Experiments and Data Analysis of Electronic Voting System

Komninist Weldemariam\textsuperscript{1,2}  
\textsuperscript{1}Dept. of Eng. and Comp. Science  
University of Trento  
Trento, Italy  
Email: sisai@fbk.eu  

Adolfo Villafiorita  
\textsuperscript{2}eDemocracy Group  
Fondazione Bruno Kessler  
(fbk-irst) Trento, Italy  
Email: adolfo@fbk.eu  

Andrea Mattioli  
\textsuperscript{2}eDemocracy Group  
Fondazione Bruno Kessler  
(fbk-irst) Trento, Italy  
Email: amattioli@fbk.eu  

Abstract—Experimental data sets related to e-voting systems are very demanding in order to improve currently deployed e-voting machines. Unfortunately, the studies of such data about the machines’ security, performance and their evolution with respect to the social and technical aspects are still unsatisfactory.

During the last four years we have been involved in the development, experimentation, and evaluation of an e-voting system. The system tried out in several regular elections, and also used in two small elections with legal value. Each experiment provided various sociological (e.g. citizens’ opinions on the system) and technical data that are related to system’s performance and behavior.

In this paper, we present various technical insights and the lessons learned during the e-voting experiment. The methodology for the various experiments we have carried-out and the data sets collection process are also discussed. This helps to confirm existing data on the subject (e.g., data related to security, procedures and logistics) and, in some cases, provide novel information or, at least, shed a new perspective on some security-critical factors concerning e-voting systems.

Keywords—Electronic Voting, E-Democracy, Experiments, Security, Assessment

I. INTRODUCTION

The development of fair and secure e-voting machines is complex (see, e.g., [1], [2], [3] and [4], [5], [6] for some discussion on the matter).

We have been involved in the development and evaluation of an e-voting system called ProVotE. During the project we developed a DRE-VVPAT mainly driven by the possibility of incorporating feedback from voters, of guiding the design based on other experiences and recommendations (e.g. [7], [8], [9]), and of choosing technologies and components. On top of that, this gave us to adopt formal verification techniques for the development of some critical components and for the assessment of procedural security (see [10], [11]).

Though improved gradually over the years, the studies of experimental data about the e-voting machines’ security, performance and their evolution with respect to the social and technical aspects are still unsatisfactory. This limits their usage in a larger scope. For example, data set on observing security threats to voter’s anonymity by following standard procedures that illustrate each machine’s behavior during elections can help raise the transparency in elections using electronic devices and increase the confidence of voters in the democratic system. Data sets related to the process of setting up experiments, running an election, and performing security evaluation across various voting machines (e.g., as in Diebold [12] and ES&S [6]) provide information about the behavior of machines under malicious circumstances, if they are not designed carefully, and provide need to consider recommendations for design alternatives.

To mitigate the risks mentioned above, we have devised a development process organized in cycles ended by running experiments during local elections. Three different (hardware) prototypes of the e-voting machine in five different electoral events were experimented so far. During these experiments various data about the social impacts and technical performances of the machines were collected.

The focus of this paper is, therefore, presenting our findings and our conclusions on the technical data (see [13], [14], for work related to the social aspects) collected during the experiments.

In the next section, we recall the voting process for the ProVotE e-voting system. Section III presents a methodology for the e-voting experiments. A detail discussion on the technological insight of the collected data is given in Section IV. Section V presents the lessons learned and findings from the experiments. Related works on the subject matter and our conclusion are presented in Section VI and VII, respectively.

II. ELECTRONIC VOTING SYSTEMS

For the purposes of this work (see [15] for a more detailed and complete view), the ProVotE e-voting system is composed of:

- A Touchscreen: Used to administer the machine and cast votes. It is a 20”-wide screen, to improve readability.
- A Thermal printer: Protected behind a glass, it has a paper roll which is 12” wide, to improve readability, and it is equipped with a cutter.
- A smartcard reader: Access to the machine is controlled by means of a smartcard reader, which is in-
stalled externally to the cabin and is only operated by the polling officer.

- **An external display**: An LCD display (20x2 characters), placed on the table of the poll officer, shows some basic information on the voting process, such as the number of voters, the status of the machine (idle or busy), and the error conditions (e.g., printer nearing end of paper).

- **An uninterruptible power supply (UPS)**: It provides emergency power for a voter to complete the voting procedure before halting. The machine is not meant to be used in case of a long lasting power outage.

- **A computer**: The computer (a standard PC) controls all the components of the voting machine.

![](image)

**Figure 1. The ProVoE Voting Machine.**

Once it has been confirmed or rejected by the voter, each vote is cut and falls in a ballot box installed inside the machine below the printer. This protects from possible breaches to anonymity due to the possibility of tracing the order in which votes have been cast, as it might happen with VVPAT with a continuous tape. Notice that the reason for protecting the printout behind a glass is to reduce risks related to tricky interpretation of votes (e.g., vote tempering) and possible compromises to the synchronization between electronic and paper data (e.g., attacks which could jeopardize the integrity of election results).

### III. A Methodology for the Experiment

One of the goal of the project is evaluating the e-voting systems in a real scenario, in order to understand the main criticalities faced by voters, poll officers, public administration officials, and systems alike.

In order to do so, the development of the e-voting system has been staged in cycles, each cycle ending in correspondence with the end of an electoral turn. This allowed to deliver prototypes of the system in time for elections and to setup experimentations during electoral turns. The advantage of the approach, very similar to the spiral development process [16], is that not only we could “observe” electronic elections from a privileged point of view, but also we could use the outputs of the experimentation to drive the development for the next cycle.

To be systematic and repeatable across experiments, we defined a methodology organized in the following order:

- **Goals Definition**, during which the technological, procedural, and sociological goals of the upcoming experiments are defined. This also served to prioritize the development of the machine functions.

- **Preparation**, during which the goals of the experimentation and the e-voting system have been presented to the voters interested by the experimentation. According to the experiment size, the preparation phase was conducted in different ways, e.g., using letters, public presentations, video advertisements.

- **Free simulation**, during which some e-voting machines have been made available in public places for citizens to try. The e-voting machines for free-simulation need to be shipped in a special configuration that allows citizens to vote more than once. The ballot shown to voters for free simulation contained randomly generated candidates names and list symbols chosen so that they could not to resemble any local or national party.

- **Training**, during which poll officers have been trained on the usage of the machines and on the procedures regulating the experiment. The procedures have been devised to resemble those to be used in electronic elections with legal value. During the project we experimented alternative procedures, partly due to logistical constraints and partly to qualitatively evaluate advantages and disadvantages of alternative ways of conducting the electronic elections. Training usually lasted half a day. Size of classes ranged quite a bit, from about four people to about thirty people.

- **The experimentation**, during which the e-voting systems and the procedures were tried out. The experiments have always been conducted in parallel to regular paper elections, with the e-voting machines installed in the polling stations where the election was taking place and with dedicated personnel responsible for managing the machines. The polling officers of the electronic election were asked to follow the procedures like they were in an election with legal value and citizen were asked to repeat the vote they had cast on paper. The resulting framework is an excellent testbed, as it puts the machines and the polling officers under the same conditions of a real election. Moreover disattention (or reduced attention) due to the experimental value of the experiment, in some cases, allowed to also collect information about various non-nominal situations.

- **Data Collection and Analysis**, during and after each experiment, data was collected about: (i) citizens’ opinions on the system, through interviews in the polling stations and follows up, (ii) systems’ performances and criticalities, (iii) procedures’ main advantages and criticalities, and, of course, (iv) the candidates resulted
elected with the electronic devices (information, however, that does not have any legal value). As mentioned earlier, all collected data were used to verify whether the goal had been achieved and, if not, the kind of corrective actions to introduce for the next cycle of development and experiment.

IV. THE EXPERIMENTATION AND DATA COLLECTION PROCESS

This section presents a brief description of the experiments we conducted, discusses the data sets collection process, and provides general remarks on the collected data sets.

A. Experiment Description

Table I lists the experiments that have been conducted during the project. The first and the last experiments are the most significant in terms of participation. They correspond to general electoral turns, for which we selected the biggest possible subset given the resources available. The selection of the polling stations in which to experiment e-voting was driven by the Faculty of Sociology, that selected polling stations in order to include the widest possible range of voters’ age, scholarship, type of job, and geographical distributions. The other experiments were conducted during special electoral turns, that often involved just one municipality. We had little or no flexibility in sizing the experiment in those cases. Matter of fact, we nearly always covered all the polling stations in which elections were being held.

The column Voter is a count of the people listed in the polling stations equipped also with the e-voting machines and Turnout is the number of people who tried out the e-voting system. Notice that affluence to elections is usually about 80%. So, for instance, during the last experimentation, about 35225 (= 0.8\times14032) people showed up at the polling stations. Of these, about 50% tried the system. All data can be found from the website of the Autonomous Province of Trento, where people interested in participation to the experiment can get the precise figures.

Common to all the experiments is the usage of e-voting machines and of automated recounting systems (in some trials the automated recounting has been paired to a manual recount). The delivery of results varied from physical transportation to the Electoral Office of the memory supports with the results, to electronic upload through a web application. Concerning the systems for determining the elected candidates, we tried both a commercial system adopted by various electoral Services in Italy and an in-house solution. In one trial we also tested a system for automating registration of voters.

Finally, notice that in two cases the system was used with legal value. Environmental conditions and requirements for legal elections differ greatly from those for experiments. The discussion and results presented in this section, therefore, do not refer to these two cases.

B. Data Collection

We mentioned that this paper only focuses on the technological data collected during the experimentations, namely on the information concerning the performances of the machines and the procedures. Other papers (see e.g. [13], [14]) describe the sociological results of the experiment, such as, e.g., the attitude of citizens towards the e-voting systems and how the introduction of new technologies shifts the meaning of void and blank ballots.

Figure 2 describes the collection and analysis mechanism adopted during the experiment. On the left hand side we have technical support center (upper part) and the polling station (lower part), where the data is produced. On the right hand side we have the data collection and analysis part, conducted after the experiment (in all experiments data has been made available for analysis only after the end of the voting procedures):

- **Requests for Intervention.** As logged by the technical center, that includes the support requests received during the experiments. Such requests might range from procedural questions to technical problems related to the implementation of the procedures for interventions due to malfunctions.

- **Printed Logbooks.** During the elections poll officers were required to compile a pre-formatted printed logbook. The logbook is meant to record the main operations, all the significant events, and the main observations about the system and the procedures.

- **Machine Logs.** The machines produce a log of the activities performed during the election. The log can be used to automate part of the logging activities currently performed on paper (e.g. when the machine was made available for voting) and contain diagnostic information about software and hardware malfunctions.

The information is collected at different levels of details:

- At the lowest level, the logs collect information about whether the machines have been turned on and off, whether they have been shut-down, whether any critical software or hardware error occurred.

- At the next level, the logs collect information about the life cycle of the machine, such as, e.g., the time at which the machine was configured, the time at which the results were tabulated.

- At the finest level of details, the logs collect information about single operations, such as the

\[1\] There are about 600 polling stations in Trentino. The general electoral turns involved all of them. The size of electronic experiments was mainly limited by the number of e-voting machines available and by logistics.

\[2\] http://www.elezioni.provincia.tn.it

107
machine being enabled for voting or the machine being locked after a vote cast.

Although obvious, it is important to remark that the logs do not collect data that could compromise secrecy and anonymity. This has been achieved by imposing requirements on what data could be collected with the logs and providing separate protection mechanisms for the critical data. For instance, votes are stored in a single, shuffled file, so that no vote has a time stamp associated; the sequence in which votes are found is not the one in which they have been cast. Moreover, the machines do not have access to any information about the voters; the voters’ register is kept separately.

- **Electronic Election Results.** All experiments have been concluded by determining the candidates elected, according to the results of the electronic votes. From the technological point of view, determining the elected candidates represented the minimum goal to achieve to ensure that all procedures had been conducted to an end. From a different perspective, the results have been compared with those of the elections conducted on paper, to verify whether the electronic voting and the paper voting had the same trends and, thus provide an indirect hint on the ease of use of the system.

- **Printed Ballots.** All the machines produced printed ballots, that contain a physical, voter verified proof of each vote cast. The right hand side of the figure shows how the data collected during the experimentation has been used. In particular:

  - the paper logbooks and the intervention requests have been analyzed by hand in order to highlight the significant events. Not all the requests of intervention are due to a technical problem, not all the events signalled in the logbook are relevant.

  - the machine logs have been analyzed with a specifically developed tool, called — not surprisingly — *Log Analyzer*, in order to highlight relevant sequences of events (turn-on, turn-off sequences) and information about such sequences (how long the machine has been available for voting). For certain kind of information (e.g., affluence), the *Log Analyzer* aggregates the data of each machine to provide total values.

  - the electronic results have been compared with the printed ballots, to make sure no discrepancies arose, that is, each electronic vote corresponded to a printout. Recounting has been performed with a specifically developed tool, called *Paper Ballot Counter*. The *Paper Ballot Counter* reads the barcodes printed on the ballots, show the correspondent ballot on screen and updates the total votes counter.

  - the electronic results, finally, have been compared with the results of the paper election, to verify whether the same trends emerged.

Notice that, since the data about the same phenomenon is collected in different ways, inconsistencies among different data might (and in various cases actually did) result. Inconsistencies, however, are not surprising and often not even a cause of great concern. Consider, for instance, the case of the Printed Logbooks, in which poll officers write information about what they think they did and the machine log that registers what was actually done on the machine (on the assumption, of course, that the machine log is correct).

Slightly more surprising, at least for us, has been the fact that in some cases the data collected is incomplete and has not been always sufficient to uniquely identify the cause or the sequence of events actually performed during the experiments. From the technological point of view, this has been caused by the limited observability we have on certain hardware components, like, for instance the printer. The “paper jammed” and the “no paper” conditions correspond to the same error code, there is no way of discerning between the two cases when analyzing the logs (e.g. was the machine started with no paper roll or did the paper roll jammed during the start-up testing operation?). In many cases resorting to the information provided by the poll officers has relieved helpful. From the procedural point of view, the incompleteness of information has been caused by various causes, among which we mention: imprecise logs (poll officers might have not been accurate or did not consider significant some events), flexibility in the implementation of the procedures, and level of detail of the information.

<table>
<thead>
<tr>
<th>Date</th>
<th>Election</th>
<th>Legal</th>
<th>Cities</th>
<th>Polling Sites</th>
<th>Machines</th>
<th>Voters</th>
<th>Turnout</th>
</tr>
</thead>
<tbody>
<tr>
<td>May 2005</td>
<td>Municipality</td>
<td>no</td>
<td>5</td>
<td>16</td>
<td>16</td>
<td>15375</td>
<td>6950</td>
</tr>
<tr>
<td>Nov 2005</td>
<td>Municipality</td>
<td>no</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>554</td>
<td>336</td>
</tr>
<tr>
<td>May 2006</td>
<td>Municipality</td>
<td>no</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>1650</td>
<td>1134</td>
</tr>
<tr>
<td>Nov 2006</td>
<td>Municipality</td>
<td>no</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>2334</td>
<td>1361</td>
</tr>
<tr>
<td>Nov 2006</td>
<td>School</td>
<td>yes</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>1392</td>
<td>1207</td>
</tr>
<tr>
<td>Nov 2007</td>
<td>Referendum</td>
<td>yes</td>
<td>2</td>
<td>4</td>
<td>1082</td>
<td>561</td>
<td></td>
</tr>
<tr>
<td>Nov 2008</td>
<td>Provincial</td>
<td>no</td>
<td>39</td>
<td>55</td>
<td>110</td>
<td>44032</td>
<td>16939</td>
</tr>
</tbody>
</table>

Table 1

**Experimental results using the ProVotE system.**
logbooks required to report.

Notice that both inconsistency and incompleteness of information can happen also in paper elections. In the case of electronic elections, however, we believe there is a concrete opportunity and need to try and improve the current situation.

C. General Remarks

Below we briefly summarize some general remarks concerning the experiments we have conducted.

- **Size.** With respect to size (number of polling stations and citizens involved), we conducted what is probably the biggest e-voting experiment in Italy. It is certainly the only one for which data about participation and electronic results are publicly available.

- **Participation.** The percentage of participation to the experiment has been higher for the small trials. In the first experiment, an explanation for the relatively low participation is that we equipped each polling station with just one machine and queues formed. The second most likely motivation is that smaller experimentations allowed for a more thorough “preparation” phase, possibly raising awareness about the importance of the trials.

- **Results Comparison.** The results of the electronic and of the traditional election are quite similar, in some cases up to the candidates elected with the fewest votes. The percentage of blank and null votes, however, is significantly higher in the electronic case. For instance, at the last experimentation, the percentage of blank votes in the electronic case is 3.8%, about four times the ratio of the traditional election (0.9%). Null votes were 2.1%, nearly twice as much as in the traditional election (1.2%). See [13] for a discussion of these results.

- **Throughput.** Voting time is about 45 seconds (+/- 34 secs) per voter. Taking into account the standard opening hours of a polling station, a single voting machine could be used by about 1200 voters; since

---

**Figure 2.** The Experiment Data Collection.
polling stations usually enlist about one thousand people, in principle, one machine per polling stations could be enough to run an electronic election. The figure, however, does not take into account different density in affluence (certain hours attract more voters than others), and, above all, liveness issues (e.g. making sure elections can run in spite of single failures).

- **Recounting.** In all the elections (including those with legal value) the printed ballots have been recounted and compared with the electronic results. In the “smaller” experimentation recounting took about one day. In the biggest experimentation, recounting was performed in 2.5 days using two recountings stations running in parallel.

V. SOME LESSONS LEARNED

In this section we summarize some of the main conclusions emerged from the data collected during the experiments. Notice that inconsistence and incoherence of the data collected imply that, in various cases, we had to speculate and make assumptions on the possible causes of certain events. The conclusions drawn here, therefore are of the authors of this paper only, they are not the only possible conclusion that can be drawn from the data, and they might or might not represent the views of the other stakeholders involved in the project. Notice also that the examples described below do not refer to events actually occurred.

1) **Choices on Conflicting Requirements.** The first consideration is probably common to all engineering systems: requirements are often conflicting and choices have to be made among which ones have to be privileged. Consider, for instance, the voter user interface we developed. In order to increase readability and ease of use, we decided to use a large touchscreen (twenty inches), a printer with a very wide paper roll, and a software interface that reproduces the paper ballot. This allowed to improve readability and ease of use sensibly. The implication of such choices, however, is increased technical complexity of the machine and increased costs. Concerning the technical complexity, consider the case of the printed ballots, which are 19 cm long and 11 cm wide. Ensuring that the ballot is properly presented to the voter, while at the same time the content of the ballot box (in which ballot boxes fall when confirmed by the voter) is not visible, required the development of an ad-hoc solution. This, in turn, raised the development and building costs.

2) **Security and Liveness.** Increased security might have a cost in terms of liveness of the process, if procedural errors occur. This cost might or might not be acceptable, according to the electoral context. Consider the life cycle of the e-voting machines. It has been devised to ensure that poll officers need to stick to a rather precise sequence of actions, starting with configuration and ending with tabulation of data. After tabulations, the machines cannot be used anymore and special intervention, that can be performed only by a few authorized profiles, is needed to make them ready for the next election. This ensures that procedural errors are less likely to occur (think, e.g., of some reported cases in the U.S. in which the machines have been used in testing mode, rather than in voting mode) and reduces some possible attacks (think, e.g., of a malicious poll officer using the machine twice to produce fake results, after the election is closed). However, if a machine is wrongly delivered to the polling station prior to having been serviced, it will not be usable, and special intervention will be needed at the beginning of the electoral day. If all the machines in a polling station are delivered in such a state, then voters showing up will have to wait for the machines to be serviced. If they are not willing to wait, they might claim that their right to vote has been denied.

3) **Improving Accountability and Traceability of the Process.** Procedures are meant to ensure that the machines are operated as expected and that elections are run fairly. The descriptions of procedures, however, might be incomplete or inaccurate and might not be fully committed to their applications. The only solution we envisage to this situation is embedding as much as possible of the procedures in the life cycle of the machine, to disambiguate the application of the procedures and increase involvement and awareness of poll officers. (See, however, the previous point.)

4) **Reduce Operational Stress.** Preliminary operations on the machines (tests and configuration) are sometimes conducted under the pressure of the first voters showing up at the polling station. Any nuisance, in the start up operations (e.g. a key to open a machine is lost; a machine has not been correctly installed in the polling station), add to the stress and increase the probability of errors. One solution to this situation is anticipating testing procedures to the day before the election. This has also the advantage of having the poll officers further familiarize with the system at an increased cost (personnel costs).

5) **Training and motivation.** Training and motivation of the poll officers is probably the most critical factor in ensuring a successful electronic election. Various activities concerning poll site management, in fact, have little or nothing to do with e-voting machines. Moreover, no matter how careful people might have been in devising systems and procedures, exceptions and problems will occur. The possibility of recovering from such situations and, therefore, ensure fair elections, remains largely under the control of poll officers, no matter what kind of systems is implemented or used.

VI. RELATED WORK

Work to rigorously define e-voting properties, attack models and languages for describing the counter-measurements is still preliminary. The application and usage of techniques (such as, formal verification methods and tools) for these
systems is even more limited. Moreover the practice of collecting, analyzing, and disseminating data sets related to various election (security) incidences that can happen before, during and/or after the elections is also uncommon. Below we list works related to the matters mentioned above (see, e.g., [4], [5], [6], [17], [18], [19], [20] for development, evaluation and criticism of e-voting systems).

Work in the past has focused on understanding what changes could be introduced in the “traditional” voting procedures to allow a secure transition to electronic elections. For instance, [21], [22] discuss risks and difficulties related to the introduction of e-voting. [23], [24] suggest possible improvements to existing procedures, and [25], [26], [27] introduce techniques to formally analyze what security breaches may be derived by executing the procedures in the wrong way.

Assessments of some existing e-voting systems have been done [6], [28], [19]. These primarily focus on the design and implementation flaws, which could be exploited to compromise and invalidate elections. The authors also suggest a drastic change in the way in which e-voting systems are designed, developed, and tested.

VII. CONCLUSIONS AND FUTURE WORK

The development of e-voting systems is extremely challenging and demanding. The need to balance conflicting requirements, such as traceability and privacy, liveness and security, adds to the complexity of building and deploying application on which our democratic rights might depend.

We built an e-voting system that has been tried out in various experiments and used in two elections with legal value. Some of the interesting aspects of our approach are the setup of a develop-test-revise cycle in the development of the machine, the setup of a standardized and planned procedure for experimentation, and the systematic collection and analysis of data about the experimentations. The information collected during the experimentations has provided us with detailed insights on what we could do to improve on systems and procedures.

As a side-effect of our approach, the type and amount of data collected with the electronic systems, moreover, allows to get a more precise picture of what happens during an experimentation. We believe that extending such mechanisms also to the case of electronic elections with legal value could help improve security, traceability and accountability, while, at the same time, preserving the fundamental rights of any democratic election.

Some of the possible future directions include the standardization of the procedures for the collection and the definition of technical standards about the data from e-voting machines. Moreover we plan to use the collected data to qualitatively and quantitatively assess the failures and their relative occurrences both at system and at procedural level. One of the results we expect is being able to prioritize faults and therefore define a minimal set of appropriate countermeasures.

REFERENCES


