On privacy and anonymity in electronic and non electronic voting: the ballot-as-signature attack.

Roberto Di Cosmo

On privacy and anonymity in electronic and non electronic voting: the ballot-as-signature attack.

Roberto Di Cosmo
PPS Laboratory (UMR 7126)
Université Paris 7 Denis Diderot
Case 7014
2, Place Jussieu
75005 Paris
http://www.dicosmo.org

Abstract

A wealth of protocols for electronic voting have been proposed in the literature over the past years. What makes these protocols difficult to conceive and verify is one fundamental property, anonymity, which is of paramount importance in the real world, in particular when performing actual political elections.

Historically, certain techniques have been used in actual elections to nullify anonymity and effectively coerce voters, by exploiting an evident weakness in many voting protocols; these techniques were used in traditional elections well before the notion of electronic voting was even proposed, yet, they still seem to be little known: as a consequence, we find recent proposals of voting protocols that can easily be attacked this way, like [Riv06], or clever attempts at formal definitions of privacy and anonymity properties that would not rule out such flawed protocols, like [DKR06].

In this paper, we describe one old technique, effectively used in Italy over twenty years ago, and show how the flaws or incompleteness in current protocols and formalization can be clearly exposed just through that simple idea.

We also show how this very same simple attack can be effectively used today on US-style elections, regardless of the presence of a VVPB or VVPT.

We hope that a wide circulation of this simple ideas will help design better protocols and formalization in the near future.

Keywords: privacy, anonymity, electronic voting, voting, signature attack
ACM Classification: K.4.1 Public Policy Issues (Privacy), D.4.6 Security and Protection

Date : 2007 – 04 – 1512 : 19 : 01

1 Introduction

There is an extensive literature on voting and elections, which is not surprising, since the need to perform voting dates far back in human history[1]. The study of voting is a particularly complex subject, that has raised interest among scientists in very different disciplines, ranging from philosophy, to politics, economics, and finally, more recently, computer science. The aspects of voting that are taken into account in each of these disciplines are quite different: some focus on designing a voting *method* that properly capture the interest of the population that votes; in computer science, one mainly focus on the *security* properties that a particular *protocol* used to implement a vote must satisfy.
It is this author's humble opinion that these different communities act as if these different aspects were orthogonal; this author strongly believes that relevant properties of the voting systems are not orthogonal: one cannot really build secure voting protocols without knowing the details of the voting method, and there is no point in designing a perfectly fair voting method if it makes it difficult to provide a secure implementation, as then its theoretical fairness will be practically impaired.
In this paper, we will first survey a few ideas and results coming from both the *security* and *social welfare* approaches, then tell the story of an old trick used in Italy over twenty years ago to break security of a voting method; the basic idea used in this old trick will allow us to construct an effective attack against the Three Ballot scheme recently proposed by Ronald Rivest [Riv06], to construct an attack on any simplistic implementation of the very appealing voting method proposed by Balinski and Laraki just a few months ago [BL06], to uncover a possible attack in all US-style aggregated elections, and finally to show the inadequacy of recent attempt at the formalization of the privacy property, like the one proposed, for example, in [DKR06].
As the reader will see, the fundamental idea is really a simple one, and this author discussed the Italian trick at length with many colleagues over the past ten years, but it is quite surprising to see that its basic principle are seemingly ignored by most concerned researchers today, hence the need for this exposition.
In the following, we will use terms like *voting method*, *voting protocol*, and *voting scheme*, which are often used interchangeably in the literature, with a more specific meaning:

voting method will indicate those aspect of a voting system that are mostly concerned with defining the rules to choose the

[1] In ancient Greece, for example, shards of pottery (*οστρακα*) were used to mark votes to decide whether to ban an individual from the community, hence the term *ostracism*.

winner(s): there are a wide range of such methods in use today, *proportional*, *single winner*, *single non transferable vote*, etc.

voting protocol will indicate those aspect of a voting system that are mostly concerned with the *implementation procedures* of a voting system: mechanisms for determining eligibility of voters, anonymity, verifiability, etc. fall here;

voting scheme will be used where the frontier among method and protocol is blurred (for example, some voting protocols really are designed to implement just one particular voting method)

1.1 Voting protocols studied from the *security* point of view

There are quite a few studies dedicated to voting *fraud* at different moments in history [Ste57], and where there is fraud, there is a need for security engineering.
This author found it difficult to identify systematic publications devoted to the analysis of the specific security problems posed by the actual design and implementation of voting procedures before the advent of computers, and it may be the case that the nature of these problems and their old origins make the relevant considerations kind of *folklore*: it is easy to find evidence in the electoral codes of many countries (for example, see [Ghe02] for a discussion of the security provisions devised in France) that there is a long standing consensus on the fact that any particular voting procedure used for political elections should at least guarantee the following properties:

anonymity the choice made by each voter is not disclosed to anyone

eligibility only legitimate voters can vote, according to the particular voting rules of the election, and all legitimate voters can vote

fairness to avoid influencing the remaining voters, no partial results can be disclosed[2]

public verifiability the voting process is open to public scrutiny so that anyone can become confident that her vote was actually counted and that the announced outcome corresponds to the votes actually made.

Due to the wide variety of voting practices in different countries, attacks on these basic properties are also quite diverse. Until 2003, the implementation of political elections in France analyzed in [Ghe02] gave good guarantees on anonymity, fairness and public verifiability, while it was seriously flawed concerning eligibility. Alleged eligibility fraud was also at the center of much debate in Florida during the 2000 US presidential elections.
In different occasions, anonymity has been reported to be attacked, even if the effectiveness and implementability of such attacks deeply depended on the particular details of the voting procedure [Jon05], or the level of coercion that can be exerted ostensibly without nullifying the election outcome (typically, in developing countries).
In computer science, voting protocols became an interesting object of study with the discovery of public key cryptography, when many researches started looking for a way to reproduce in a purely electronic voting method, possibly enabling remote voting, the same security properties that are reasonably assumed with physical voting.
Due to the unique combination of security requirements that arise in the design and implementation of voting protocols, and in particular with remote voting, this quest is far from being concluded, and along the way a certain amount of extra properties of an (electronic) voting system have been formulated, which are intuitively subsumed by the anonymity property above (each of these properties breaks anonymity), but make sense in an electronic setting, where the attackers may have significant more ways to disrupt anonymity:

receipt freeness : a voter cannot prove to a third party, *after the vote*, that she voted in a certain way [BT94];

coercion-resistance : a voter cannot prove to a third party that she voted in a certain way, *even if the third party is allowed to electronically interact with her during the voting process* [JCJ02].

Looking at the long list of papers published with protocol proposals that were proven flawed just a few month after their publication[3], one can get a clear idea of how difficult the problem is.
This quite unsatisfactory state of affairs prompted for some formal methods to enter the scene: a seminal work in this direction is that by Lowe [Low95] and recently, some notable work advocates the use of variants of the π-calculus to formalize the definitions of these properties, and make them suitable for machine verification [KR05].
What is quite striking in most studies concerning the *security* aspects of (electronic) vote is the fact that one does not care much about the actual *content* of a ballot, nor on the particular aggregation function used to obtain the final result of the election: these are considered *details* that can be abstracted away for the sake of the security analysis; in many cases, it is simply assumed that the ballot is just a yes/no choice, and the result is simply the sum of the ballots (yes counting as a one and no as zero) [Sch99].
As we will see in the next section, these *details* are actually extremely important from the political, economical, social and mathematical point of view, so one cannot assume that if some security properties of a voting protocol on one hand, and some fundamental properties of a voting method on the other hand turn out to conflict, it will be the latter that will have to be adapted.

1.2 Voting methods studied from the *ethical* or *social welfare* point of view

Philosophers, politicians and economists focus on building voting methods with good *ethical* properties, which can be formalized mathematically as a quest for a function that transform a set of individual preferences into a global ranking of the candidates having some good properties that model the common interest; in these disciplines, it is common to see the term *social welfare function* used to denote them.
A particularly significant period for the analysis and discussion of voting methods is the 18th century, when Nicolas de Condorcet, a french nobleman and distinguished mathematician, devoted a number of studies to the problem of finding an election method that guaranteed an outcome representative of the voters' will. He observed, in particular, a significant shortcoming of single-winner election methods, where each voter votes for just one choice, and

[2] This is not so consensual a property as it might seem, as some authors have proposed to let the voters change their choice continuously over a period of time, according to partial results, even if this author could find no evidence of the actual usage of such systems in political elections.

[3] See for example the history of receipt-freeness reported in the introduction of [BT94]

the winner is whomever gets more votes: it may well be the case that the winner so selected is actually the *least preferred* in the population. This is best seen in an example: suppose that we have a race among three candidates, Adam, Bill and Cynthia, and the real feeling of the voters is represented as follows

20	Adam	Bill	Cynthia
15	Cynthia	Bill	Adam
10	Bill	Cynthia	Adam

Since each voter can only cast one preference, Adam will win with 20 votes, followed by Cynthia with 15 and Bill with 10. The political commentators will then conclude that Adam is preferred over Cynthia, and Cynthia is preferred to Bill.

But if one looks at the real feelings of the voters, it is easy to see that:

- 30 people prefer Bill to Cynthia, while only 15 prefer Cynthia to Bill;
- 25 people prefer Bill to Adam, while only 20 prefer Adam to Bill;
- 25 prefer Cynthia to Adam, while only 20 prefer Adam to Cynthia.

So it would be more representative to have Bill ranked first, followed by Cynthia and Adam last, that is, the exact inverse of the result of the single-winner method.

Condorcet concluded that it was *necessary*, if one really wants to get a representative outcome from an election, to choose a voting method where the voters can express *all of their preferences, in ranked order*.

He also managed to show that this is far from sufficient: after proposing a particular voting method, the Condorcet method, he uncovered what is known as *Condorcet's paradox*: there are cases where even using his method, we cannot pick a clear winner; for example, with the following voting outcome it is the case that Bill is preferred to Cynthia, which is preferred to Adam, *which is preferred to Bill*:

20	Adam	Bill	Cynthia
15	Cynthia	Bill	Adam
10	Bill	Cynthia	Adam
6	Cynthia	Adam	Bill

Indeed, we have

- 30 people prefer Bill to Cynthia, while only 21 prefer Cynthia to Bill;
- 31 prefer Cynthia to Adam, while only 20 prefer Adam to Cynthia.
- 26 people prefer Adam to Bill, while only 25 prefer Bill to Adam.

We can even have situations where the apparent winner is really hiding a perfect tie, like in the following example where Ada, Bill and Cynthia really split the population's will (checking this is left as an exercise to the reader):

20	Adam	Bill	Cynthia
20	Bill	Adam	Cynthia
20	Cynthia	Adam	Bill
20	Cynthia	Bill	Adam

This rather unsatisfactory state of affairs led to the proposal of a wealth of different voting methods that try to cope with this paradox, including one from Borda, a contemporary of Condorcet's; we now know, thanks to a result of Kenneth Arrow [Arr51], Nobel Prize for economics in 1972, that no voting method exists, based on the assumption that the voters can only express their preferences with a ranked list of candidates, that satisfies all of the following reasonable assumptions one might expect of an election, as soon as there are at least three candidates:

Unanimity if every voter prefers A to B, than in the election result A is better ranked than B

Non dictatorship there is no voter whose choice can systematically alter the election result

Independence of Irrelevant Alternatives whether A is ranked above B in the election outcome depends only on the voters' ranking of A respective to B (so, adding or removing the ranking of an extra candidate C, all the rest being equal, must not change the order of A and B in the result).

This might explain why we find around the world today quite a large number of voting methods, chosen according to the different appreciation of the importance of each feature of a voting method (and there are much more features than just the three above): most of them have in common the possibility for the voter to express her preferences via a ranked list of candidates (the length of this list may be constrained, for example, to one, but the above exposition clearly state that such a limitation leads to the worst election system), and many of the differences are to be found in the algorithm used to compute the outcome of the election.
As we will see in the next section, some of these variants allow to easily and blatantly break privacy.
Very recently, Michel Balinski and Rida Laraki made a clear analysis of the roots of the limitations of voting systems based on preferences expressed solely by means of ranked lists of candidates, and show in [BL06] that one can avoid Arrow's limiting result with voting methods that allow each voter to express her view of a candidate in a more refined way, by *grading* each candidate, using a common language[4]. For a detailed exposition of their proposal, the *majority ranking*, we refer the interested reader to [BL06]; for the purpose of the present work, it will suffice to remember that this method *requires* that the voter grade *every* candidate.
As we will see in what follows, naïve implementations of this method allow to easily break privacy.
Finally, independently on the voting method chosen, it is the case, most notably in the US, that the authorities try to reduce costs, and increase voter participation, by federating together unrelated elections: one may find in the same ballot on major elections in the US questions ranging from the choice of the next President of the United States, next to questions about homosexual marriage, or the local sheriff, which definitely do not carry the same weight.
As we will see in what follows, this practice also allows to easily break privacy, no matter what the voting method for each individual question is.

[4]Grading is more powerful than ranking; we all know how to rank students, given their grades, but it is quite another task to extract a grade from a ranking: knowing that A is considered better than B tells us nothing on the *absolute value* of A on a ranking scale.

There is a lesson to be learned from this exposition: election systems based on simplistic yes/no choices, even if actually used in practice, like in the French two round system, or worse, in the UK single run system, definitely represent the worst, least satisfactory election system of all, from the social welfare point of view, so it is definitely not acceptable to limit research in computer science to electronic voting protocols that are thoroughly studied only in the case of yes/no choices.

1.3 Additional implementation details

When proceeding to implement a voting system used for political elections, one is faced with quite a lot of additional practical design choices which may seem marginal, but are of paramount importance, and are well known to political parties.
The first of these practicalities is the necessity to choose a method to assign political representatives to geographical areas; for this, the global territory where elections are performed is divided into *electoral districts*[5], each district being assigned one or more seats.
Second, each district is subdivided into many *electoral precincts*[6], which is the area corresponding to one single polling station; its *size* is the number of voters that are eligible to vote in that area.
The typical size of an electoral precinct is around 1000, while electoral districts have sizes that vary widely depending on the country (in France, there are around 100.000 voters in a district), and the importance of the way the territory is divided into districts is well known to all politicians.

2 An old attack on the ranked voting method in Italy

The Italian law fixing the voting method for the election of the chamber is the *Testo unico delle leggi elettorali*, which was first established in the *Decreto del Presidente de la Repubblica del 30 marzo 1957, n. 361*, published in the *Gazzetta Ufficiale* n. 139, the 3rd of June 1957. This law has undergone a wealth of modifications over the years, but at the time of reference that this author recall for the fraud in object (in the 70s and 80s), it was consistently the case that representatives of the parliament were elected with a *proportional* method, that can be simplified as follows:

- there are a few large electoral districts, and each of them is assigned a number of seats that will be given to the candidates with the best scores
- in each district, every party may present a ranked list of candidates
- by means of a special ballot, the voter chooses a party, and can either agree on the ranking of the candidates proposed by the party, or optionally express different *preferences* with respect to the party's proposal, by marking a (limited) list of candidates, out of the chosen party's list, in any order she likes
- a rather complex algorithm is then used to determine the winners in the district; the details of such algorithms are out of the scope of this paper, but it suffices to know that the party with more votes will have more seats, and for each party, the candidate with more preferences (either explicitly marked by the voters, or implicitly indicated by the party's ranked list) will have more chances to get a seat.

[5] French: circonscription éléctorale, Italian: circoscrizione elettorale, possibly divided into collegi elettorali

[6] French: bureau de vote; Italian: seggio elettorale

To give an idea of what a paper ballot looked like in the Italian elections, I present in figure 1 an actual ballot still used for the European elections today. The voter chooses a party by drawing a cross on the party's symbol, and may optionally express a number of ranked preferences by writing down candidate names on the lines on the right of the party's symbol. In the current European election method the number of these preferences is limited to three; in the Italian parliamentary elections in the 80's the voter could express up to 4 preferences and lists with over 40 candidates were commonplace.

Figure 1. Facsimile of a paper ballot for european elections used in Italy

2.1 Voiding anonymity by converting the ballot into a signature

This author was unfortunately unable to find an online reference to the fraud that was unveiled sometime in the 80's in Italy[7], and led to the modification of the election procedures that reduced the number of preferences to only one[8], even if they still can express three in the european election, but the following story should not be too far from the actual events (if you have an Italian friend over 40 years old, he will remember it too).
In various towns, it had been observed a high number of votes for the local political *boss* during the elections conducted using paper ballots similar to the one in figure 1, with the possibility of expressing up to 4 preferences. It was also observed a high usage rate of preferences in the ballots.
Taken together, these hints may be interpreted as pointing to some kind of fraud being performed taking advantage of the preferences system, but it is not evident: after all, a high usage rate of preferences may simply indicate that preferences are useful for the voters, and hence used! Indeed, it took a long time before the actual fraud procedure was largely understood, leading to a change of the system: the father of this author served as a volunteer officer in an electoral precinct for over 20 years, and actually recalls remarking in several occasion people taking notes of the preferences expressed in the votes during the public counting phase, but only realized what was going on when the whole affair become public in the '80s.
The actual fraud procedure can be reconstructed as follows:

- the *boss* is candidate C_1 in the list of candidates $C_1, \ldots, C_n$ of the fraud party P_F;

[7] The author would be very grateful to whomever may have a copy of the old newspaper articles, and would be so kind to scan and send him a copy.

[8] Actually, since the last reform, in 2006, preferences are gone altogether.

- before the elections, the boss, accompanied by Carlo and Giovanni, two well built, massive bodyguards with black suits and sunglasses, makes a courtesy visit to a sizeable number of electors susceptible of not voting for party P_F;
- in each visit, he instructs elector E_k to vote for P_F and express preferences $C_1, C_{\sigma_k(2)}, C_{\sigma_k(3)}, C_{\sigma_k(4)}$, *in this precise order*, where each $\sigma_k : 2\cdots 4 \to 2\cdots n$ is a different injection of the interval $2\cdots 4$ into the interval $2\cdots n$;
- he then tells E_k that, in case a ballot with this precise sequence of preferences does not come up during the tallying operations (that are, of course, public and publicly verifiable), Carlo and Giovanni will pay E_k a visit to make sure she will exhibit better memory efficiency when the next election will come;
- now, either E_k has some basic mathematical skills and is able to assess precisely her chances to disobey the boss' instructions, and complies, or she has not; in this second case, she may try her chance at disobeying the first time, but will comply after Carlo and Giovanni's visit.

To subsume this story, the possibility for the voter to express complex preferences allows to construct a *signature* out of the very content of the ballot, so that a fraud is easily set up by assigning each voter a different signature to check that she obeys the instructions, and all this *without* changing the overall outcome of the election (the boss will be elected no matter the order of the preferences of inferior weight).

2.2 Assessing the attack's effectiveness

In case E_k has some mathematical skills, and wants to take a chance to disobey, she will first assess the seriousness of the boss' threat, and the risk she incurs of being discovered by the boss.

First of all, she will try and check if it is really the case that the boss has found a space large enough to provide a *different* signature for every voter. For this, she has readily available a few informations:

- how many people vote in her same precinct: this varies from country to country, but in Italy the average number is $N = 1.000$, and is usually less than that;
- the list of candidates from which to choose the $C_{\sigma_k}(i)$ is on the average of length 40

The boss can produce as many signatures as there are different ways of choosing the 3 meaningless candidates out of the 39 which are left after putting the boss as the first of the 4 allowed preferences. This gives

$$3!\left(\begin{array}{c} 40 \\ 3 \end{array}\right) = 59280$$

different signatures, which is largely enough to cover 1.000 voters (a list of 13 names would be already enough), so the boss has a point.
Our voter now turns to assessing the actual risk of being caught in case she decides to disobey nonetheless: the only way she will not be discovered is when somebody else will actually cast a vote for P_F with exactly her assigned preferences $C_1, C_{\sigma_k(2)}, C_{\sigma_k(3)}, C_{\sigma_k(4)}$; how likely it is that nobody of the N voters actually cast this vote? Supposing (which is far from being reasonable) that votes are uniformly distributed, the probability that all of the N voters cast a vote different from $C_1, C_{\sigma_k(2)}, C_{\sigma_k(3)}, C_{\sigma_k(4)}$ is

$$\left(\frac{V-1}{V}\right)^N$$

where V is the number of possible different votes. The actual value of V depends on the number P of parties, and the number n_p of candidates in each party's list; under the assumption that the voters cannot express more than 4 preferences, this amounts to

$$\begin{aligned} V &= \sum_{p=1}^{P} (A^0_{n_p} + A^1_{n_p} + A^2_{n_p} + A^3_{n_p} + A^4_{n_p}) \\ &= \sum_{p=1}^{P} (1 + \frac{n_p!}{(n_p-1)!} + \frac{n_p!}{(n_p-2)!} + \frac{n_p!}{(n_p-3)!} + \frac{n_p!}{(n_p-4)!}) \end{aligned}$$

As our voter knows exactly P and all the values n_p, she can compute this probability, either precisely, if she has access to a good computer with an arbitrary precision arithmetic library, or via the approximation (valid for N/V small, which was the case in Italy for $N = 1000$ and $n_p = 40$).

$$\left(\frac{V-1}{V}\right)^N = \left(1 - \frac{1}{V}\right)^N \approx 1 - N/V$$

In table 1 we give some values of this probability, for a race with just ten parties (at the time, in Italy, ten parties were a minimum, and if you look at the specimen of paper ballot for last european elections, things are not getting any better), but also for a race with just one party.
As you can see, the probability of getting caught decreases as the number of voters increases and increases as the length of the list or the number of parties grows, which corresponds to the intuition that the chances of finding a saver decrease when there are more choices for the votes, and less voters around.
Even under the unrealistically optimistic assumptions that all votes are equi-probable, with lists of length 40 and a precinct size of 1.000 our potential hero has vanishingly small chances of avoiding Carlo and Giovanni's visit if she does not comply with the boss' order.
It is clear that if a sizeable amount of the voters decide to comply, our hero's chances are quite smaller: all voters that comply cast a vote *different* from $C_1, C_{\sigma_k(2)}, C_{\sigma_k(3)}, C_{\sigma_k(4)}$, so that we can effectively subtract their number from N, and look up again table 1 to assess the damage.

The little calculations of this section show that the boss has managed to void anonymity by turning the very ballot into an extremely effective voter signature[9].
To summarize the findings of this section

REMARK 1. *When the space of possible votes is large enough to construct a unique signature for each voter in the voting precinct, and the probability of a collision of the publicly verifiable ballots in the precinct is low, then anonymity can be voided, no matter how the voting protocol is designed.*

So, contrarily to what is implicitly assumed in a vast literature on security of voting protocols, *details* of the voting "method" (as opposed to the voting "protocol"), do matter when assessing the security of the system.

[9] A colleague of mine recently told me that he knew one of the engineers that were asked by some political parties to perform these calculations in the south of Italy back in the '80s...

Candidates	*Precinct size*	P_C 10 *parties*	P_C 1 *party*
40	100	0.999996	0.999956
40	1000	0.999956	0.999556
40	2000	0.999911	0.999113
40	3000	0.999867	0.998670
40	4000	0.999823	0.998227
40	5000	0.999778	0.997784
30	100	0.999985	0.999854
30	1000	0.999854	0.998537
30	2000	0.999707	0.997076
30	3000	0.999561	0.995617
30	4000	0.999415	0.994160
30	5000	0.999268	0.992706
20	100	0.999919	0.999191
20	1000	0.999191	0.991937
20	2000	0.998382	0.983939
20	3000	0.997574	0.976005
20	4000	0.996767	0.968135
20	5000	0.995960	0.960329
10	100	0.998295	0.983081
10	1000	0.983083	0.843130
10	2000	0.966452	0.710869
10	3000	0.950102	0.599355
10	4000	0.934029	0.505334
10	5000	0.918227	0.426062
5	100	0.952604	0.614701
5	1000	0.615356	0.007703
5	2000	0.378663	0.000059
5	3000	0.233012	0.000000
5	4000	0.143385	0.000000
5	5000	0.088233	0.000000
4	100	0.857302	0.212159
4	1000	0.214457	0.000000
4	2000	0.045992	0.000000
4	3000	0.009863	0.000000
4	4000	0.002115	0.000000
4	5000	0.000454	0.000000

Table 1. Exact probability of getting caught when disobeying the boss, assuming a limit of 4 expressed preferences in the race, as a function of the number of candidates, the precinct size and the number of parties in the race.

Add a footnote somewhere to explain the trick for voiding votes in France...

3 Signature attack on Rivest's Three Ballot voting protocol

In a recent paper, Ronald Rivest proposed a paper ballot voting protocol inspired by ideas coming from cryptographic voting protocols designed for electronic voting [Riv06]. The essential idea is to try and provide a voting protocol that will allow the voter to come back from the voting booth bringing with him a piece of information that is not enough to discover the vote cast, but allows the voter, with high probability, to detect any malevolent manipulation of the tally that would have as effect of ignoring her particular vote.

We will not reproduce here the discussion of the protocol, for which we refer the reader to the original paper, but we only resume the voting steps:

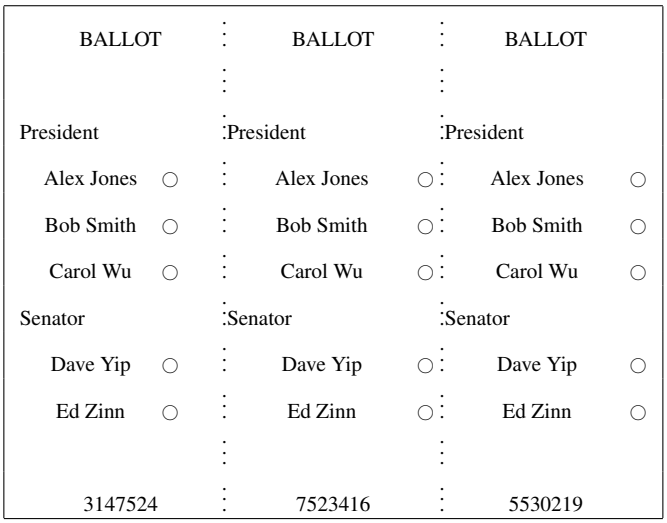

Figure 2. A sample multi-ballot from the Three Ballot voting protocol

1. the voter gets an empty multi-ballot,
2. to vote FOR a candidate, she fills with her black pencil *exactly two* of the ○ next to the name of that candidate
3. to vote AGAINST (or not vote for) a candidate, she fills with her black pencil *exactly one* of the ○ next to the name of that candidate
4. it is not allowed to leave empty all the ○ corresponding to a candidate, or to fill all three of them; after being filled in, the multi-ballot would look like the one in figure 3
5. a special machine (which may look like the ones used to read the Lottery ballots, checks that the filling constraints are fulfilled, marks the bottom of the multi-ballot, then it cuts the multi-ballot in three, spitting out the resulting three ballots, and also produces a copy of only ONE of the three column, *at the voter's choice*
6. the voter keeps the copy, and puts the three ballots separately in the ballot box (which is of course is supposed to be regularly shuffled to mix the ballots in such a way that nobody can link together the three ballots to the original voter's multi-ballot)
7. once the vote is over, a summary of ALL the ballots is posted publicly, together with the list of all the voters, and every vote can check that her receipt is actually present in the public summary that is used to computer the election outcome

In the original paper, Ronald Rivest arguments that this protocols allows the voters to spot frauds (modifications of the ballots) with high probability by simply checking that her (secretly chosen) receipt does appear in the summary, while not giving up anonymity because the receipt alone does not give any indication on the actual value of the vote cast (for each possible configuration on the receipt can be completed with two other columns in order to yield whatever result one likes).

But what about the old 1980's Italian trick? Let's see: the boss will come to you and give you a precise configuration you must use to vote for him; this means *a full configuration of the multi-ballot*, not

BALLOT		BALLOT		BALLOT	
President		President		President	
Alex Jones	○	Alex Jones	○	Alex Jones	●
Bob Smith	●	Bob Smith	●	Bob Smith	○
Carol Wu	○	Carol Wu	●	Carol Wu	○
Senator		Senator		Senator	
Dave Yip	●	Dave Yip	○	Dave Yip	○
Ed Zinn	○	Ed Zinn	●	Ed Zinn	●
3147524		7523416		5530219	

Figure 3. A sample filled multi-ballot from the Three Ballot voting protocol: the votes go to Bob Smith for President, and Ed Zinn for Senator.

just the receipt you may choose to take home.

3.1 An analysis of the ThreeBallot

After having studied in detail the *Italan trick*, we know what is left to be done: we need (i) to check whether the boss has a large enough "ballot space" available to tag every voter in the precinct, and (ii) compute the probability for a voter of getting caught if she decides to disobey, or, equivalently, what is the probability that one of the tags chosen by the boss will come up nonetheless, even if our voter decides to disobey.

3.1.1 Tagging space in ThreeBallot

Counting how many *different* valid ballots one can cast is not difficult: for a race with k rows and $c \leq k$ maximum choices, we have the following possible outcomes

- ballots refusing all candidates: these have just one ● in every row, and there are 3 independent possibility for placing it in a row, so we have 3^k such ballots
- ballots choosing only one candidate: these are like the above ones, *but* for the fact that, on one and only one row, we find two ●, that can be arranged in one of three different ways only; so we can have produce 3^k different single-winner multi-ballots for every chosen row r, and there are k possibilities for choosing the row r; hence, there are $k \cdot 3^k$ single-winner multi-ballots
- in general, for a multi-ballot designating k' winners, we have $\left(\begin{array}{c} k \\ k' \end{array} \right)$ ways of choosing the k' winners, so we get

$$\left(\begin{array}{c} k \\ k' \end{array} \right) \cdot 3^k$$

multi-ballots designating k' winners (notice that the formulas for the two special cases above are specializations of the general formula for $k' = 0, 1$).

- the number of possible valid multi-ballots one voter can choose from is hence given, for c maximum choices, by the following formula

$$mb_{k,c} = 3^k \cdot \sum_{k'=0}^{c} \left(\begin{array}{c} k \\ k' \end{array} \right) \quad (1)$$

which gives, by the Binomial theorem, $mb_{k,k} = 3^k \cdot 2^k = 6^k$, and corresponds to the fact that when $c = k$ there are exactly 6 possible configurations for every line, that gives again 6^k possible multi-ballots.

Observing actual US ballots like the one in figure 4, we see that typical values for k are bigger than 40, and typical values for c can be bigger than 10. Our boss can then happily tag millions of different voters.

3.1.2 Probability of getting caught

For this, we need to spend a little time analyzing what the structure of Rivest's ThreeBallot implies on the probability distribution of the ballots one may find in the ballot box at the end of the day.
We make in the following the simplifying assumption that the probability distribution of the multi-ballots is uniform[10].
We first remark that each ballot that will be found in the ballot box at the end of the election necessarily comes form a valid multi-ballot (as the protocol checks that only valid multi-ballots may be cast). This has precise consequences on the probability distribution of the individual ballots in the ballot box:

- in an election where every multi-ballot cast is one refusing all candidates, we know that all the k lines on each multi-ballot have two ○ and one ●, so the probability of seeing a *column* with j occurrences of ○ and $k-j$ occurrences of ● is

$$\left(\frac{2}{3}\right)^j \cdot \left(\frac{1}{3}\right)^{k-j}$$

- in an election where every multi-ballot cast is one accepting all candidates, we know that all the k lines on each multi-ballot have two ● and one ○, so the probability of seeing a *column* with j occurrences of ○ and $k-j$ occurrences of ● is

$$\left(\frac{1}{3}\right)^j \cdot \left(\frac{2}{3}\right)^{k-j}$$

- in general, in an election with *up to c possible winners*, supposing a uniform distribution of the multi-ballots, the probability of seeing a *column* with j occurrences of ○ and $k-j$ occurrences of ● is given by

$$(1 - p_k^c)^j \cdot (p_k^c)^{k-j}$$

where p_k^c is the ratio among the total number of occurrences of ● over the number of total positions in all multi-ballots with k rows and up to c winners; the value of p_k^c is evidently greater or equal than $1/3$ (obtained with 0 winners) and strictly smaller than $2/3$ (never attained, as it is a limit for k winners), but since we have already undertaken some combinatorial show in this paper, here is the exact calculation that formalizes and narrows these bounds:

[10] When this is not the case, the probability of getting spotted may be *higher*.

– there are, by (1), $3^k \cdot \sum_{k'=0}^{c} \binom{k}{k'}$ possible such multi-ballots, each with $3k$ positions, hence the number of positions in all these multi-ballots is $3k \cdot 3^k \cdot \sum_{k'=0}^{c} \binom{k}{k'}$

– the number of ● occurring in total is then given by

$$3^k \cdot \sum_{k'=0}^{c} \binom{k}{k'} (k+k')$$

as we have in each valid ballot with k' winners k occurrences of ● (no line can be empty), plus k' more occurrences of ● (for the rows designating a winner)

– so p_k^c is

$$\frac{3^k \cdot \sum_{k'=0}^{c} \binom{k}{k'} (k+k')}{3k \cdot 3^k \cdot \sum_{k'=0}^{c} \binom{k}{k'}} = \frac{1}{3} \cdot \left(1 + \frac{\sum_{k'=0}^{c-1} \binom{k-1}{k'}}{\sum_{k'=0}^{c} \binom{k}{k'}}\right)$$

This value is greater or equal to $1/3$ (attained for $c = 0$), and strictly smaller or equal than $1/2$ (attained for $c = k$, where by the binomial theorem $\sum_{k'=0}^{c-1} \binom{k-1}{k'} / \sum_{k'=0}^{c} \binom{k}{k'} = 2^{k-1}/2^k = 1/2$).

So, $1/3 \leq p_k^c \leq 1/2$, and the probability of seeing any one given column is between $\left(\frac{1}{3}\right)^k$ and $\left(\frac{1}{2}\right)^k$.

Now, if our voter decides to resist the boss, she must change at least one of the two columns for which she does not obtain the receipt (the third one must be conformant to the boss' request, as she is supposed to turn it in to him).

So the probability P_C of getting caught is equal to the probability that nobody of the other N voters will by chance produce the column she decided to modify. We already know how to estimate this, and we can properly conclude this section by stating the following

PROPOSITION 1. *Using Rivest's ThreeBallot voting scheme, it is possible to sell a vote (or to coerce a voter), with probability P_C of identifying a breach in the contract in the following range*

$$\left(1 - \left(\frac{1}{3}\right)^k\right)^N \geq P_C \geq \left(1 - \left(\frac{1}{2}\right)^k\right)^N$$

where N is the precinct size and k the number of different rows in the ballot.

With typical values of k (greater than 50) and N (less than 1.000) from US elections, even neglecting the actual values of c, our hero will be caught with probability greater than 0.999999999999. And this, of course, assuming every other voter is choosing freely: voters obeying the boss will surely produce ballots *different* from the one our hero got, thus effectively reducing the effective value of N.

Notice that this is not the only problem with this voting scheme: Charlie Strauss [Str06b, Str06a] and Andrew Appel [App06] present various other approaches to break it.

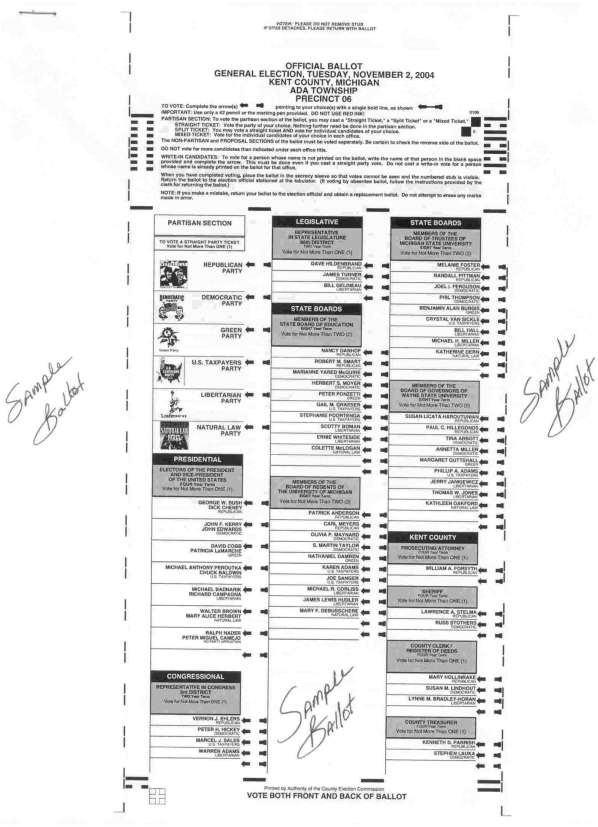

Figure 4. Example sample ballot from an actual US election

4 Ballots as signatures in the main US elections

We are now well familiar with the ballot-as-signature attack based on the *Italian trick*: we can perform it as soon as the ballot space, the precinct size and the tallying method allow to construct a high probability signature directly out of the very same ballot data that is necessary for a correct tallying of the election, yet still allow a coercer, or a vote seller, to cast a vote that is useful to the fraud party.

If we look around a bit, we easily find candidates for an attack even where the voting method has nothing to do with the old italian one, nor with the cryptographically inspired Three Ballot method proposed by Rivest. Indeed, the regular US voting method is a very good candidate for another instance of our attack, and it is, as we will see, quite an interesting candidate, as it uncovers yet another conceptual error in the design of concrete voting protocols.

We have already looked at a typical US voting ballot, shown in figure 4, but we did it only in order to get concrete values for k and c in our formula computing the probability of a signature for the Three Ballot voting method in section 3, which involved determining the probability of seeing a particular column obtained from a legal multi-ballot.

Now, let's look again at this ballot and try to find a way to compute a signature *directly*; our first remark is, a typical US ballot is *mixing* elections of very different nature and weight: in the example ballot of figure 4 we find 6+1 choices for president, 4+1 choices for representative in congress, 3+1 choices for representative in state legislature, 10+2 choices for the state board, and so on down the weight scale until we get to 9+2 choices for the board of the state university,

3+1 choices for county clerk and 2+1 choices for county treasurer, not to mention the different proposals that are to be found on the back of the ballot (which is not reproduced here). The "+1" or "+2" part in the paragraph above stands for the blank places where one can mark a write-in candidate.
For any attentive reader, there is no point in insisting further: this voting method is a real nightmare when it comes to anonymity of the voter... given a large enough supply of names for the write-in places, it is extremely easy to reliably tag a few hundred voters, without even bothering to do any sophisticated calculation.
But even if we had no places for write-in candidates, it would be overly easy to use the low-weight choices to tag every voter and force her to vote for *any chosen configuration* in the high-weight choices.
If we just take the example of the ballot in figure 4, we have $\binom{9}{2}+\binom{9}{1}+1$ possible configurations for the board of trustees, and $\binom{9}{2}+\binom{9}{1}+1$ possible configurations for the board of governors of the state university, then 2 configurations for attorney, 3 for sheriff, 4 for county clerk and 3 for treasurer. This gives

$$T = 46*46*2*3*4*3 = 152352$$

possible configurations to choose from to tag the voters, and force all of them to vote for the same precise configuration in the high-weight elections (columns one and two of the ballot). It is easy to see that there are

$$V = T*8*5*4*46*56 = 516374528$$

possible different configurations for a ballot (excluding write-in candidates), so, for each voter, the probability of getting caught in a precinct of size $N = 1000$ when disobeying the boss' orders would be,

$$\left(1-\frac{1}{V}\right)^N = \left(1-\frac{1}{516374528}\right)^{1000} \approx 0.999998063423$$

And again, this is an optimistic approximation: the more tagged voters comply, the more infinitesimal becomes the chance of not getting caught, not to mention the fact that most untagged voters will probably simply get tired of filling lines before actually getting to the low-weight marks.
We can resume this analysis in the following

PROPOSITION 2 (PRIVACY PROPERTIES DO NOT COMPOSE).
Voting protocols that aggregate different elections of different importance into a single ballot may seriously compromise the privacy of the voters. Even if each election taken alone has good privacy properties, the aggregated election using a single ballot breaks privacy with a probability that increases with the number of aggregated elections.

4.1 On the boss' strategies to overturn an election

We have concentrated up to now only on the issue of determining whether the boss has enough ballot space to tag all voters in a precinct, and whether he can spot any potential voter resisting his threats with high accuracy.
Once these conditions are satisfied, it is clear that the boss has a means to violate privacy and anonymity, and can then mandate the value of the votes of the voters he decides to threaten.

But then there are quite a wealth of different strategies that the boss can choose from. If we focus on the US election example above, for example, the following strategies offer increasing degree of control at the price of a bit of creativity:

1. he can threaten (or instruct) the voters in a precinct, using tags that will give winning votes *only* to the designated candidates in the high choices: this way he gets the wanted result in the high choices, at the price of giving up control on the low choices;
2. or he can take advantage of the large numbers of ballot configurations he can use as tags, and the complex composition of the US ballot to actually determine the *full* outcome of the vote by choosing these tags carefully to make sure in every sub-race they give more points to the boss' designated candidates; in the example above, one possibility among many would be the following one:
   - choose 930 uniformly distributed configurations out of the 2116 possible ones for the first 2 sub-races (the ones with 46 choices each); these configuration will give approximately 50 points to each of the 18 candidates in these sub-races;
   - construct the first 930 tags using just the first 2 sub-races as the varying part of the tag and fixing the values of the 4 smallest sub-races to the wanted ones;
   - construct the remaining 70 tags by fixing the values for the first two races and letting the others vary.

   The net result is that in the first two sub-races the fraud candidates will have 70 points more than the competitors, and in the remaining sub-races the fraud candidates will have an overwhelming majority.

5 Signature attack for the Majority voting method

A few months ago, Michel Balinski and Rida Laraki proposed a way to overcome the limitations of voting systems based on preferences expressed solely by means of ranked lists of candidates, and Arrow's limiting result [BL06]. This is achieved using a voting method that allow each voter to express her view of a candidate in a more refined way, by *grading* each candidate, using a common language.

For a detailed exposition of their proposal, the *majority ranking*, we refer the interested reader to [BL06]; for instantiating our attack, it will suffice to remember that this method *requires* that the voter grade *every* candidate using a grade g chosen from a common set $g_1, \dots g_n$.

A possible ballot for voting using this method could be similar to the one shown in figure 5: there must be exactly one ● on every line, indicating the grade that the voter gives to that candidate[11].

With n grades and k candidates there are n^k possible configurations,

[11] This method is undergoing a real-world testing for the 2007 French presidential elections while this author is writing this article, using a ballot exactly like the one presented here; more information can be found in `http://www.ceco.polytechnique.fr/jugement-majoritaire.html`

MAJORITY BALLOT					
Candidate	Excellent	Good	Average	Bad	Very Bad
President					
Alex Jones	○	○	●	○	○
Bob Smith	●	○	○	○	○
Carol Wu	●	○	○	○	○
Senator					
Dave Yip	○	●	○	○	○
Ed Zinn	○	○	○	○	●

Figure 5. Sample marked ballot for voting with the majority method

of which $(n-1)^{k-w}$ assign the grade *Excellent* only to w chosen candidates; the boss than can tag $(n-1)^{k-w}$ different voters, and here again we have our limiting result for privacy:

PROPOSITION 3. *In any implementation of the Balinski-Laraki method using a single ballot, a voter can sell (or be coerced to cast) a vote with probability P_C of detecting a breach in the contract, assuming all configuration uniformly distributed, given by*

$$P_C = \left(1 - \frac{1}{n^k}\right)^N$$

This formula takes the following values in some realistic cases:

n	N	*french presidential*, $k=12, w=1$		*US election*, $k=50, w=12$	
		P_C	*number of tags*	P_C	*number of tags*
5	1000	0.999995904008	4194304	$< 10^{-17}$	$7.55e+22$
5	300	0.999998771201	4194304	$< 10^{-17}$	$7.55e+22$
3	1000	0.998120091053	2048	$< 10^{-17}$	274877906944
3	300	0.999435655844	2048	$< 10^{-17}$	274877906944

Hence, the privacy offered by this voting method, as implemented using the ballot presented here, is negligible under real world assumptions.

6 Countermeasures

Can we design *practical* countermeasures to the ballot-as-signature attack? The strength of the ballot-as-signature attack is also its weakness: the signature *is* the voter's ballot, when the voter's choices are complex. We will briefly outline in this section some possible countermeasures one could take against the boss' voter tagging endeavor.

We cannot change the voting *method* (for example, by reducing the voter choices), as this is a *social welfare* matter.

But we can try to decorrelate the voter's original ballot (that carries the signature chosen by the boss) from the *observable* part of the ballot (the information actually made public when tallying, and is the only information the boss get about the voter's actual choice).

In the voting schemes presented here, at the price of some serious impracticalities, this may be attempted, with varying degrees of success:

Ranked voting, Italian style : the voter has to cast 4 preferences; instead of using a single ballot holding up to four preferences, to be dropped in a single ballot box, one could use 5 separate ballots, of different colors, to be dropped in 5 different ballot boxes: one for the no preferences option, one for the first preference, one for the second, one for the third and one for the fourth; the voter will then cast one, two, three or four ballots into the corresponding ballot boxes. At the price of some mathematical gymnastics, one may still recover the correct tally (with preference transfer to the list as appropriate), provided one carefully checks that the voter casts her ballots according to the rules (she cannot, for example, cast a 4th preference without casting a 1st, 2nd and 3rd one!). Now, the observable part of the ballot is just the set of 1st, 2nd, 3rd and 4th preferences, and the boss is unable to check the presence of the expected signatures as easily as before: he may still spot a missing signature, but with much lower probability;

Three Ballot : we can modify the original protocol adding a machine that, after verifying that the multi-ballot is valid, will cut it not only in columns, but along the lines of different sections, or better, along all different lines; this way one effectively breaks the boss' signing ability; to retain the added benefits of the scheme, each snippet must carry a different serial number, and the copy of one ballot we bring home as a receipt, together with the serial numbers of the corresponding snippets, will still allow us to check that our retained ballot really is part of the tally; unfortunately, we now *really need* the serial numbers to perform our checks, and a simple counter-strategy for the boss is to team up with the maker of the machine that prints the serial numbers to be able to effectively reconstruct all the ballots starting from the serial numbers (this is much easier than one would think: it is just a matter of agreeing beforehand on some serial numbering scheme, and this is a practical attack that would work very well on the unmodified ThreeBallot system too);

US elections : here, the countermeasure outlined above (cut along the lines the different ballot sections) just boils down to the evident one: *do not aggregate* different elections; instead of a huge ballot with dozens of choices, one should get dozens of small ballots with few choices each;

Balinski-Laraki ranked voting : the same decorrelation technique will mandate to use a different ballot for each candidate, containing only the grade given to that specific candidate; if this author correctly understood the aggregation function sketched in [BL06], there is no need to know who gave which grade to a candidate in order to compute it, so the tallying will not be particularly more cumbersome than using a single ballot, and one needs only use one ballot box.

Interestingly, the Balinski-Laraki method, that is presented as the fairest one known, is also the easiest to practically and satisfactorily implement in a way that avoids ballot-as-signature attacks. It would be a good thing to see it deployed widely soon, provided, of course, it is implemented as suggested above.

7 Formalizations of privacy properties and the ballot-as-signature attack

The privacy and anonymity properties, with all their receipt-freeness and coercion-freeness electronic variants, are notoriously difficult to obtain, and this is why we find an increasing interest in formal methods to actually prove or disprove them, possibly (semi)automatically, for a given voting scheme.

One interesting work along this line is [DKR06], that proposes the following approach to formalization and verification.

First, a voting protocol is modeled as a process of the applied π-calculus [AF01] having the shape

$$VP = \nu \overline{n}.(V_{\sigma_1} | \ldots | V_{\sigma_n} | A_1 | \ldots | A_m)$$

Here V_{σ_i} are the voter processes, all obtained as instances, via the substitutions σ_i, of the same process V that contains free a variable

v, the intended vote of the voters; A_j represent the different election authorities, and $\overline{n}$ are channel names. The voting process VP will at some moment make public the result of the vote (by sending it on a designated public channel for example). One notes $S[]$ a context obtained from VP by suppressing any two voter processes.
Then, a protocol is defined to have the *privacy* property when the following holds

$$S[V_A\{a/v\}|V_B\{b/v\}] \approx S[V_A\{b/v\}|V_B\{a/v\}]$$

In simpler words, a protocol is considered to respect privacy when nobody can tell whether two voters exchanged votes.
In [DKR06] this definition is retained because it is considered to have several merits: it is amenable to (sometimes automatic) formal proof, and it is still satisfied even in *some* of the cases when one ends up knowing the actual vote cast by V_A and V_B no matter what the protocol is (for example, if the vote is unanimous).
Unfortunately, this definition lets seamlessly go through all of the voting schemes we have analyzed in this paper: in no single one of these it is possible to tell the difference among two cowards following the boss' instructions, and two lucky heros that both disobey the boss but end up casting exactly the vote that the other one was requested to cast. And yet, not any single one of these schemes provide by any means, in practice, the least level of anonymity.
Indeed, if one looks at all the issues more carefully, the extremes to which this proposed definition is not sensitive, like the unanimous vote, or the single voter election, are nothing more than particular cases of the ballot-as-signature attack: one voter elections is the extreme attained with precinct size equal to 1, and then the chances of getting caught when disobeying are also equal to 1; an election ending in an unanimous vote is the extreme where the number of possible ballot configurations is equivalent to one, and there also the chances of getting caught are 1. But in between, with more than one voter, and more than one *observable* ballot configuration, we have a full range of privacy probabilities, that a satisfactory formal model should properly account for.
We do not need a formalization that *avoids* the two extremes by being insensitive to them, but one that can *explain* and *compute* these two extremes, *and all the range in between them*.
Of course, all this applies as well to the subsequent definitions of *receipt freeness* and *coercion freeness* in [DKR06].
Building on the analysis we have done of the different voting scheme, and on the ideas put forward about possible countermeasures, we can formulate the following

REMARK 2. *The anonymity of a vote cast during an election strongly depends on the election method used; in traditional ranked vote methods, anonymity is not an absolute property, but a probabilistic one, which is a function of*

- *the size N of the electoral precinct,*
- *the number Nv of distinct votes that can be cast,*
- *the probability distribution of these distinct votes,*
- *the number Nov of the* observable *distinct individual vote outcomes,*
- *the probability distribution of these observable votes.*

This may give a first guideline for designing a satisfactory formal specification of the anonymity property, even if, in practice, the degree of risk the voters are likely to accept has also an impact on the level of anonymity of the scheme, which makes a precise estimate quite challenging, as this last element is in turn function of the *perceived* probability of remaining anonymous.

8 Conclusions

We have presented in detail a powerful signature attack on privacy and anonymity of voting methods based on the *Italian trick*: building on the lesson learned from actual election fraud occurred in Italy in the 80's, we construct a signature out of *the very ballot data* that is necessary for a correct tallying of the election results.
This approach is clearly very different than the usual attacks on privacy and anonymity in electronic voting protocols, where, by the very nature of electronic voting, and in particular remote voting, one usually concentrates on possible information leak concerning the identity of the voter coming from the authentication phase, which can be complex and very error prone, but has nothing to do with the actual *content* of the voter's ballot.
It is also much more powerful: we have proven its effectiveness against a wealth of voting methods: some of them, like the US election system, have been in used for quite a long time; others were proposed just a few months ago by top experts in cryptographic protocols on one side and well known economists and mathematicians on the other side.
The attack described here makes it definitely clear that, unlike what is done in a great deal of research works on electronic voting, one *cannot* abstract away so-called *details* of a voting method when performing an assessment of properties like privacy and anonymity: the actual *content* of a voter's ballot matters; the *size* and *nature* of the possible voter's choices contained in the ballot matters; the *precinct size* matters.
As a consequence, before actually implementing *any* concrete solution based on the vast existing literature on electronic voting protocols claiming to preserve privacy and anonymity, one should carefully check the *precise* hypothesis made by the proponents of each protocol: do the good properties still hold as soon as one has more than two choices (yes/no) when voting? what assumptions are made about the ballot content, the public verifiability of the vote, the precinct size? It might be the case that under some restrictive assumptions the protocol really preserves anonymity and privacy, and that the instance one implements violates these assumptions, as it happens quite often when the implementation proposes multiple choices or multiple elections in a single ballot.
We have seen in the case of the US elections that aggregating several elections into one single ballot may seem a good idea when it is difficult to bring citizens to the voting booth more than once every few years, but is actually catastrophic in terms of privacy and anonymity.
The lesson learned there is quite clear: privacy and anonymity *do not compose*: if we have a voting method with a high probability of privacy, for example a binary vote on a single candidate, or a single-winner choice among several candidates, we *cannot* compose them in a single ballot preserving the same degree of privacy.
We have also seen at length in all the examples given in this paper that privacy and anonymity are actually *probabilistic* and not *absolute* notions, so the current efforts at formalizing these properties that are known to this author are not yet satisfactory: one needs adequate formalisms taking these features in account when looking for a model of such properties.
It would be probably beneficial not only to the research community, but to the full population, if the arguments presented here are widely circulated and discussed, leading hopefully to more secure voting methods and/or protocols, but especially and urgently warning once again people that security is a risky business, and so much so when properties like privacy and anonymity are at stake.

9 Acknowledgements

The author is grateful to Jerôme Vouillon, Pierre Letouzey, Alexandre Miquel, Stéphanie Delaune, Delia Kesner, José-Maria Fullana, Dominique Poulalhon and many others for bearing with me while telling this story so many times, and dropping in their offices to discuss this or that detail of the arguments or the calculations. Special thanks to Andrew Appel for pointers to relevant related work on the Three Ballot scheme [Str06b, Str06a, App06].

10 References

[AF01] Martín Abadi and Cédric Fournet. Mobile values, new names, and secure communication. In *28th ACM Symposium on Principles of Programming Languages (POPL'01)*, pages 104–115, 2001.

[App06] Andrew W. Appel. How to defeat rivest's threeballot voting system. Technical report, Princeton University, October 2006. Available as `http://www.cs.princeton.edu/~appel/papers/DefeatingThreeBallot.pdf`.

[Arr51] Kenneth J. Arrow. *Social Choice and Individual Values*. New York: John Wiley, 1951.

[BL06] Michel Balinski and Rida Laraki. A theory of measuring, electing and ranking. Cahier 2006-11, Ecole Plytechnique, Laboratoire d'Economtrie, 1, Rue Descartes, 75005 Paris, 28November 2006. Available as `http://ceco.polytechnique.fr/fichiers/ceco/publications/pdf/2006-11-29-1528.pdf`.

[BT94] Josh Cohen Benaloh and Dwight Tuinstra. Receipt-free secret-ballot elections (extended abstract). In *STOC '94: Proceedings of the twenty-sixth annual ACM symposium on Theory of computing*, pages 544–553, New York, NY, USA, 1994. ACM Press.

[DKR06] Stéphanie Delaune, Steve Kremer, and Mark D. Ryan. Verifying properties of electronic voting protocols. In *Proceedings of the IAVoSS Workshop On Trustworthy Elections (WOTE'06)*, pages 45–52, Cambridge, UK, June 2006.

[Ghe02] Richard Ghevontian. La sincrit du scrutin: La notion de sincrit du scrutin. *Cahiers du Conseil Constitutionnel*, (13), 2002. Available online as `http://www.conseil-constitutionnel.fr/cahiers/ccc13/ccc13somm.htm`.

[JCJ02] Ari Juels, Dario Catalano, and Markus Jakobsson. Coercion-resistant electronic elections. Cryptology ePrint Archive, Report 2002/165, 2002. `http://eprint.iacr.org/`.

[Jon05] Douglas W. Jones. Threats to voting systems. `http://www.cs.uiowa.edu/~jones/voting/nist2005.shtml`, 7October 2005. Presented at the Workshop on Developing an Analysis of Threats to Voting Systems National Institute of Standards and Technology.

[KR05] Steve Kremer and Mark D. Ryan. Analysis of an electronic voting protocol in the applied pi-calculus. In Mooly Sagiv, editor, *Programming Languages and Systems — Proceedings of the 14th European Symposium on Programming (ESOP'05)*, volume 3444 of *Lecture Notes in Computer Science*, pages 186–200, Edinburgh, U.K., April 2005. Springer.

[Low95] Gavin Lowe. An attack on the needham-schroeder public-key authentication protocol. *Inf. Process. Lett.*, 56(3):131–133, 1995.

[Riv06] Ronald L. Rivest. The threeballot voting system. Technical report, Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, 1October 2006.

[Sch99] Berry Schoenmakers. A simple publicly verifiable secret sharing scheme and its application to electronic voting. In *Advances in Crytoplogy – CRYPTO '99*, volume 1666 of *Lecture Notes in Computer Science*, pages 148–164. Springer-Verlag, 1999.

[Ste57] Lincloln Steffens. *The Shame of the Cities*. Hill & Wang, 1957. Reprint of the 1905 edition.

[Str06a] Charlie Strauss. A critical review of the triple ballot voting system. part 2: Cracking the triple ballot encryption. Available as `http://www.cs.princeton.edu/~appel/voting/Strauss-ThreeBallotCritique2v1.5.pdf`, 8October 2006.

[Str06b] Charlie Strauss. The trouble with triples: A critical review of the triple ballot (3ballot) scheme, part 1. Available as `http://www.cs.princeton.edu/~appel/voting/Strauss-TroubleWithTriples.pdf`, 5October 2006.