{n}-k}\dots dw_1 =$$

$$\frac{k!}{(k-2)^+!}\frac{(\tilde{n}-k)!}{(\tilde{n}-k-1)^+!}\int_{\tilde{v_1}=0}^{1}\int_{\tilde{v_2}=0}^{\tilde{v_1}}\int_{\tilde{w_1}=0}^{1}$$

$$f_{\alpha,\beta}(\tilde{w_1})f_{\alpha,\beta}(\tilde{v_1})f_{\alpha,\beta}(\tilde{v_2})I_{\tilde{w_1}}(\alpha,\beta)^{(n-k-1)^+}$$

$$I_{\tilde{v_2}}(\alpha,\beta)^{(k-1)^+}\tilde{u}(\tilde{w_1},\tilde{v_2},\tilde{v_1})d\tilde{w_1}d\tilde{v_2}d\tilde{v_1}.$$

Recall that $\frac{k!}{(k-2)^+!} = \begin{cases}1 & k=0,1\\ k(k-1) & k=2,\dots,\tilde{n}\end{cases}$, $\frac{(\tilde{n}-k)!}{(\tilde{n}-k-1)^+!} = \begin{cases}1 & k=\tilde{n}\\ \tilde{n}-k & k=0,\dots,\tilde{n}-1\end{cases}$.

We then have

$$P^n_{\alpha,\beta}(x,y,\theta{=}1,q){=}E^{q,Beta(\alpha,\beta),n}_{(1,1),truth}[\tilde{u}]{-}E^{q,Beta(\alpha,\beta),n}_{(1,1),attack(x,y)}[\tilde{u}] =$$

$$\sum_{k=0}^{\tilde{n}}\binom{\tilde{n}}{k}q^k(1-q)^{\tilde{n}-k}P^{k,n}_{\alpha,\beta}(x,y,\theta=1),$$

$$Q^n_{\alpha,\beta}(x{=}1,y,\theta,q){=}E^{q,Beta(\alpha,\beta),n}_{(2,\theta),truth}[u]{-}E^{q,Beta(\alpha,\beta),n}_{(2,\theta),attack(1,y)}[u] =$$

$$\sum_{k=0}^{\tilde{n}}\binom{\tilde{n}}{k}q^k(1-q)^{\tilde{n}-k}Q^{k,n}_{\alpha,\beta}(x=1,y,\theta),$$

with

$$P^{k,n}_{\alpha,\beta}(x,y,\theta=1) = E^{k,Beta(\alpha,\beta),n}_{(1,1),truth}[u]{-}E^{k,Beta(\alpha,\beta),n}_{(1,1),attack(x,y)}[u],$$

$$Q^{k,n}_{\alpha,\beta}(x=1,y,\theta) = E^{k,Beta(\alpha,\beta),n}_{(2,\theta),truth}[u]{-}E^{k,Beta(\alpha,\beta),n}_{(2,\theta),attack(1,y)}[u].$$

Notice that $P^{k,n}_{\alpha,\beta}$ and $Q^{k,n}_{\alpha,\beta}$ are polynomial expressions for any α,β,n and $0\le k\le\tilde{n}$. That is since the probability density $f_{\alpha,\beta}(\theta_i)$ is polynomial in its variable, and the incomplete beta distribution function $I_{\theta_i}(\alpha,\beta)$ is polynomial in its parameter. Also, the full form of the $\tilde{u}_{\hat{\theta},Bids}(\tilde{w_1},\tilde{v_1},\tilde{v_2})$ are expressions which are multiples of polynomials in $\tilde{w_1},\tilde{v_1},\tilde{v_2},x,y,\theta$ and indicator functions with bounds polynomial in these parameters. We can thus rewrite the triple integral as a linear combination of triple integrals of the following form:

$$\int_{\tilde{v_1}=T_1(x,y,\theta)}^{T_2(x,y,\theta)}\int_{\tilde{v_2}=T_3(\tilde{v_1},x,y,\theta)}^{T_4(\tilde{v_1},x,y,\theta)}$$

$$\int_{\tilde{w_1}=T_5(\tilde{v_1},\tilde{v_2},x,y,\theta)}^{T_6(\tilde{v_1},\tilde{v_2},x,y,\theta)}T_7(\tilde{w_1},\tilde{v_1},\tilde{v_2},x,y,\theta)d\tilde{w_1}d\tilde{v_2}d\tilde{v_1},$$

with all T_i polynomials. Since the class of Polynomials is closed under integration and composition, we have that all $P^{k,n}_{\alpha,\beta},Q^{k,n}_{\alpha,\beta}$ are polynomial in x,y,θ. By the definition of $P^n_{\alpha,\beta}(x,y,\theta=1,q),Q^n_{\alpha,\beta}(x,y,\theta=1,q)$ as linear combinations of polynomials in q multiplied by the k-parameterized polynomials, they are also polynomials.

Solving polynomial positivity using computer algebra

The question of whether truthfulness is a Bayesian Nash equilibrium is now reduced to whether there are no assignments (a true type and attack bids) where the polynomial is negative in the domain. For this purpose, any computer algebra method that is able to prove polynomial positivity suffices. The method we found most useful for our setting is Partial Cylindrical Algebraic Decomposition (Caviness and Johnson 2012). Given a guess of the threshold value q^*_guess for a given $Beta(\alpha,\beta)$ distribution and n value, we feed Maple (Chen and Moreno Maza 2016) with the following problems:

$$PartialCylindricalAlgebraicDecomposition$$
$$(P^{q,n}_{\alpha,\beta}(x,y,1),[x-y,y,1-q,1-x,q-q^*\_guess],$$
$$PolynomialRing([x,y,q]),$$

$PartialCylindricalAlgebraicDecomposition$
$(Q^{q,n}_{\alpha,\beta}(1, y, theta), [v - theta, y, 1 - q, q - q^*\_guess],$
$PolynomialRing([theta, y, q]))$.

Partial CAD finds representing points in the Cylindrical Algebraic Decomposition. It then suffices to check each of the representing points to be positive for the corresponding polynomial, in order to verify that it does not attain negative values in the entire domain. The guess q^*_guess is derived either from a bisection of the $[0, 1]$ interval, or from the split attack on which we elaborate in the next section.

Our findings are given in the following figures. As can be seen, the granularity thresholds in all cases we tested are bound away from 1, in particular, lower than 0.75.

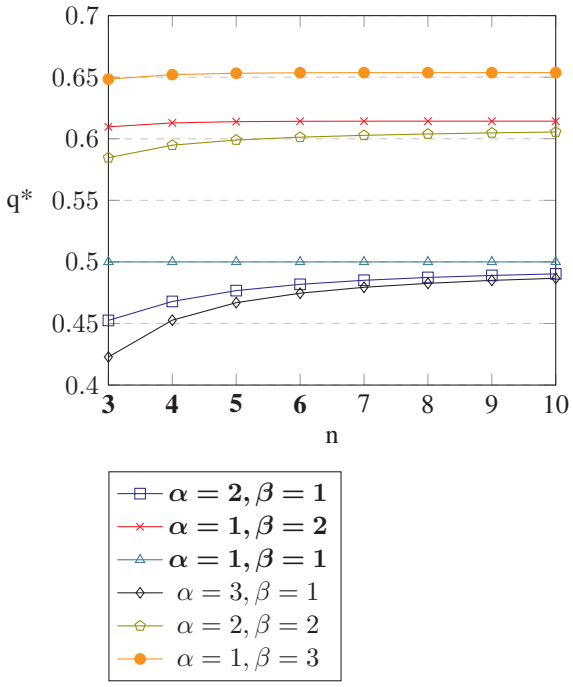

Figure 1: q* values for beta distribution parameter values

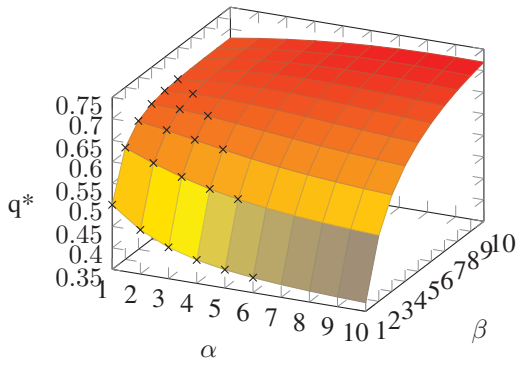

Figure 2: q* values for beta distributions n=3

Some of the symbols in the figures (in particular, some of the n's in the x-axis and some of the α's and β's in the legend) are in bold while others are not. Bold symbols represent cases that were fully verified by the partial CAD method described above, and they hold for all possible attacks. In the three-dimensional figure, the fully verified cases are those marked with x symbol. For these fully verified cases, we observe that the granularity threshold is typically decreasing (improving) as a function of α and increasing as a function of β. Interestingly, in all fully verified cases, the attack most persistent in respect to higher values of q ("best attack") was the split attack. We therefore extend the figures to include q^* values that were evaluated only for the split attack – these are the non-bold parameter values in the figures.

The Mathematica and Maple files to attain and verify the figures' exact values independently can be found at https://github.com/yotam-gafni/vcg_bayesian_fnp.

6 Split Attacks

As mentioned, all the q^* values given in the figures of the last section have a shared property — they are in fact all derived from one attack, which proves itself in all verified cases to be the one that yields the highest q* values. This attack is the split attack — where a 2-type attacker with value 1 splits her bid into two 1-type bids $x = y = 1$. In this section, we focus our attention on the split attack, which gives us a few advantages. First, the problem formulation for the CAD becomes uni-variable, which allows us to computationally reach higher parameter values. Also, we are able to prove analytic results for general n values. We establish a few lemmas that yield such a general result for the uniform distribution case.

Lemma 6. *For the uniform distribution, $Q^{k,n}_{1,1}(1, 1, 1)$ is monotone increasing as a series in k.*

Proof. By direct calculation we have

$$F(\tilde{n}, k) = Q^{k,n}_{1,1}(1, 1, 1) =$$
$$\begin{cases} \frac{1}{(\tilde{n}+1)2^{\tilde{n}-1}} - \frac{2}{\tilde{n}+1} & k = 0 \\ \frac{8}{\tilde{n}(\tilde{n}+1)} - \frac{2}{\tilde{n}} - \frac{3}{\tilde{n}(\tilde{n}+1)2^{\tilde{n}-1}} & k = 1 \\ \frac{1}{\tilde{n}+1} & k = \tilde{n} \\ \frac{2k}{2^{\tilde{n}-k}} \sum_{i=0}^{\tilde{n}-k} \binom{\tilde{n}-k}{i} (\frac{1}{i+k} + \frac{1}{i+k+1}) + & \\ \frac{2k(k-1)(\tilde{n}-k)}{(\tilde{n}+1)(\tilde{n}-k+1)2^{\tilde{n}-k+1}} \sum_{i=0}^{\tilde{n}-k+1} \binom{\tilde{n}-k+1}{i} \frac{1}{\tilde{n}-i} - & \\ \frac{2}{\tilde{n}} - \frac{4k(\tilde{n}-k)}{(\tilde{n}-k+1)2^{\tilde{n}-k+1}} \sum_{i=0}^{\tilde{n}-k+1} \binom{\tilde{n}-k+1}{i} \frac{1}{i+k} - & \\ \frac{k(k-1)}{(\tilde{n}+1)2^{\tilde{n}-k}} \sum_{i=0}^{\tilde{n}-k} \binom{\tilde{n}-k}{i} (\frac{1}{\tilde{n}-i} + \frac{1}{\tilde{n}-i-1}) & o/w. \end{cases}$$

We show that for any $\tilde{n} \geq 2, 0 \leq k \leq \tilde{n} - 1$,

$$F(\tilde{n}, k) < F(\tilde{n}, k + 1).$$

It can be directly verified that $F(\tilde{n}, 0) < F(\tilde{n}, 1) < F(\tilde{n}, 2)$ for every $\tilde{n} \geq 2$, and also that for $\tilde{n} = 2$ the series is monotone. For the $\tilde{n} \geq 3, 2 \leq k \leq \tilde{n} - 1$ case, we can define f

$$G[\tilde{n}, k] = F[\tilde{n}, k + 1] - F[\tilde{n}, k],$$

for which the following recurrences hold:

$$(2\tilde{n} - 2k)G[\tilde{n}, k] + (-4 + k - 3\tilde{n})G[1 + \tilde{n}, k] + (3 + \tilde{n})G[2 + \tilde{n}, k] = 0, \quad (3)$$

$$(2+\tilde{n})G[1+\tilde{n},\tilde{n}] = 2\tilde{n}G[\tilde{n},\tilde{n}-1], \quad (4)$$

$$(\tilde{n}^2+4\tilde{n}+3)G[\tilde{n}+2,\tilde{n}] = (2\tilde{n}^2+4\tilde{n})G[\tilde{n}+1,\tilde{n}-1]. \quad (5)$$

We now show that $\forall \tilde{n} \geq 3, 2 \leq k \leq \tilde{n}-1, G[\tilde{n},k] > 0$. For that, we denote $m = \tilde{n} - k$. It's enough to show that for any $m \geq 1, k \geq 2$,

$$G[k+m+1,k] > G[k+m,k] > 0.$$

We prove by induction on m. First we show the induction step. Assume for some m the assumption holds. Then by the recurrence relation (3)

$$G[k+m+2,k] =$$
$$\frac{(3m+2k+4)\cdot G[k+m+1,k] - 2m\cdot G[k+m,k]}{3+k+m} >$$
$$\frac{(m+2k+4)\cdot G[k+m+1,k]}{3+k+m} \geq G[k+m+1,k] \geq 0.$$

As for the base case $G[\tilde{n}+1,\tilde{n}-1] > G[\tilde{n},\tilde{n}-1] > 0$, we prove by induction on $\tilde{n}$. For $\tilde{n}=3$ it holds. Now assume for some $\tilde{n}$ the induction assumption holds, then

$$G[\tilde{n}+2,\tilde{n}] = \frac{2\tilde{n}^2+4\tilde{n}}{(\tilde{n}^2+4\tilde{n}+3)}G[\tilde{n}+1,\tilde{n}-1] >$$
$$\frac{2\tilde{n}^2+4\tilde{n}}{(\tilde{n}^2+4\tilde{n}+3)}G[\tilde{n},\tilde{n}-1] = \frac{(2+\tilde{n})(2\tilde{n}^2+4\tilde{n})}{2\tilde{n}(\tilde{n}^2+4\tilde{n}+3)}G[\tilde{n}+1,\tilde{n}] >$$
$$G[\tilde{n}+1,\tilde{n}] = \frac{2\tilde{n}}{2+\tilde{n}}G[\tilde{n},\tilde{n}-1] > 0$$

holds by (5), induction assumption, (4), arithmetics, (4) and induction assumption respectively. □

The recurrences in equations 3 to 5 were found using RISCErgoSum's Guess package (Kauers 2009).

Lemma 7. *Recall that*

$$Q^{k,n}_{\alpha,\beta}(1,1,1) = E^{k,Beta(\alpha,\beta),n}_{(2,1),truth}[u] - E^{k,Beta(\alpha,\beta),n}_{(2,1),attack(1,1)}[u],$$

and regard it as a series of real numbers parameterized by k. If the series is monotone increasing in k, then if there exists q such that $Q^n_{\alpha,\beta}(1,1,1,q) = 0$, it follows that $q^*_{n,g=2,\theta=1,x=1,y=1,F=Uni([0,1])} = q$.

Proof. It suffices to prove that $\forall q' < q, Q^n_{\alpha,\beta}(1,1,1,q') < 0$ and $\forall q' > q, Q^n_{\alpha,\beta}(1,1,1,q') > 0$. This is due to first order stochastic dominance of a binomial distribution with a higher q parameter over another binomial distribution with a lower q parameter (Wolfstetter 1999). □

Theorem 6.1. *For the uniform distribution, and any number of bidders $n \geq 3$, $q^*_{n,g=2,\theta=1,x=1,y=1,F=UNI([0,1])} = \frac{1}{2}$.*

Proof. We examine the expressions $F(\tilde{n},k)$ derived before and notice that when $q = \frac{1}{2}$ the sum

$$S[\tilde{n}] = \sum_{k=2}^{\tilde{n}-1}\binom{\tilde{n}}{k}q^k(1-q)^{\tilde{n}-k}F(\tilde{n},k)$$

satisfies the recurrence

$$-2(1+\tilde{n})S[\tilde{n}]+(2+\tilde{n})S[1+\tilde{n}] = -7+32^(1-\tilde{n})+2\tilde{n}, \quad (6)$$

.

We solve the recurrence and check initial $\tilde{n}$ values. We then see that indeed

$$S[\tilde{n}] = -\frac{F(\tilde{n},0)+\tilde{n}F(\tilde{n},1)+F(\tilde{n},\tilde{n})}{2^{\tilde{n}}},$$

which means $\forall n \geq 3$,

$$Q^n_{\alpha,\beta}(1,1,1,\frac{1}{2}) = \frac{1}{2^{\tilde{n}}}\sum_{k=0}^{\tilde{n}}\binom{\tilde{n}}{k}F(\tilde{n},k) =$$
$$S[\tilde{n}] + \frac{F(\tilde{n},0)+\tilde{n}F(\tilde{n},1)+F(\tilde{n},\tilde{n})}{2^{\tilde{n}}} = 0.$$

By Lemma 7 and Lemma 6 this proves that

$$\forall n \geq 3, q^*_{n,g=2,\theta=1,x=1,y=1,F=UNI([0,1])} = \frac{1}{2}.$$

□

We used RISCErgoSum's HolonomicFunctions Mathematica package to derive and solve the recurrence in equation 6 (Koutschan 2009).

7 Conclusions and Future Directions

In this work, we investigate the most simple yet general model for a Bayesian analysis of VCG under Sybil attacks. Interestingly, our results imply false-name resistance can be obtained without sacrificing social welfare in many cases, which we capture using the notion of granularity threshold. While we were able to show that the split attack is in a sense the best false-name bid attack in our study, an interesting open question is under what conditions does the split attack yield the highest $q*$ values. Can a formula be derived for the asymptotic behavior of $q*$ as a function of α, β for the beta distributions, or for specific n values? Can we find larger classes of distributions that admit a nice analysis? Can we exploit approximations of other general distributions by Beta (or more generally polynomial) distributions to yield precise bounds for their respective q^* values?

In section 3 we restrict attention to the case $F := F_1 = F_2$, i.e., the same per-item value distribution for different item demand types, which allows for an elegant analysis. If they are bound together by a looser correlation or stochastic dominance criterion, a result in the same spirit might still be attainable, even though it's not clear in what form.

It is possible given the computer algebra methods for the beta distribution to discuss a larger number of items. In the two item case we were able to prove Lemma 1 that shows a bidder has two alternatives - to submit a truthful bid or submit a pair $(1,x),(1,y)$ with $0 \leq y \leq x \leq 1$. This results in a tri-variate polynomial with q,x,y. With more items (e.g., 3) one might consider other scenarios: The bidder submits (3,x), or $(2,x),(1,y)$, etc - but still a small finite space that depends on the number of items. For each case one derives a polynomial as done in section 5, and using the computer algebra solvers can come up with analytic results. Though computationally more demanding, in theory this could extend to any number of items. Notice that in the two item

case, there is a single number q that measures 'granularity' by quantifying the probability of a 1-item demand bidder to appear. Once we analyze 3 items or more, one needs choose a finer way of defining granularity - for example, we could say that having the probability vector (0.4,0.5,0.1) over the amounts (1,2,3) of item demand is more granular than (0.4,0.1,0.5), even though the "q" value as defined for both is 0.4.

8 Acknowledgements

We wish to thank Doron Zeilberger and Christoph Koutschan for their kind advise regarding symbolic identity proving. We also wish to thank Noam Neer for his advise regarding polynomial positivity proof techniques.

Yotam Gafni and Moshe Tennenholtz were supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant No. 740435).

Ron Lavi was partially supported by the ISF-NSFC joint research program (grant No. 2560/17).

References

Alkalay-Houlihan, C., and Vetta, A. 2014. False-name Bidding and Economic Efficiency in Combinatorial Auctions. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*, AAAI'14, 538–544. AAAI Press.

Caviness, B. F., and Johnson, J. R. 2012. *Quantifier Elimination and Cylindrical Algebraic Decomposition*. Springer Science & Business Media.

Chen, C., and Moreno Maza, M. 2016. Quantifier Elimination by Cylindrical Algebraic Decomposition Based on Regular Chains. *J. Symb. Comput.* 75(C):74–93.

Gupta, A. K., and Nadarajah, S. 2004. *Handbook of Beta Distribution and its Applications*. CRC press.

Iwasaki, A.; Conitzer, V.; Omori, Y.; Sakurai, Y.; Todo, T.; Guo, M.; and Yokoo, M. 2010. Worst-case Efficiency Ratio in False-name-proof Combinatorial Auction Mechanisms. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: Volume 1 - Volume 1*, AAMAS '10, 633–640. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems.

Kauers, M. 2009. *Guessing Handbook*. Citeseer.

Koutschan, C. 2009. *Advanced Applications of the Holonomic Systems Approach*. Ph.D. Dissertation, RISC, Johannes Kepler University, Linz, Austria.

Krishnamoorthy, K. 2016. *Handbook of Statistical Distributions with Applications*. Chapman and Hall/CRC.

Nisan, N.; Roughgarden, T.; Tardos, E.; and Vazirani, V. V. 2007. *Algorithmic Game Theory*. New York, NY, USA: Cambridge University Press.

Wolfstetter, E. 1999. *Topics in Microeconomics: Industrial Organization, Auctions, and Incentives*. Cambridge University Press.

Yokoo, M.; Sakurai, Y.; and Matsubara, S. 2001. Robust Combinatorial Auction Protocol against False-name Bids. *Artificial Intelligence* 130(2):167 – 181.

Yokoo, M.; Sakurai, Y.; and Matsubara, S. 2002. The Effect of False-Name Bids in Combinatorial Auctions: New Fraud in Internet Auctions. *Games and Economic Behavior* 46:174–188.

Yokoo, M. 2003. Characterization of Strategy/False-name Proof Combinatorial Auction Protocols: Price-oriented, Rationing-free Protocol. In *Proceedings of the 18th International Joint Conference on Artificial Intelligence*, IJCAI'03, 733–739. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.