Preventing Coercion in E-Voting: Be Open and Commit

Masoud Tabatabaei, Wojciech Jamroga, and Peter Y. A. Ryan

Interdisciplinary Centre for Security, Reliability and Trust University of Luxembourg

Abstract

Coercion resistance is one of the most important features of a secure voting procedure. Recently several coercion resistant voting schemes have been introduced, and some measures for quantifying the level of the coercion resistance in a voting protocol were defined. Also the relationship between coercion resistance and privacy in elections has been discussed in the literature.

In this paper, we take a new approach to coercion resistance from the point of view of an honest election authority that chooses between various procedures with different levels of resistance and different implementation costs. We give a game-based model for an election in which the authority chooses its strategy against a set of coercers. We also propose a preliminary analysis of the game in its simplest variant, namely one where only a single coercer is present. It turns out that, in these simple games, Stackelberg equilibrium is always more attractive to the society than Nash equilibrium. This suggests that the society is better off if the security policy is made open and public, and the authorities commit to it.

1 Introduction

It was recognised early on in the history of voting that ballot privacy is an essential property of voting systems to counter threats of coercion or vote buying. More recently, cryptographers and security experts have been looking at using cryptographic mechanisms to provide voter-verifiability, i.e. the ability for voters to confirm that their votes are correctly registered and counted. It strengthens to integrity properties, but, if it is not done carefully can introduce new threats to ballot secrecy. This lead to the introduction of more sophisticated privacy style notions: receipt-freeness and coercion-resistance. The later is the strongest property and can be defined informally as: a voting system provides coercion-resistance if the voter always has a strategy to vote as they intend while appearing to comply with all the coercer's requirements. The coercer is assumed to be able to interact with the voter throughout the voting process: before, during and after.

After [3], we accept the following meaning of the fundamental concepts:

Privacy: The system cannot reveal how a particular voter voted. Thus, privacy guarantees that the link between a voter and her vote remains secret.

Receipt-freeness: The voter does not gain any information (a receipt) which can be used to prove to a coercer that she voted in a certain way.

Coercion-resistance: The voter cannot cooperate with a coercer to prove to him that she voted in a certain way. Coercion resistance requires that the coercer cannot become convinced of how the voter has voted, even if the voter cooperates with him.

Achieving coercion-resistance is extremely challenging, especially in the context of internet and remote voting (e.g. postal). A number of schemes have been proposed that provide it, but typically this comes at a cost, in particular in terms of usability. In this paper we take a game theoretic approach to analyse the trade-offs between the costs of implementing coercion-resistance mechanisms on the one hand, and the cost to the society in terms of threats to the legitimacy of the outcome due to coercion attacks.

Unlike most existing papers, we do not propose a new coercion-resistant voting scheme, nor prove that a scheme is secure in that respect. Instead, we focus on the context of coercion attempts in e-voting, namely costs and benefits of involved parties. The main question is: Should the society invest in coercion-resistant procedures, and if so, in what way? We do not aim at devising a secure voting procedure, but rather at exposing conditions under which security of a procedure is relevant at all.

This work is still at a preliminary stage. We do not represent privacy explicitly, and we do not investigate its relation to coercion. Instead, we model the coercion resistance level as a simple scalar, usually indicating how much effort/cost it would take to break it. Also, our game-theoretic models are extremely simple: the society's interests are represented by a single agent that we call the "election authority", there is only one coercer in the game, and the structure of the game (i.e., strategies and their outcomes) are common knowledge among the participants. In fact, this is more of a starting point than a complete study, but we believe that it offers some interesting insights even at that stage.

1.1 Related Work

Related work can be roughly divided into 3 strands: definitions of the concept of coercion resistance (and its relation to privacy), proposals of coercion-resistant voting procedures, and studies of the context of coercion resistance. The notion of coercion-resistance was first introduced in [6]. In [3], a formalization of coercion resistance was proposed, and its relation to receipt-freeness and privacy was studied. [4] gave a formal definition of coercion resistance for the end-to-end voting schemes. In [7], a game-based cryptographic definition of coercion

resistance was proposed.¹ The same authors added a game-based cryptographic definition of privacy in [8], and showed that the relationship between privacy and coercion resistance can be more subtle than it is normally assumed. Finally, [5] proposed a general framework for specifying different coercion resistance and receipt freeness properties. The framework uses the process algebra CSP and is general enough to fit a wide range of definitions and properties given in the literature for coercion resistance.

The second strand overlaps with the first: [6, 4] all propose voting protocols that satisfy their definitions of coercion resistance while [7] proves coercion resistance of two previously existing protocols. Another coercion-resistant voting scheme was introduced – and proved – in [1]. Several other papers proposed voting schemes which provably satisfy privacy as an intuitive argument for coercion-resistance, cf. e.g. [11].

Putting coercion resistance in a broader economic or social context has been, to our best knowledge, largely left untouched. The only paper in this strand that we are aware of is [2]. The authors compare two voting systems using game models, more precisely zero-sum two player games based on attack trees. Two actions are available for the attacker (performing the attack or not); the authority is presumably choosing one of the two voting systems. The utility of the attacker is the expected probability of successful coercion minus the expected probability of being caught. The value is computed for the two systems using empirical data. In contrast, we consider a more general game where coercion—and resistance measures—come at a cost (instead of simply assuming probability distributions for the possible events), and we look for the rational choices of the players using game-theoretic solution concepts. We also argue that the coercion game is not zero-sum, with important consequences for the best policy to be chosen.

1.2 Structure of the Paper

In this paper, we propose a preliminary game-theoretic analysis in coercion resistance in voting. The main idea is to model the voting process as a noncooperative game [10] between the society and coercers. We begin by proposing a relatively general model of possible actions and incentives in Section 2. Then, in Section 3, we assume that there is only one coercer, which is a gross simplification but lets us avoid reasoning about the complex interaction between different coercion attempts. Under this assumption, we look at two specific instances of the model: one where a perfect coercion-resistance policy is available, and another one where all coercion resistance measures can be overcome by coercers if they invest sufficient resources.

Our analysis is based on two game-theoretic solution concepts: Nash equilibrium [9] and Stackelberg equilibrium. Nash equilibrium corresponds to the

¹The definition was game-based in the technical sense, i.e., the security property was defined as the outcome of an abstract game between the "verifier" and the "adversary" (similar e.g. to game semantics of mathematical logic). In this paper, we use game models to study the interaction between the actual participants of the protocol.

behavior of players that should emerge "organically" when they adapt to the behavior observed from the other players over a period of time. Stackelberg equilibrium [13], which has become very popular in anti-terrorist security recently [12], corresponds to a scenario when a designated "leader" commits openly to a selected strategy and thus forces the response from the other players. In all the games that we consider, Stackelberg equilibria for the society turn out to be different and more beneficial than Nash equilibria. From this, we conclude that the authority representing the society should not adapt its policy to what it expects from the coercer. Instead, it should commit to its Stackelberg policy. Moreover, if the authority announces its commitment publicly, it forces the coercer to give up coercion attempts. This is because the coercer knows now that coercion has become unprofitable.

2 Coercion as Noncooperative Game

We begin by presenting our game-based framework for modeling coercion games. The framework is still preliminary; in particular, we do not give any account of how actions from different coerces may interfere. This is why we will use only models with one coercer in our analysis of coercion resistance policies in Section 3. In a way, the framework presented in this section is more a survey of postulates on game models of coercion, rather than a concrete model.

2.1 Players and Strategies

Having a set of election candidates $\Omega = \{\omega_1, ..., \omega_g\}$ and a set of n voters, we model the election as a strategic game $\langle \{A, C_1, ..., C_c\}, (\Sigma_A, \Sigma_1, ..., \Sigma_c), (u_A, u_1, ..., u_c) \rangle$, with the following ingredients.

Players. A and C_1 to C_c are the players in the game. The player A is an honest election authority on behalf of the society. We assume that the goal of this player is in line with what we may call "the good of the society" as a whole. A has no preference for any of the candidates, and tries to make the result of the election as similar as possible to the result of the election without any coercion, i.e when the voters vote only on basis of their own preferences over the candidates.

Players C_1, \ldots, C_c represent potential coercers. These players may try to change the result of the election by threatening or bribing voters in order to make them vote based on the coercer's plan, rather than the voters' own preferences over the candidates.

Note that we do not represent the voters explicitly as players in the game. Their interests are globally represented by the preferences of A.

Actions and strategies. $\Sigma_A = \{\alpha_0, ..., \alpha_{Max}\}$ is a set of privacy methods available to be implemented by the election authority A. It is assumed that α_0 represents the case of no privacy. We assume that each α_i is more costly to implement (but also more costly to break) than α_{i-1} . Moreover, we will consider

two interpretations of α_{Max} : one which guarantees that coercion cannot happen (perfect resistance) and another which does not guarantee that (imperfect resistance).

 $\Sigma_i = \{b_0^i, ..., b_{n_i^*}^i, ... b_{n_i}^i\}$ is the set of actions for the coercer C_i . Each action specifies the number of voters that the coercer attempts to bribe. Thus, b_t^i represents the attempt of coercer C_i to bribe t voters. The minimal number of voters that the coercer C_i needs to bribe in order to change the result of the election in his favour is n_i^* , where $0 \le n_i^* \le n_i$. Note that coercing more than n_i^* voters only increases the cost of coercion without increasing its effectivity, and it is always strongly dominated for the coercer by $b_{n_i^*}^i$. Therefore, when analyzing the model, we will usually omit actions $b_{n_i^*+1}^i, \ldots, b_{n_i}^i$.

Besides pure strategies (i.e., actions), a player can decide to select a *mixed* strategy, that is, a randomized choice represented by a probability distribution over actions. We will denote strategies of player i (be it pure or mixed) by s_i . A strategy profile is a combination of strategies from all the players (one per player). A strategy profile consisting of only pure strategies is sometimes called an action profile.

2.2 Preferences

Preferences are represented by utility functions over possible combinations of strategies. For each player, their utility combines several factors. The utility for the election authority A is defined as $u_A(\sigma) = v_A(out(\sigma)) - imp(\alpha_j) - \beta_A \sum_{t=1}^c k_t$, where:

- $\sigma = (\alpha_j, b_{k_1}^1, ... b_{k_c}^c) \in \Sigma = \Sigma_A \times \Sigma_1 \times ... \times \Sigma_c$ is an action profile, chosen by the players in the game.
- $out: \Sigma \to \Omega$ gives the result of the election.
- $v_A: \Omega \to \mathbb{R}$ gives the value of the election for player A. We assume that $v_A(out(\alpha_j, b^1_{k_1}, ..., b^t_{k_t}, ..., b^c_{k_c})) \leq v_A(out(\alpha_j, b^1_{k_1}, ..., b^t_{k_{t'}}, ..., b^c_{k_c}))$ if $k_{t'} < k_t$. It means that bribing one more voter can only decrease the value of the election for player A. Moreover, $v_A(out(\sigma)) = v_A^*$ when the result of the election is the same as its result without coercion, and $v_A(out(\sigma)) = v_A^* \epsilon_A$ for some $\epsilon_A > 0$ when the result of the election has changed because of coercion.
- $imp: \Sigma_A \to \mathbb{R}$ is the implementation cost function. $imp(\alpha_j)$ shows the cost of applying the privacy method α_j for the player A. It is assumed that $imp(\alpha_0) = 0$, and $imp(\alpha_t) \leq imp(\alpha_{t'})$ if t < t'.
- The term $\beta_A \Sigma_{t=1}^c k_t$ represents the "corruption damage" to the society due to coercion attempts. We require that $\beta_A \Sigma_{t=1}^c n_i < \sigma_A$. In other words, the society always prefers unsuccessful coercion over successful coercion, and no coercion attempts over unsuccessful coercion.

 $^{^{2}}$ In what follows, we use terms "coerce" and "bribe" interchangeably.

Likewise, the utility function for the coercer C_i is defined as $u_i(\sigma) = v_i(out(\sigma)) - (d_i(\alpha_i) + \beta_i) \cdot k_i$, where:

- $\sigma = (\alpha_j, b_{k_1}^1, ...b_{k_c}^c) \in \Sigma = \Sigma_A \times \Sigma_1 \times ... \times \Sigma_c$ is an action profile, chosen by the players in the game.
- $out: \Sigma \to \Omega$ returns the result of the election, given an action profile.
- $v_i: \Omega \to \mathbb{R}$ is a function that gives the value of the election for the player C_i . We assume that $v_i(out(\sigma)) = v_i^*$ when the result of the election is in favor of the coercer i, and $v_i(out(\sigma)) = v_i^* \epsilon_i$ for some $\epsilon_i > 0$ when the result of the election is not in favor of the coercer.
- k_i is the number of voters that player C_i attempts to coerce.
- β_i is the cost of bribing one voter for the player C_i . It is assumed that this number is constant for any number of bribed voters and for all the privacy methods used.
- $d_i: \Sigma_A \to \mathbb{R}$ is the cost of breaking privacy. More precisely, $d_i(\alpha_j)$ shows the cost of verifying the vote of one voter by the coercer C_i , when the privacy method α_j is applied. It is assumed that $d(\alpha_0) = 0$, and $d(\alpha_t) \leq d(\alpha_{t'})$ if t < t'.

Utility of a mixed strategy profile is defined as the expected value of utility, assuming that different players randomize independently.

2.3 Solution Concepts: Nash vs. Stackelberg

In game theory, solution concepts are used to define which collective behaviors (represented by strategy profiles $\sigma=(s_A,s_1,\ldots,s_c)$) are "rational" and should (or may) be selected. Different solution concepts correspond to different notions of rationality – more precisely, to different models of the deliberation process that leads to selecting one or the other strategy. Our analysis of simple models of coercion in Section 3 proceeds by comparing the predictions obtained by two solution concepts: Nash equilibrium and Stackelberg equilibrium.

Nash equilibrium represents a play which can emerge when players adapt their choices to what they expect from the other players. Formally, σ is a Nash equilibrium iff no player can unilaterally change her strategy in σ so that she increases her utility (the strategies of the other players are assumed to stay the same). Nash equilibrium often captures the collective behavior which emerges "organically", through a sequence of strategy adjustments from different players that leads to a point when nobody is tempted to change their strategy anymore.

Stackelberg equilibrium represents a rational play in a game with a designated leader player. The leader has the power to choose her action first. She also publicly commits to her choice so that the other players are fully aware of it. Formally, Stackelberg equilibrium is defined as the best response to best response. That is, for every strategy s_i of the leader we find the response $resp(s_i)$

$$\begin{array}{|c|c|c|c|c|}\hline & b_0^1 & b_{n_1^*}^1 \\\hline \alpha_0 & (v_A^* \ , \ v_1^* - \epsilon_1) & (v_A^* - \epsilon_A - \beta_A n_1^* \ , \ v_1^* - \beta_1 n_1^*) \\ \alpha_1 & (\boldsymbol{v_A^*} - \boldsymbol{imp}(\boldsymbol{\alpha_1}) \ , \ \boldsymbol{v_1^*} - \epsilon_1) & (v_A^* - \boldsymbol{imp}(\boldsymbol{\alpha_1}) - \beta_A n_1^* \ , \ v_1^* - \epsilon_1 - n_1^* \cdot (d_1(\boldsymbol{\alpha_1}) + \beta_1)) \\ \end{array}$$

Figure 1: Game model for perfect privacy. The Stackelberg equilibrium for the authority is shown in bold. There is no Nash equilibrium in pure strategies

that maximizes the utilities of the opponents; then, we select the s_i which maximizes $u_i(s_i, resp(s_i))$. Additionally, we will use the term Stackelberg minimum to denote the strategy profile consisting of the leader's Stackelberg strategy and the most damaging response of the other players. Intuitively, Stackelberg minimum represents the worst that can happen if the leader plays her Stackelberg strategy minimum minimu

3 Analyzing Simple Models of Coercion

In this section we look at two simple game models that represent a simplified view of the social context of coercion. In both cases, we assume that there is only one coercer. That is, we look at the two-player game that arises between the society (represented by the election authority A) and the sole coercer C_1 .

We distinguish two possible settings: with perfect privacy (and hence perfect coercion resistance) available for A, and the other one, where A can only increase the cost of breaking the privacy but it cannot rule out the possibility completely.

3.1 Perfect Privacy

The first case that we study, is the situation where only one coercer is acting in the game, and the election authority has a choice between no privacy and perfect privacy. In this case the game model is $\langle \{A,C_1\},(\Sigma_A,\Sigma_1),(u_A,u_1)\rangle$, in which $\Sigma_A=\{\alpha_0,\alpha_1\}$, where α_0 represents no privacy and α_1 represents prefect privacy. As the actions b_1^1 to $b_{n_1^*-1}^1$ are all dominated by the action b_0^1 , we remove them from the game table. As a result the only actions considered for the coercer are b_0^1 and $b_{n_1^*}^1$. The payoff table for this game is depicted in Figure 1.

This game has no pure Nash equilibrium. In its mixed Nash equilibrium the authority chooses "no privacy" with the probability $p = \frac{n_1^*(d_1(\alpha_1) + \beta_1)}{\epsilon_1 + n_1^*d_1(\alpha_1)}$, and the coercer attempts to coerce the voters with the probability $q = \frac{imp(\alpha_1)}{\epsilon_A}$. One way for the authority to improve its utility is bringing the game to the Stackelberg equilibrium. It can do that by making its commitment to choose α_1 public. As a result, the coercer should give up coercion attempts as they only

Figure 2: Game model for breakable privacy. The Nash equilibrium is shown in red, and the Stackelberg equilibrium for the authority is shown in bold

increase the cost with no hope for successful coercion. By using the Stackelberg equilibrium, the utility of the authority is improved by $\frac{imp(\alpha_1)\beta_A n_1^*}{\epsilon_A}$. Thus, A has an incentive to select the coercion-resistance policy in advance, without adapting to the strategy of the coercer.

Moreover, it is easy to see that the Stackelberg equilibrium does not coincide with Stackelberg minimum (because $v_A^* - imp(\alpha_1) > v_A^* - imp(\alpha_1) - \beta_A n_1^*$). Thus, A is better off when publicly committing to strategy α_1 , rather than keeping the policy secret and risking random response of the coercer.

3.2 Imperfect Privacy

In this case only one coercer is acting in the game, and the election authority has several choices for the privacy method, none of them providing perfect protection against coercion. In other words, under all the privacy methods the coercer is able to coerce successfully, the only difference is how much he has to pay. The game model is $\langle \{A, C_1\}, (\Sigma_A, \Sigma_1), (u_A, u_1) \rangle$, in which $\Sigma_A =$ $\{\alpha_0,...,\alpha_{Max}\}$. As the authority changes the privacy method from α_0 to α_{Max} , the coercer's cost of coercing. For a given privacy method α , if $v_1^* - \epsilon_1$ is smaller than $v_1^* - n_1^* \cdot (d(\alpha) + \beta)$ then the coercer prefers coercing. But it might be the case that from some privacy method α_m on, the cost of coercing for the coercer is greater than ϵ_1 . In this case the coercer, although able to coerce successfully, prefers not to attempt coercion in the election. We assume that, among the available privacy methods, the coercer prefers to coerce when the privacy method is among α_0 to α_{m-1} , and prefers not to coerce when the privacy method is among α_m to α_{Max} . In other words for $\alpha_i, m \leq i \leq Max$, we have $(v_1^* - \epsilon_1) \ge (v_1^* - n_1^* \cdot (d(\alpha_i) + \beta))$. In consequence, it is sufficient to consider only $b_0^1, b_{n_*}^1$ as the available actions of the coercer. The resulting payoff table is depicted in Figure 2.

This game has a unique pure Nash equilibrium in $(\alpha_0, b_{n_1}^1)$. In this equilibrium, the coercer attempts to coerce enough voters, and the authority chooses the zero-level privacy policy. This can hardly be seen as a socially desirable outcome of the election. Like in the previous game, we observe that the authority's Stackelberg equilibrium is strictly more beneficial to the society than the Nash equilibrium. Thus, if the authority commits to the privacy policy α_{m^*} and makes this commitment public, such that the coercer believes the commitment of the authority to use α_m , the coercer chooses b_0^1 . By using the Stackelberg

strategy, the authority improves its utility by $\epsilon_A + \beta_A n_1^* - imp(\alpha_m)$. Note also that the Stackelberg equilibrium is again different from Stackelberg minimum. This suggests that even if the authority chooses a privacy method which is difficult enough to break, such that the coercers should not attempt to coerce, the authority can only benefit from that when it makes its decision public and makes the coercers know its choice and believe in its commitment.

4 Conclusions

In this paper, we look at simple game-theoretic models of coercion resistance in voting procedures. Elections are modeled as two-person non-zero-sum non-cooperative games, where one player represents the society and the other a potential coercer in the election. The work is still preliminary, and the models are arguably too simple. Still, even at this stage some interesting patterns can be observed. Most importantly, we show that in all games that we consider, Stackelberg equilibrium is different from Nash equilibrium. In other words, it is in the interest of the society not to adapt to the expected strategy of the coercer. Rather, the society should decide on its coercion-resistance policy in advance. Moreover, the Stackelberg equilibrium does not coincide with Stackelberg minimum, which suggests that the society will benefit from announcing its policy openly. This way, the coercer is best off when refraining from coercion altogether.

References

- [1] Roberto Arajo, Narjes Rajeb, Riadh Robbana, Jacques Traor, and Souheib Youssfi. Towards practical and secure coercion-resistant electronic elections. In Swee-Huay Heng, RebeccaN. Wright, and Bok-Min Goi, editors, Cryptology and Network Security, volume 6467 of Lecture Notes in Computer Science, pages 278–297. Springer Berlin Heidelberg, 2010.
- [2] Ahto Buldas and Triinu Mägi. Practical security analysis of e-voting systems. In *Proceedings of IWSEC*, volume 4752 of *Lecture Notes in Computer Science*, pages 320–335. Springer, 2007.
- [3] S. Delaune, S. Kremer, and M. Ryan. Coercion-resistance and receipt-freeness in electronic voting. In *Computer Security Foundations Workshop*, 2006. 19th IEEE, pages 12–pp. IEEE, 2006.
- [4] Ryan W. Gardner, Sujata Garera, and Aviel D. Rubin. Coercion resistant end-to-end voting. In Roger Dingledine and Philippe Golle, editors, Financial Cryptography and Data Security, volume 5628 of Lecture Notes in Computer Science, pages 344–361. Springer Berlin Heidelberg, 2009.

- [5] J. Heather and S. Schneider. A formal framework for modelling coercion resistance and receipt freeness. FM 2012: Formal Methods, pages 217–231, 2012.
- [6] A. Juels, D. Catalano, and M. Jakobsson. Coercion-resistant electronic elections. In *Proceedings of the 2005 ACM workshop on Privacy in the* electronic society, pages 61–70. ACM, 2005.
- [7] R. Küsters, T. Truderung, and A. Vogt. A game-based definition of coercion-resistance and its applications. In *Proceedings of the 2010 23rd IEEE Computer Security Foundations Symposium*, pages 122–136. IEEE Computer Society, 2010.
- [8] R. Küsters, T. Truderung, and A. Vogt. Verifiability, privacy, and coercion-resistance: New insights from a case study. In *Security and Privacy (SP)*, 2011 IEEE Symposium on, pages 538–553. IEEE, 2011.
- [9] J.F. Nash. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences U.S.A.*, 36:48–49, 1950.
- [10] M. Osborne and A. Rubinstein. A Course in Game Theory. MIT Press, 1994.
- [11] Peter Y. A. Ryan. The computer ate my vote. In *Formal Methods: State* of the Art and New Directions, pages 147–184. Springer, 2010.
- [12] M. Tambe. Security and Game Theory: Algorithms, Deployed Systems, Lessons Learned. Cambridge University Press, 2011.
- [13] Heinrich Freiherr von Stackelberg. Marktform und Gleichgewicht. Vienna, 1934.