Journal of Information Security and Applications 35 (2017) 168–174

Contents lists available at ScienceDirect

Journal of Information Security and Applications

journal homepage: www.elsevier.com/locate/jisa

The risk to breach vote privacy by unanimous voting

Peter Ullrich

Universität Koblenz-Landau, Fachbereich 3, Mathematisches Institut, Universitätsstraße 1, 56070 Koblenz, Germany

ARTICLE INFO

Article history:

Keywords:
Breach of vote privacy
Unanimous voting
Small subsets of voters

ABSTRACT

The paper studies the risk that all members of a set of voters give their votes unanimously and thereby breach the privacy of the voting procedure. This problem becomes relevant in the situation that the voting behavior of a small (sub)set of voters can be identified by the way they transmit their votes, e. g., when at least two possible ways to give votes (like voting with ballot boxes, postal voting, and electronic voting) are admitted in theory but one of them is used by only a small minority of voters in practice.

For the situation of a simple alternative between "yes" and "no" it turns out that as long as the probability of approval lies between 25% and 75% the probability of a breach of vote privacy by unanimous voting is smaller than 1% if there are at least 17 voters and even smaller than 0.1% if there are at least 25 voters. If, however, the rate of approval or disapproval increases, even to values already observed in reality, then the probability of such a breach of vote privacy can no longer be neglected. And even small values for the probability of a breach of vote privacy sum up when several thousands of these situations appear in parallel.

Furthermore, if there is a three valued situation "yes" – "no" – "abstention" present, then, depending on the concept of vote privacy, a breach of it becomes considerably more probable even if the probability of approval remains within the boundaries mentioned above.

1. Introduction

One of the basic conditions a voting process has to fulfil is to guarantee privacy for all voters in the sense that for no voter one can find out the individual voting behavior.

Contrary to this statement on the principle one has to take into account the possibility of unanimous results which can happen even if the voters give their votes independently of each other. In these cases it is revealed either that each voter has given a "yes" to the proposal or that each voter has given a "no". Both cases have the same result that the privacy of the voting procedure is breached.

For large sets of voters, of course, it is very improbable that all voters give a "yes" or all a "no" to the proposal. And one tends to accept such incidents if there are only very few voters who also know each other well enough so that the loss of vote privacy is more or less fictitious.

The phenomenon can no longer be neglected, however, if the two situations described above mix in the following way: Suppose that there are altogether enough voters that an unanimous voting of the complete set is improbable enough. But within the complete set of voters let there be a small subset such that the result of the voting of this subset can be identified because of whatsoever reason. If this subset is small enough then the risk has to be considered that the members of this small subset vote unanimously so that also the individual voting behavior of all of its members becomes known and the privacy of the voting is breached, hence.

This problem is not new in principle: It becomes present as soon as there are at least two ways to transmit votes like using ballot boxes or, alternatively, postal voting.

In Germany, for example, legal texts concerning election procedures already explicitly mention this problem: The regulation for federal elections in Germany ("Bundeswahlordnung") generally prescribes that the voting precincts ("Wahlbezirke"), i. .e., the subsets of voters for which partial results of the election are determined, have to encompass enough voters that it is impossible to find out how single voters have voted [1, Section 12, (2), sentence 3]. This prescription is made concrete for the case of postal voting by demanding that such a subset should consist of at least 50 voters [1, Section 7, No. 1]. For the case of voting by use of ballot boxes, however, smaller numbers are accepted in legal practice. The most extreme example is the "Hallig Gröde", a small island in the North Sea, where only 9 voters live. (For the last federal elections in 2013 all of them chose postal voting in order to save the privacy of their votes. For the state elections of 2017, however, they used again their ballot box, which made them an object of public interest since their voting behaviour differed considerably from the

E-mail address: ullrich@uni-koblenz.de

http://dx.doi.org/10.1016/j.jisa.2017.07.001
2214-2126/

average: Three of the nine gave their votes to a party that got only 1.2% of all of the votes in the state of Schleswig-Holstein.)

E-voting firstly gives a further way to transmit votes so that the number of possibilities is increased. But there is also a much more serious effect: In the case of voting with the use of ballot boxes vs. postal voting one can shape (not only the ballots themselves but) also the envelopes for the ballots in exactly the same manner for both ways of transmission so that one can mix them and by this make the votes indistinguishable. For votes transmitted electronically, however, the different media make it technically difficult to mix the votes given by the different ways of transmission and thereby hide the partials results for these ways.

Furthermore, because of the advantages of e-voting for the voter its introduction to the election system may lead to a rapid decrease in the number of voters who use the other ways: In the 2012 elections of the German Informatics Society ("Gesellschaft für Informatik", GI) for its council ("Präsidium") and its board ("Vorstand"), for example, 2671 of the members used electronic voting, but only 42 postal voting, which is less than 2% of the voters altogether.

In principle, however, the argument in the present article is not specific for e-voting or for any other way to transmit the votes. In particular, it does not depend on the concrete voting system and its implementation at all besides the only fact that there are at least two different ways to transmit the votes. Contrary to this, the paper [2] describes a possibility to breach the privacy of a particular paper ballot voting system by using its particularities. According to these results the vote of up to 96% of the voters can be correctly recovered. Contrary to this, the present paper is concerned with the breach of vote privacy for only small subsets of voters, namely those that transmitted their votes by the barely used way. But the only information that is used for this breach is that the voter under consideration has used this way of transmission, not any other information on the voting system or the particular act of voting, e.g., the time when it took place.

For sake of simplicity the paper starts by examining the situation that there is only one voting decision to be made. At first, only a dichotomic voting is considered, i. e., that only the two alternatives "yes" and "no" can be realized (Section 3.1). The consequences of the theoretical result are evaluated on the basis of emprical data gained from elections of several scientific societies (Section 3.2).

Then the situation of trichotomic voting is discussed in a similar way, i. e., that besides approval and disproval there is the possibility of abstention. It turns out that this makes necessary a discussion of what "keeping of vote privacy" should mean in this context (Section 4.1). Also a short generalization of this discussion is given for the case that three or more alternatives are presented to the voter (Section 5).

In the sequel, these results are generalized to the case of several voting decisions at a time, both if these are independent (Section 6.1) and dependent of one another (Section 6.2). Furthermore, an application is given for the situation of a voting on the scale of a state which is divided into many voting precincts.

The paper closes with a discussion of the conclusions and the possibilities to reduce the risk to breach vote privacy in the way described above.

2. Notation and assumptions

For the discussion two assumptions are made:

- There exists a way to find out the probabilities with which the small set of voters under consideration takes the relevant actions. At least, one should be able to give boundaries for these probabilities.
  This is the case, for example, when the small group of voters on the average has the same voting behavior as the group of all voters and when one can use the outcomes of previous voting procedures as a predictor for the voting under consideration.
- All voters, in particular those from the small subset, give their votes independently from each other.
  This is plausible since the results of the voting process will only become known after the possibility to give votes has ended. In Section 6, however, also the situation is studied that several voting procedures take place at a time or that the whole community of voters is divided into several precincts where the probabilities are not necessarily equal.

Throughout this paper the following symbols will be used

n for the number of members of the small set of voters (which is supposed to be strictly positive),
p for the probability that a vote is "yes" (or something similar definitely positive),
in the dichotomic case: $q = 1 - p$ for the probability that a vote is not "yes",
in the trichotomic case: r for the probability that a vote is "no" (or something similar definitely negative),
in the trichotomic case: $s = 1 - (p + r)$ for the probability that a vote is "neither – nor", and
P for the probability that a breach of vote privacy takes place by an unanimous voting of the n voters.

In the case of several decisions their respective probabilities will be indicated by indices.

3. One single dichotomic vote

3.1. Determination of the probability

At first, consider the following situation:

Each voter is confronted with only one voting decision and has exactly two alternatives: to say "yes" or to do the contrary, so to speak, to give a "non-yes" regardless whether this means that (s)he simply does not say "yes" or that (s)he explicitly says "no" (where in the latter situation (s)he does not have the possibility of abstention).

Let p denote the probability that a voter says "yes" and $q = 1 - p$ the probability of the contrary.

Since it has been assumed that all members of the subset of voters under consideration give their votes independently, the law of multiplication gives that the probability that all n voters unanimously give a "yes" equals

$$p^n$$

whereas the probability that all n voters unanimously give a "non-yes" equals

$$q^n.$$

Since the two possibilities of unanimous voting described above are mutually exclusive, the resulting total probability of a breach of vote privacy by unanimous voting equals

$$P = p^n + q^n$$

where p and q underly the condition $p + q = 1$. Therefore as a function of p alone one has

$$P = P(p) = p^n + (1 - p)^n.$$

This expression for P has the following properties: