Automated Election Auditing of DRE Audit Logs

P. Baxter[1], A. Edmundson[2], K. D. Ortiz[3], A. M. Quevedo[4], S. Rodríguez[5], C. Sturton[6], and D. Wagner[6]

[1]*Clemson University*
[2]*Cornell University*
[3]*University of Puerto Rico-Arecibo*
[4]*Miami Dade College*
[5]*University of Puerto Rico-Mayagüez*
[6]*University of California-Berkeley*

Abstract

Voting audit logs produced by Direct Recording Electronic (DRE) machines are often unwieldy and unintelligible to a human reader. These logs detail all events recorded on the DREs and include data on ballots cast and post-election procedures. The authors of the paper "Auditing a DRE-Based Election in South Carolina" [5] demonstrated that these logs can be analyzed to uncover both procedural errors and election anomalies. In this study, we replicate the results from the aforementioned paper and develope additional analyses. These include identification of procedural errors by election officials, DRE hardware problems and precinct statistics. We have integrated these reports into a public website that produces a detailed report from the ES&S iVotronic log files. We intend this work to stand as proof-of-concept software for future auditing tools and as an immediately accessible tool to assist those working with election auditing and integrity.

1 Introduction

Currently in the United States, Direct Recording Electronic (DRE) units are used widely. A DRE is a type of electronic voting machine in which the voter interacts directly with the machine, typically through a touch screen. DREs provide a friendly interface to assist the voter with the ballot marking process. Compared to other widely used voting systems such as optical scan systems, DRE units can prevent overvoting and reduce undervoting. Additionally, audio DREs can assist visually impaired voters.

However, one drawback of using a paperless DRE system is that it does not generate a paper trail for use in post-election audits, there is no hardcopy version of each voter's ballot and the DRE's electronic record of votes cast must be accepted as correct. As a result, Federal standards require that electronic voting machines generate detailed event logs which can be used during post-election audits. These logs record events as they occur on the voting machine such as, "terminal opened", "vote cast" and "terminal closed". The log data may also include an electronic record of every ballot cast in the terminal. The authors of the paper, Auditing a DRE-Based Election in South Carolina [5], have shown how these logs can be analyzed to uncover procedural errors and anomalies that occur during the election. Manual analysis of raw data is usually cumbersome and time consuming, making countywide post-election analysis impractical and prone to human error. Therefore, at the present time, election officials do not regularly perform these types of analyses.

We aim to make DRE audit log analyses more useful and more accessible to election officials and other interested parties. In this work, we created a web application, available to anyone, that performs a variety of analyses on election data to detect procedural errors and system deficiencies. Our tool allows election officials to use machine audit logs to recreate election day events, identify possible voting machine errors or unexpected behavior and to identify areas of election day protocol that may need improvement.

Our research parallels with, and builds on, a similar study conducted with DRE audit data collected by fourteen South Carolina counties during the 2010 primary and general elections [5]. The authors of that study were able to determine, solely by analyzing the audit logs, that 1127 votes did not get included in the official certified tally in Richland county. These findings were possible because DRE systems used in South Carolina produce three different types of audit logs, each capturing slightly different information. By cross checking the logs against each other, the authors found inconsistencies that led them to uncover the missing votes. In our research we used the same data set and were able to replicate their results. We took this matter further and found fifteen memory devices containing votes that were not

uploaded to the tabulation systems from seven counties during the 2010 General election. These memory devices tallied 2082 total votes. Without additional information we could not verify whether alternate procedures were used to add those missing votes to the aggregated totals.

Our research was conducted with data from the iVotronic DRE system which is manufactured by Election Systems and Software (ES&S). The iVotronic system is a standalone, portable, touchscreen system that records vote totals, ballot images and an event log on internal flash memory. The event log records, in chronological order, the system events including unit configuration, polls opened, votes cast, polls closed, calibration or battery issues, warnings and system errors.

Election officials can use the analyses our tool provides to identify memory cartridges containing precinct totals that were not uploaded on election night, machines that may have experienced hardware problems during the election, locate polling locations that experienced lines of voters and determine which locations closed late. A brief description of our analyses follows.

Votes not uploaded. This analysis warns election officials of any memory cartridges used to close iVotronic terminals that did not have their data uploaded to the tabulation system. Our tool produces a report containing the precinct's name, the serial number of the iVotronic terminals collected in the cartridge and the total number votes stored on the cartridge. With this information, the election officials can quickly locate and upload the missing cartridges into the cumulative totals resulting in accurate election night reporting.

Machines not closed. This analysis produces a report listing the terminals that were not closed at the polling location. This analysis outputs the precinct's name and iVotronic serial number. With this information, election officials can quickly locate the terminals, close them out, and then upload their votes to the cumulative totals.

Missing terminals from the audit database. This analysis identifies any iVotronic terminals used during the election whose event log or ballot images have not been uploaded to the election reporting software. Election officials can locate the terminals or removable media containing the missing audit data so that the files can be uploaded to the database of the election reporting software. Complete DRE ballot images and event logs will allow for more accurate and complete post-election audits.

Polling location related analyses. Our tool provides a series of analyses related to polling location activity. We identify locations that closed late as well as locations that may have experienced long lines during the day. This information can help county officials to identify those locations which may need additional resources in the future.

DRE terminal configuration and hardware problems. Our tool performs several analyses that can identify iVotronic terminals that may need testing, repair or reconfiguration. These analyses include: identifying possible calibration issues, machines with low batteries, terminals that were forced to close early and machines with incorrect date and time settings.

Pollworker training related issues. Four analyses are available on our website which can be used to enhance the pollworker training curriculum. These analyses identify incorrect procedures at the precincts such as: using the wrong cartridges to close terminals in a precinct, failure to print the precinct's zero tape, activating ballots with the incorrect cartridges and the frequency of different reasons for cancelling ballots.

We based our study solely on the audit data logs which are considered the ground truth for election analyses and disputes. For this reason, it is important that audit logs are protected from accidental or malicious tampering. In this work we assume that DRE audit logs are complete, accurate and trustworthy. Detecting and preventing audit log tampering is a field of research outside of the scope of this study.

In summary, this study implements ways in which this data can be used meaningfully and in an automated fashion to enhance the accuracy and efficiency of elections. We believe our tool will provide intelligent feedback to election administrators during the canvassing process and post-election audits. We hope that this study serves to influence similar audits that can be expanded to other election technologies.

2 Background

2.1 Introduction to the iVotronic

Approximately 422 jurisdictions in the United States used the ES&S iVotronic electronic voting terminal in 2010. A brief description of its functionality and main system components follows:

- Voting terminal. The voting terminal is a standalone touchscreen voting unit. The ports available in the back of the terminal include: serial port, compact flash card slot and power supply port. The terminal is equipped with an internal battery which keeps the terminal operational during periods of power failure. To comply with federal standards, at least one audio terminal is placed in each precinct to assist visually impaired voters.

- Personalized Electronic Ballot (PEB). The PEB is a proprietary cartridge designed by ES&S to operate the iVotronic terminal. The PEB is placed in a slot located to the left of iVotronic's touchscreen. The terminal and the PEB communicate through the

infrared port. The South Carolina counties deploy two types of PEBs to the precinct: a) the green band master PEB and b) the red band activator PEB. Both types of PEB have the same functionality, however, poll workers are trained to perform the following tasks with each PEB type.

- Master PEB. Poll workers use the master PEB to open polls on election day. When the PEB is placed in the terminal, the touchscreen displays the precinct's name programmed in the PEB so that poll workers can verify the polling location information and date/time registered in the terminal's internal clock. If the information displayed is correct, the poll workers open the terminal for voting. The same master PEB should be used to open all terminals of the polling location. In the same fashion, the master PEB should be used to close all terminals of the polling location at the end of the voting day. When the terminal closes, it uploads its totals onto the master PEB. The master PEB accumulates the precinct totals which are used in the official tally.

- Activator PEB. This PEB is used by poll workers to activate ballots for voters. The number of activator PEBs that the election officials program for each precinct is proportional to the number of terminals and poll workers assigned to the precinct. The ratio varies depending on the jurisdiction criteria.

- Removable Compact Flash card (CF). The CF cards are programmed at Election Central and installed in the back of the voting terminal prior to precinct deployment. The CF cards contains graphic (bitmap) files read by the voting terminal during the voting process. The CF cards are also used as an external memory device: the audit log and ballot images are written to the CF card when the terminal is closed for voting. Once the polls close, the CF cards are removed from the back of the terminal and delivered to election headquarters on election night.

- External printer module. This module is connected to the serial port on the back of the voting terminal. The thermal printer produces the precinct zero tape and the results tape. Poll workers are instructed to print the zero tape once all iVotronics of the precinct are opened for voting. In the same fashion, the results tape should be produced when all voting terminals are closed for voting on election night.

2.2 Description of logs

We used three iVotronic system logs to perform the analyses described in the next section. The event log (EL152.lst), ballot image file (EL155.lst) and the ES&S election reporting manager system log (EL168a.lst). The header of the log files identify the County's name, the type and date of the election, the date the report was generated and the election ID. The election ID is a parameter generated by the ES&S election programming software to uniquely identify each election.

The event log (152.lst) lists all iVotronic terminals used on the election. The log records the terminal configuration at headquarters prior to precinct deployment which begins with the "clear and test"of the terminal to delete previous election data from the terminal's memory. The log also records, in chronological order, all relevant election day events including polls open and polls closing and the number of ballots cast. The event log contains several columns which include: iVotronic's terminal serial number, PEB serial number, PEB type, date, time, event code and event description. An excerpt of an event log is given in Appendix A.

The ballot image file (155.lst) contains all ballot images saved by the iVotronic terminals during the voting process. The ballot images are segregated by precinct and terminal where the votes were cast. The ballots are saved in a random order to protect the privacy of the voter. An asterisk (*) indicates the beginning of each ballot. An excerpt of a ballot image file is given in Appendix B.

The system log file (EL168a.lst) tracks activity in the election reporting database since its creation at the election headquarters. Its chronologically ordered entries reflect the commands executed by the operators during pre-election testing, election night reporting and post-election canvassing. This log contains the totals accumulated in the various precincts during election night reporting, as well as any warnings or errors reported by the reporting software system during the tabulation process. The system log also tracks the uploading of the PEBs and CF cards to the central election reporting database. Manual adjustment of precinct totals are also documented in the system log file. An excerpt of a system log file is given in Appendix C.

3 Analyses

In this section we present a description of our analyses and important findings. The iVotronic log files used for testing these analyses were downloaded from the website titled "South Carolina Voting Information"[1].

[1] www.scvotinginfo.com

3.1 Votes Possibly Not Uploaded

3.1.1 PEBs Not Uploaded

This analysis generates a list of PEBs used to collect votes on election day. It warns the user of any PEB, master or non-master, used to close terminals, which did not had their data uploaded to the election reporting system. The iVotronic files used by the analysis are: EL152 to search for terminal closing information and votes saved to each PEB, and EL68A for PEB upload details.

The South Carolina counties deploy two types of PEBs to each precinct on election day: a green band master PEB to open and close terminals and red band PEBs to activate ballots once the iVotronic terminals are opened for voting. The precinct procedures, dictate that a single PEB should be used to open and close all machines at a polling location. Failure to strictly follow this protocol led to problems identified in a recent study [5]. Similar problems were experienced in Miami-Dade County during the 2002 Primary election [10]. In that case, poll workers used two or more PEBs to open and close terminals at their precinct. However, election officials only uploaded one of these PEBs, because they were expecting pollworkers to follow procedures and close all machines with the same PEB. As a result, the votes from some machines were not collected on election night. Election officials were forced to spend several days at the warehouse collecting all PEBs used in the election, printing tapes of every PEB, and uploading the votes from the PEBs that were not transported to election headquarters on election night. This caused a significant delay in the reporting of election results.

This analysis is intended to address the procedural errors described above. The information it produces includes: the serial number of the terminals collected in the PEBs, the number of votes contained in the PEBs and the precinct's name and number. With the information the election officials can located the missing PEBs and add those votes to the aggregated count resulting in accuracy of certified totals and voter confidence.

Table 1 summarizes the PEBs not uploaded during the General 2010 elections in South Carolina. The system file EL168a was used to identify which PEBs containing votes were not uploaded to the election reporting software. If the South Carolina counties had access to our tool during their canvass audits they could have quickly located the PEBs.

The following are some recommendations for system improvement that would make this type of analysis easier in the future. It would be useful if the PEBs used to close terminal(s) can upload not only the total votes collected but also the serial number of the terminals it closed. Additionally, it should be possible to import a text file containing the list of iVotronic machines and master PEBs deployed to each polling location. That list could produce a crosscheck table for verification of iVotronics and PEBs uploaded during election night reporting.

3.1.2 Machines Not Closed

One of the most important aspects of any election audit is ensuring that all votes are counted. There are two main pieces of the election system that need to be analyzed to determine if votes were left out of the count. We have discussed the first essential piece: the PEB. The second piece of the election system that must be taken into account is the voting machine itself. If a machine is not closed, then a PEB has not collected this terminal's data. By checking the event log, we can determine which machines had not been closed; our tool will display the precinct name and number, and the machine serial number that was not closed.

This analysis is important because for two reasons: it helps detect the cases where some votes were not counted and gives officials enough information to collect those votes; and also highlights cases of incomplete audit data. Our tool will report the precinct name and number and the machine serial number.

While these analyses will aid officials in finding votes that may have been lost, they cannot guarantee to find all uncounted votes. Depending on the circumstances, some votes may remain uncounted because the event logs are not suitable to other auditing techniques. In order to account for more missing votes, a list of machines used in each precinct would be extremely helpful. With the files used by most of our analyses, we assume audit data is complete; by conducting a different analysis, we know that this is not true. When the audit logs are incomplete we cannot account for every machine, thus we cannot ensure all votes are counted. If there was a master list of machines used in each precinct, then there would no longer be a problem of keeping track of all machines.

The tool we implemented has reported a few instances of machines not being closed. There was a single machine that wasn't closed in each of the following counties: Greenville County, Horry County, and Sumter County. These are only detected if they are closed at some point before uploading the audit data to the database in order to print the event log and ballot images. There may be other cases that are undetectable by our analysis because a machine's audit data will not be in the event log or ballot images file if it has not been closed in the normal circumstances.

3.2 Incomplete Audit Data

In order to conduct an accurate audit of an election, the audit data must include everything that was recorded on

County	PEBs used to collect votes	PEBs not uploaded	PEBs combines votes
Anderson	77	1	163
Colleton	36	1	122
Georgetown	36	1	92
Greenville	154	3	500
Horry	121	2	189
Richland	128	5	648
Sumter	60	2	368

Table 1: PEBs not uploaded

all voting machines during the election. One of our analyses checks both the event log and ballot images to see if the appropriate machines are present. Not only is it important that there are ballot images for every machine with votes cast on it, but we need to verify that there is an equal number of vote cast events and ballot images per machine. When a machine occurs in the event log with recorded votes cast on it during election day, but does not appear in the ballot images file, then there must be data missing from the ballot images file. The reverse situation reveals the opposite error- the event log is not complete.

The importance of complete audit data lies in the accuracy of auditing elections. In South Carolina, one of the few components of the election paper trail are the audit logs. While we can sometimes identify missing information from the event log and the ballot images file, we assume the data is complete when conducting our other analyses. If a county supplies an incomplete log, our tool's results will be less accurate than they would be otherwise. This means anomalies may go undetected; missing votes may not be found and officials may not be able to identify why errors occurred during the election.

There are cases where it may be nearly impossible to tell if the data is incomplete or not. For example, if a machine was opened on election day, experienced severe problems, had no votes cast on it, and was not included in the event log, it would be undetectable unless the voting system was altered; the creation of a list of machines used at each precinct would benefit this analysis. In the case that the files are incomplete, we assume it to be a result of uploading vote data into the database at different times. Because there are two different databases (one for the event log and one for the ballot images), the system allows for new data to be uploaded between the creation of the two files. If there were one database used for both logs, this would reduce the problem. On the other hand, if there were just one database, it would be nearly impossible to detect incomplete data.

There were a number of counties' audit logs from the South Carolina 2010 elections that showed incomplete data. Our analysis detected six counties that did not have the same set of machines in both the event log and ballot images file. Florence County had the most inconsistencies with 65 machines that had votes cast on them according to the event log, but no ballot images. We also saw cases where there were ballot images for votes cast on machines that did not record any events on the event log. We also found a couple of very odd situations, such as in Sumter County, where there were two machines that were detected; one of these machines was in the event log, but not in the ballot images file, and the other machine was in the ballot images, but not in the event log. In addition to an unusually large amount of missing data, the analysis of Florence county showed machines in both files that did not have the same number of votes cast as ballot images. If election officials find this error when running an analysis, they should re-upload the data to ensure a set of complete files.

3.3 Polling Location Related Analyses

3.3.1 Polling Locations That Closed Late

The polling locations in South Carolina must be opened for voting from 7 am until 7 pm. However, the electors waiting in line after 7 pm should be allowed to vote [1]. Therefore, polling locations may stay open late in order to accommodate those voters prior to closing the voting machines on election night. If election officials knew which places were likely to experience long lines, they could deploy more equipment or personnel to those polling locations. Our tool can assist them by providing information about long lines that occur in this election. Officials can use this information to make predictions about where long lines might occur in future elections.

This analysis gives election officials information about how many polling locations had to stay open late and for how long. It uses two iVotronic log files, EL152 and EL155. It also uses the iVotronic time/date verification function described in section 3.6 to exclude any terminals whose time stamp is probably inaccurate.

This analysis generates a histogram detailing the number of polling locations that stayed open after 7 pm

# Precint	Possibly not long lines experienced (p-value <10%)	Possibly long lines experienced (p-value >10%)
26 Huger	9:00 a.m. - 10:00 a.m. 12:00 p.m. - 1:00 p.m. 4:00 p.m. - 5:00 p.m. 6:00 p.m. - 7:00 p.m.	7:00 a.m. - 9:00 a.m. 10:00 a.m. - 12:00 m. 1:00 p.m. - 4:00 p.m. 5:00 p.m. - 6:00 p.m.
10 Cordesville	7:00 a.m. - 8:00 a.m. 12:00 m. - 1:00 p.m.	8:00 a.m. - 12:00 m. 1:00 p.m. - 7:00 p.m.
24 Hilton Cross Rd	12:00 m. - 1:00 p.m. 5:00 p.m. - 6:00 p.m.	7:00 a.m. - 12:00 m. 1:00 p.m. - 7:00 p.m.
22 Hanahan 3	7:00 a.m. - 5:00 p.m. 6:00 p.m. - 7:00 p.m.	5:00 p.m. - 6:00 p.m.
20 Hanahan 1	12:00 m. - 1:00 p.m. 2:00 p.m. - 3:00 p.m. 5:00 p.m. - 6:00 p.m.	7:00 a.m. - 12:00 m. 1:00 p.m. - 2:00 p.m. 3:00 p.m. - 5:00 p.m. 6:00 p.m. - 7:00 p.m.

Table 2: Long Lines in Berkeley County

grouped by 10 minute increments. This information can be useful to election officials as they can identify how many polling locations that closed late and develop a strategy to mitigate those circumstances in future elections. Assigning additional resources to those polling locations can address bottlenecks at the precinct resulting in a more expeditious voting process during in future elections.

3.3.2 Long Lines

Election officials assign voting machines and polling location supplies based on the number of voters registered in each precinct. However, voter turnout can vary and as a result, some polling locations may end up overstocked with equipment, supplies or poll workers while others may lack resources or personnel on election day. Monitoring all the polling places in a large county can be a daunting task. Often, election officials don't have any process in place to monitor polling location usage. South Carolina counties have experienced voting machine bottlenecks during the 2008 and 2010 elections [9, 12, 2]. Those counties can benefit from a tool that can analyze DRE audit data to identify peak times at the precincts. This analysis can infer a steady flow of voters from two iVotronic log files (EL152 and EL155) and produce a report detailing possibly busy timeframes. Such information will assist election officials with the planning of future elections.

Our tool focuses on lines of voters by detecting heavily used voting terminals. When there are consecutive ballots cast with no time delay in between, we are able to infer that there is a line of voters at the voting machines.

Once our tool groups the iVotronic units by polling location, based on the information contained in the ballot images report (EL155), it determines that all the machines in the voting location were in use. This analysis also uses a function that finds polling locations which closed late in addition to a time/date verification function that invalidates and excludes voting machines with anomalous date/time settings as described in section 3.6. The date and time of iVotronic terminals is set manually and subject to human error; therefore creating inaccuracy in the timing of the events identified by the audit log. The time verification function is significant in determining which machines were heavily used at the precinct in a specific time period; therefore inferring the possibility of long lines at the precinct.

To infer long lines, we focused on the polling locations that stayed opened after 7 P.M. as we could conclude they were busy processing the voters standing in line at that time. Our analysis calculates the time between consecutive votes before 7 P.M.; we keep track of the time that these votes occurred and the time difference between votes. For all of the consecutive votes after 7 P.M., we only store the time difference between votes. This data is found per machine, which allows us to match it to its respective polling location. Then, we organize the time differences into one-hour time windows starting at 7 A.M. until 7 P.M.; all of the after-7 P.M. data was grouped together. Then, focusing on the polling locations that close very late, we use the two sample Kolmogorov-Smirnov test to determine whether the votes cast in a particular time window come from the same distribution as the after-7 P.M. votes. The result of the statistical test returns two values; one of them is the p-value. A p-value less than 10%, indicates the two

samples are unlikely come from the same distribution, and therefore there probably weren't long lines. Otherwise, a p-value higher than 10% is consistent with the two samples coming from the same distribution, however, we can not make any concrete conclusions based on a high p-value, we can only note that it is possible there were long lines in that polling location.

Table 2 summarizes the times periods when the Berkeley County polling locations experienced long lines before 7 p.m. We found that 17 precincts were closed after 7:30 P.M. and decided to run the analysis in these locations to determine when there were long lines. We will emphasize the top five precincts that close very late. Our analysis reveals that the first precinct Huger #26, which closed at 8:43:44 P.M., has higher p-values at 7:00 A.M., 8:00 A.M., 10:00 A.M., 11:00 A.M., 1:00 P.M., 2:00 P.M., 3:00 P.M. and 5:00 P.M.. The minimum p-value is 15.4% at 2:00 P.M. and the maximum is 66.0% at 11:00 A.M. Precinct Cordesville #10 could possibly have had long lines throughout almost the entire day, except for 7:00 A.M. and 12:00 P.M. with a minimum p-value of 13.6% at 10:00 A.M. and the rest of the time had p-values higher than 20%. Precincts Hilton Cross Rd #24, and Hanahan 1 #20 also had long lines almost all day. Precinct Hanahan 3 #22 only experienced long lines at 5:00 P.M. with a p-value of 38.5%. We can conclude that these precincts have experienced long lines during the whole day and it may be the reason for which those precincts closed very late.

We strongly recommend election officials to conduct this analysis after each election to plan for future elections of the same type. Assigning additional personnel, whether poll workers or rovers, and machines to the busy polling locations may reduce long lines of voters. Our analysis can detect when there is a steady flow of voters, but it does not determine if the long line of voters is caused by a slow registration process or too few voting machines.

3.4 Hardware Issues

Election officials may be interested in identifying machines that have hardware problems, such as screen calibration issues, machines with a low battery, terminals that closed early, and machines that recorded unknown, but possibly severe events. The first of these analyses detects machines with recurring calibration errors and machines that had recorded votes while possibly not calibrated. By finding the events that correspond to a screen that is not calibrated and to the recalibration of that screen, we can find if votes were cast in between those times. The second analysis regarding hardware issues looks for machines with an unusually large number of events titled "Terminal shutdown - IPS Exit."We infer that these machines have a low battery because they experience a more-than-normal number of events related to the Internal Power Supply. Additionally, our tool searches for machines that recorded a warning event about the terminal closing early. In order for this event to occur, a trained technician must enter a password to access the service menu and make a particular selection to close the machine. If a machine is closed in this manner during election day, there must be something wrong with it that is preventing votes from being cast correctly. Lastly, there is a set of events that have questionable meanings, but could potentially represent hardware issues.

Analyses such as these can help officials identify machines that may require maintenance or need to be replaced. In the case of a machine having ballots cast when it is not calibrated, it may not have correctly captured the voter's intent. Depending on the magnitude of the situation, this could cause a different outcome in the election. If our analyses detect other hardware problems with a machine, it may not be recording votes accurately; these votes may not even appear in the event log or ballot images. If the event log and the ballot images do not record ballots being cast, then it is nearly impossible for officials to realize votes are not being counted.

Due to the available resources and the nature of these analyses, we made assumptions regarding the meaning of events and the severity of the situation. Currently, there is no user manual or detailed description of the events that appear in the event log; because of this, we are not able to guarantee that the event "Terminal shutdown - IPS Exit"means the machine has a low battery. This assumption is also applicable to our calibration analysis and our detection of unknown warning events. If a description of each event was available, we could be more definitive in our results and possibly implement analyses that report other useful hardware failures.

The machines used in South Carolina have experienced frequent potential hardware issues. For example, the combination of machines in Berkeley County experienced votes cast on a machine when the machine was not calibrated, machines with possible low batteries, and at least one machine that closed early. Our analysis found that there were seven counties where at least one machine was possibly not calibrated when votes were cast on that machine; these errors spanned 12 different polling locations. We suggest an election official or technician inspect these machines for possible calibration issues. We had similar findings when searching for terminals that recorded a "Warning - Terminal Closed Early" event. There were machines with this warning in seven counties and 13 polling locations. Terminals should not close before 7 P.M. in South Carolina on election day; for this reason, we recommend that these machines be evaluated

for potential problems that would have caused early closure. When our tool reports machines with possible low batteries, election officials should verify that the machine is working properly and does not need maintenance. Florence County and Greenville county experienced a number of Internal Power Supply - related events; at least one machine in each precinct had 53 and 63 instances of this event, respectively. This could be a possible indicator that the battery is running low; therefore, the election officials should take action to ensure all machines work properly in future elections.

3.5 Procedural Errors

Our tool can detect procedural errors and poll worker mistakes. A few of these are: precincts that do not print zero tapes on the morning of election day; using a master PEB to activate ballots; opening and closing machines with different PEBs. According to the South Carolina poll worker training video (citation), poll workers are required to print at least one zero tape per polling location on the morning of the election. Using the event log, our tool checks each polling location for this event and reports the locations that did not record this event. Another way our tool finds procedural errors is by crosschecking the master PEBs with the PEBs used to activate ballots. Poll workers should be using non-master PEBs to activate ballots so that the PEBs do not get switched. Along the same lines, we report incidents of opening and closing a machine with different PEBs. A machine should be opened and closed with the same master PEB; if not, it may be more likely that this PEB does not get uploaded.

When poll workers cancel ballots, they must select a reason why; this is another way to detect errors. There are seven options for canceling a ballot: wrong ballot, voter left after the ballot was issued, voter left before the ballot was issued, voter request, printer problem, terminal problem, or an unspecified reason. If there are any instances of canceling a ballot due to a printer problem, it could be an indicator of a procedural error because ballots are not printed. In other cases, if there is a large number of a specific reason, such as having the wrong ballot, this could indicate the poll workers are repeatedly issuing the wrong ballot.

It may be beneficial to election officials if they could detect at which locations poll workers are following the required procedures. Procedural errors can cause many problems including lost votes, incorrect vote counts, disgruntled voters, and long lines. If election officials are aware of the procedures that are not being followed, they could review their precinct checklist. This will allow for more efficient audits as well as a better voting experience for voters.

While our analyses detect an important set of errors, there are certainly many more procedures that can be analyzed. In addition to printing zero tapes in the morning, poll workers are required to print results tapes at the end of the election; unfortunately, this is not detectable due to the way the event log is produced. We have inferred from the event logs that the poll workers are extracting the compact flashes before printing the results tapes, therefore the event log shows no record of the event.

Our findings reveal the improvements needed for poll worker training and for various procedures. Although we found that pollworkers were following procedures concerning the printing of zero tapes, there were a number of counties with procedural errors. When opening and closing a machine, the same master PEB should be used, but in 11 counties there were cases of opening and closing machines with different PEBs. Our results showed a correlation between this error and certain precincts, where pollworkers made those mistakes repeatedly. Colleton County had five instances of this procedural error, but four of those instances took place at one polling location; Walterboro No 4 had machines 5129946, 5133679, 5138439, and 5138563 opened with PEB 155914, but closed with PEB 155925. This should raise a red flag to the election officials that they may need to emphasize this procedure in poll worker training. When poll workers activate ballots for voters, they should do so with a non-master PEB; we saw two counties that had an unusually high number of violations of this procedure. Horry County and Richland County had 22 and 32 instances of this violation, respectively. When election officials see this result, they may wish to revise poll worker training. Our tool also analyzes the reasons why votes were canceled, which could give insight to procedural errors. There is likely to be a certain number of vote cancellations due to a number of reasons, but our tool will only report the machines that recorded an abnormally large amount of vote cancellations for a specific reason. Colleton County had a machine that recorded 12 instances of vote cancellations due to a terminal problem; in this case, we would recommend the officials to inspect the machine for potential hardware problems. A machine in Lexington County experienced an unusually large number of vote cancellations due to a ”wrong ballot”; this could be result of many problems. The machine may have a calibration issue, or there may be a procedural error in that the poll workers are repeatedly selecting the wrong ballot.

3.6 Systematic Date and Time Errors

The iVotronic DREs append each audit event to the log in chronological order. Each event is marked with a timestamp based on the DRE's internal clock. We discovered and report on a variety of errors related to these time

stamps. Correct time-stamps are critical in post election audits and often incorrect stamps can not be automatically corrected post election. Previous work identifies and remarks on some of these date errors [5, 11]. We further attempt to classify and automate the identification of these issues.

Erroneous time-stamps can invalidate the audit-logs and often preclude data from being used in automated analyses. Determining machines to have either valid or invalid time-stamps has a lot of gray area and different errors will affect different analyses. For example, some machines experienced time-stamps that would blank to '00/00/00 00:00:00' for only a couple of events. This wouldn't affect data looking at opening and closing times, but would create outliers or gaps in data measuring time between votes cast.

We found it simplest to classify date errors into two categories: those errors resulting from machines not having their clocks set appropriately; and those resulting from apparent bugs in the iVotronic time-stamp mechanism itself. Our website includes an automated a report that attempts to identify and group as many of these errors as possible.

First, machines shipped or set with incorrect dates were identified. These errors all suggest the need for a more thorough pre-election check of each machine's clock. Any machines experiencing manual clock adjustments on election day were included in the report as well as those machines which opened for voting on a date that was wildly incorrect (i.e. dates well before pre-election or dates after election day). This mostly included machines that did not account for Daylight Saving Time or those machines that didn't set their initial date until after opening for voting. All the above machines were checked as closing on a valid election day. Early voting dates were not considered as they appeared to be inconsistent among the different counties. Machines which open and close on an improbable dates are separately identified as machines that had bad dates that went uncorrected..

Machines experiencing date errors were classified in a separate report. Many machines were found to have anomalous date changes that weren't paired with the normal date set event. [Figure 1 will show the 4/12 jump in Berkeley county] Often before the clock on a machine is first set the dates will show up as being many years into the future or as a zero date. This isn't a problem as most of these machines are set correctly before being opened for voting. However, there are many places in the logs were the date will seemingly randomly jump to a date far into the future or the past and remain there until manually corrected. [Figure 2 1221 in Berk County] shows a case where the date jumps ahead to 12/21/2010 for two events before changing back. These machines were automatically identified by looking for any major date jumps that occur on election day or zero stamps being recorded after machines are open. Machines experiencing many date jumps may require troubleshooting from ES&S. It may be a more systematic issue that the date on an iVotronic can apparently change for no reason.

We strongly advise that procedures for setting the clocks on machines are reviewed. The unknown date jumps seen in the logs are concerning, but generally are not creating as many audit issues compared to machines whose date were never set correctly.

Accurate clocks directly impact the usefulness and correctness of the audit logs. Ensuring that every single machine is set correctly is not necessarily a simple task. We would recommend that each machine is configured accurately before being sent to the precincts. Additionally, all machines should be double checked for a correct time before opening for voting. Daylight Saving Time settings are also a potential concern. Many machines in Anderson County were found to be adjusted forward by an hour during election day.

The authors of "Casting Votes in the Auditorium" [11] propose a distributed network between DREs. This 'Auditorium'provides a far more robust system to ensure accurate and verifiable audit logs. They propose a system where all the election machines are networked together and append to a common audit log verified by each machine. This allows for more error redundancy and removes the logistical issue involved in making sure every single machine has their date correctly set.

4 Related Work

Many election technology systems provide a possible means of auditing elections. For example, in optical scanning systems the cast ballots themselves form a paper record of the votes cast on election day. On the other hand, DRE machines do not provide this type of paper trail. Some DREs provide a Voter Verified Paper Audit Trail (VVPAT), which stores a hard copy version of each ballot cast. The paper ballot can be read, but not modified, by the voter at the time of casting their vote. A third type of audit trail, which is produced by all DREs, are the event logs stored electronically on each DRE. Our work pertains to the elections in South Carolina, which do not require the creation of a paper trail, but do provide the audit logs from the machines. In this section we discuss related work on the analysis of audit logs for post-election auditing.

Two recent studies, which analyzed the iVotronic audit logs, focused on the verification of election results [5, 11]. The authors of the first study [5] performed an audit of the same South Carolina elections that we analyzed. Using these audit logs, they discovered un-

counted votes and problems with the audit data. By consulting additional audit materials, such as the printed results tapes, the authors were able to offer possible reasons and explanations as to why the problems occurred. Our work takes a slightly different approach. We focus on providing a fully automated web-based tool that election officials might use and therefore we do not refer to printed results tapes or other material aside from the audit logs in our analysis. While our tool did discover and report these same problems, we simply report what was wrong, but can not provide a possible explanation for the cause of the error.

The authors of the second study [11] provided an analysis of vote tallies using the protected count of votes on each machine and comparing this to the printed results tapes. Their report also finds dates that were most likely inaccurate. With further investigation, they concluded that the hardware clock was incorrect. Our research provides analyses to identify similar problems, but in a way that could be automated.

There has also been research on using the audit logs to analyze election-day procedure and activity. For example, one recent publication showed how event logs could be used to determine if a machine acted "normally"on election day [4]. The authors of this research studied the event logs of the AccuVote Optical Scanning system and used those logs to build a finite state machine that models the sequences of events a well-behaved machine might produce. This type of analysis would be useful to provide for the iVotronic systems that we studied. However, the AccuVote machines have considerably fewer possible event types than the iVotronics so the analysis would become considerably more complex.

A common problem on election day, which we try to identify in our analysis, is the occurrence of long lines. Many studies have researched ways to mitigate long lines at polling locations [3, 6, 13, 15]. One such study has simulated the flow of voters through the voting process at polling location [7]. The authors use this simulation to determine the optimal number of voters per voting machine, and correspondingly, the correct number of voting machines per polling location based on the number of registered voters at that particular location. Their work is predictive: the authors make some assumptions, such as the average time it takes to vote and when peak voting hours will occur, and use those as a basis for predicting where long lines are likely to occur. Our analysis is descriptive: given the audit logs from election day, we infer the average time it took to vote and use that information to determine whether a particular polling location experienced long lines or not. The two methods are complementary. Predictive models can be used to prevent long lines, while descriptive models can be used to check and refine the prediction algorithms.

Voter Verified Paper Audit Trails (VVPATs) are a different type of audit log. Unlike the audit logs we used in our analyses, VVPATs are viewed and verified by the voter and are more suited to audits concerning a DRE incorrectly capturing a voter's intent. Our work is more concerned with identifying cases of cast votes not being included in the final count, or issues at the polling place that might prevent the voter from casting their vote in the first place. With VVPATs, as long as a certain percentage of voters do check their paper ballot [8], the voting machine need not be assumed correct, whereas our analyses do make this assumption.

5 Future Voting System Suggestions

[To complete by Wednesday].

6 Conclusion

In this study we developed a tool to analyze audit data from DRE voting machines. Our web based application, accessible to anyone, performs a variety of analyses on the audit data to detect procedural errors and system deficiencies. In addition, our tool can identify terminals that were not closed and their votes not uploaded to the cumulative count. This information can be very useful during the canvass process as election officials can locate the missing terminals, close them and add their votes to the election totals.

Having performed analyses with the iVotronic logs from South Carolina, we also report statistics on polling location procedures. These statistics include: polling locations that closed late or may have experienced long lines of voters, precincts which did not report the zero tape and polling locations which used the wrong device to activate ballots on election day. Our tool can also report statistics concerning possible DRE hardware problems such as calibration issues, low battery and incorrect date and time settings. With this information election officials can improve their poll worker training or schedule voting machine repairs as needed.

Dr. Wagner's commissioned report , "Voting Systems Audit Log Study" extensively documents and evaluates many different types of audit logs produced by six different voting systems. In the findings there were no machines that provided tools, support, or generated summary reports for analyzing audit logs. [14]. While, the authors are very familiar with the strengths and weaknesses of the iVotronic's audit logs we would direct anyone interested in the future design of audit logs to this report. Fully documenting the strengths and weaknesses of the iVotronic audit systems is outside the scope of this project. Our website is only the first step in creating a

process for automated election auditing. We hope that future third-party audit log tools can build on some of our work to create a useful and robust solution for deriving meaningful audits directly from the logs.

We recommend that election administrators conduct routine reviews of the audit logs generated by the voting machines as they are ground truth for election disputes. By automating our analyses and making it as simple as uploading the iVotronic audit logs to a website, we believe our tool can standardize the post-election audits performed by the iVotronic system users. Our website can quickly provide intelligent feedback to election officials during the canvassing process and serve to influence future audit procedures.

7 Acknowledgments

Special thanks to Dr. Kristen Gates, National Science Foundation and the TRUST program staff.

References

[1] South Carolina Voting Information Page. `http://www.scvotes.org/south_carolina_voting_information_page/`.

[2] Heavy voter turnout reported statewide, few problems reported. `http://www.wistv.com/Global/story.asp?S=13428174/`, November 10, 2010.

[3] T. Allen and M. Bernshteyn. Mitigating Voter Waiting Times. *Chance Magazine*, 19(4):25–36, 2006.

[4] T. Antonyan, S. Davtyan, S. Kentros, A. Kiayias, L. Michel, N. Nicolaou, A. Russell, and A. Shvartsman. Automating voting terminal event log analysis. In *Proceedings of EVT 2009*, August 2009.

[5] D. A. Buell, E. Hare, F. Heindel, C. Moore, and B. Zia. Auditing a DRE-Based Election in South Carolina. In *Proceedings of EVT 2011*, August 2011.

[6] K. Dow. Study of Voter Flow at the 2006 General Election, Columbia County, NY. `http://tinyurl.com/6dozpq/`, 2007.

[7] W. A. Edelstein and A. D. Edelstein. Queuing and Elections: Long Lines, DREs and Paper Ballots. In *Proceedings of EVT 2010*, August 2010.

[8] J. L. Hall. Design and the Support of Transparency in VVPAT Systems in the US Voting Systems Market. `http://vote.nist.gov/jlh-vvpat-design-transparency.pdf/`, 2006.

[9] A. Kreitman. Precincts across SC report heavy voter turnout. `http://www.live5news.com/story/13428203/live-updates-election-2010/`, November 02, 2010.

[10] C. Mazella. Inquiry into circumstances surrounding the September 10, 2002 election in Miami-Dade County. Memorandum to Honorable Alex Penelas, Mayor, September 10, 2002.

[11] D. Sandler and D. S. Wallach. Casting Votes in the Auditorium. In *Proceedings of EVT 2007*, August 2007.

[12] D. Slade and G. Smith. Long lines, patient voters. `http://www.postandcourier.com/news/2008/nov/04/voters_speak60331/`, November 4, 2008.

[13] D. M. Spencer and Z. S. Markovitz. Long Lines at Polling Stations? Observations from an Election Day Field Study. *Election Law Journal*, 9(1):3–17, 2010.

[14] D. Wagner. Voting systems audit log study. Report commissioned by the California Secretary of State, June 1, 2010.

[15] R. Wilson. 2004 and 2006 Long Voter Lines in Prince Georges County. `http://tinyurl.com/46m2ng/`, September 10, 2008.

Appendix A: Event Log File

Votronic	PEB#	Type	Date	Time	Event
5124751	151386	SUP	11/02/2010	08:25:24	0001510 Vote cast by voter
		SUP	11/02/2010	08:30:34	0001510 Vote cast by voter
	137507	SUP	11/02/2010	08:36:59	0001510 Vote cast by voter
	151386	SUP	11/02/2010	08:42:37	0001510 Vote cast by voter
		SUP	11/02/2010	08:47:07	0001510 Vote cast by voter
		SUP	11/02/2010	08:52:06	0001510 Vote cast by voter
		SUP	11/02/2010	08:55:06	0001510 Vote cast by voter
	137507	SUP	11/02/2010	09:01:30	0001510 Vote cast by voter
	151386	SUP	11/02/2010	09:05:33	0001510 Vote cast by voter
	137507	SUP	11/02/2010	09:12:19	0001510 Vote cast by voter
		SUP	11/02/2010	09:16:40	0001510 Vote cast by voter
		SUP	11/02/2010	09:18:52	0001510 Vote cast by voter
	155466	SUP	11/02/2010	09:19:52	0001721 PEB pulled while getting PEB type
		SUP	11/02/2010	09:19:52	0002405 Failed to get PEB type
		SUP	11/02/2010	09:19:52	0002400 PEB access failed
	137507	SUP	11/02/2010	09:21:20	0001515 Vote canceled - voter left before ballot
		SUP	11/02/2010	09:24:17	0001516 Vote cancelled - voter request
		SUP	11/02/2010	09:29:25	0001510 Vote cast by voter
		SUP	11/02/2010	09:32:33	0001516 Vote cancelled - voter request
		SUP	11/02/2010	09:36:18	0001510 Vote cast by voter
		SUP	11/02/2010	09:40:17	0001510 Vote cast by voter
		SUP	11/02/2010	09:43:32	0001510 Vote cast by voter
	155466	SUP	11/02/2010	09:45:26	0001510 Vote cast by voter
	151386	SUP	11/02/2010	09:51:45	0001510 Vote cast by voter
	137507	SUP	11/02/2010	09:56:10	0001510 Vote cast by voter
	151386	SUP	11/02/2010	10:04:22	0001510 Vote cast by voter
		SUP	11/02/2010	10:09:17	0001516 Vote cancelled - voter request
		SUP	11/02/2010	10:12:03	0001510 Vote cast by voter
	155466	SUP	11/02/2010	10:13:32	0001516 Vote cancelled - voter request
		SUP	11/02/2010	10:17:36	0001510 Vote cast by voter
	137507	SUP	11/02/2010	10:21:52	0001511 Vote cast by poll worker
		SUP	11/02/2010	10:27:24	0001510 Vote cast by voter
		SUP	11/02/2010	10:30:35	0001510 Vote cast by voter
		SUP	11/02/2010	10:37:13	0001510 Vote cast by voter
		SUP	11/02/2010	10:40:29	0001510 Vote cast by voter
		SUP	11/02/2010	10:45:26	0001510 Vote cast by voter
		SUP	11/02/2010	10:50:07	0001510 Vote cast by voter
		SUP	11/02/2010	10:52:48	0001510 Vote cast by voter
	151386	SUP	11/02/2010	10:58:06	0001510 Vote cast by voter
		SUP	11/02/2010	11:04:08	0001510 Vote cast by voter
	137507	SUP	11/02/2010	11:14:11	0001510 Vote cast by voter

Appendix B: Ballot Image File

RUN DATE:03/30/11 08:05 AM PRECINCT 57 - Liberty Hall ELECTION ID: 08110210

VOTR.	B/I	CANDIDATES RECEIVING A VOTE	
5129343	2	46 Tom E Elliott	Commissioner of Agriculture
5129343	2	49 Tom Clements	U.S. Senate
5129343	2	57 Ben Frasier	CONG001 U.S. House of Rep. Dist 1
5129343	2	66 Bill Crosby	HOU0117 State House of Rep Dist 117
5129343	2	72 Wayne Dewitt	Sheriff
5129343	2	75 Keith Kornahrens	Probate Judge
5129343	2	78 Mary P Brown	Clerk of Court
5129343	2	88 Diane Edwins	Soil and Water District Commission
5129343	2	92 Yes	Amendment 1
5129343	2	95 Yes	Amendment 2
5129343	2	98 Yes	Amendment 3
5129343	2	101 Yes	Amendment 4
5129343	2	104 Yes	Local Question
5129343	2 *	11 Vincent A Sheheen	Governor
5129343	2	19 Mark Hammond	Secretary of State
5129343	2	24 W/I MICKEY MOUSE	State Treasurer
5129343	2	28 Matthew Richardson	Attorney General
5129343	2	32 Robert Barber	Comptroller General
5129343	2	38 Frank Holleman	State Superintendent of Education
5129343	2	43 W/I MARTIN SHEEN	Adjutant General
5129343	2	46 Tom E Elliott	Commissioner of Agriculture
5129343	2	51 Alvin M Greene	U.S. Senate
5129343	2	57 Ben Frasier	CONG001 U.S. House of Rep. Dist 1
5129343	2	92 Yes	Amendment 1
5129343	2	95 Yes	Amendment 2
5129343	2	99 No	Amendment 3
5129343	2	101 Yes	Amendment 4
5129343	2	104 Yes	Local Question
5129343	2 *	4 Democratic	STRAIGHT PARTY
5129343	2	11 Vincent A Sheheen	Governor
5129343	2	16 Ashley Cooper	Lieutenant Governor
5129343	2	20 Marjorie L Johnson	Secretary of State
5129343	2	23 Curtis Loftis	State Treasurer
5129343	2	28 Matthew Richardson	Attorney General
5129343	2	32 Robert Barber	Comptroller General
5129343	2	38 Frank Holleman	State Superintendent of Education
5129343	2	42 Bob Livingston	Adjutant General
5129343	2	46 Tom E Elliott	Commissioner of Agriculture
5129343	2	51 Alvin M Greene	U.S. Senate

Appendix C: System Log File

```
11-02  09:03 pm  START PACK ACCUMULATION  (Replace Mode - restarting)
11-02  09:03 pm  0023-Time stamp mismatch (Reply was: Update)
11-02  09:03 pm                                          PRC 0023 PACK RECEIVED VTR  (BALS=728  TOT=728)
11-02  09:03 pm  0040-Time stamp mismatch (Reply was: Update)
11-02  09:03 pm                                          PRC 0040 PACK RECEIVED VTR  (BALS=816  TOT=816)
11-02  09:03 pm  0049-Time stamp mismatch (Reply was: Update)
11-02  09:03 pm                                          PRC 0049 PACK RECEIVED VTR  (BALS=1093  TOT=1093)
11-02  09:03 pm  0043-Time stamp mismatch (Reply was: Update)
11-02  09:03 pm                                          PRC 0043 PACK RECEIVED VTR  (BALS=604  TOT=604)
11-02  09:03 pm  0044-Time stamp mismatch (Reply was: Update)
11-02  09:03 pm                                          PRC 0044 PACK RECEIVED VTR  (BALS=739  TOT=739)
11-02  09:03 pm  0044-Time stamp mismatch (Reply was: Update)
11-02  09:03 pm  0044-Precinct already updated (Pack suspended)
11-02  09:03 pm  0058-Time stamp mismatch (Reply was: Update)
11-02  09:03 pm                                          PRC 0058 PACK RECEIVED VTR  (BALS=833  TOT=833)
11-02  09:03 pm  STOP PACK ACCUMULATION
11-02  09:03 pm  START PROCESS PEBS
11-02  09:07 pm  PEB votes retrieved for P0137152
11-02  09:07 pm  SPP file record created for P0137152
11-02  09:07 pm  PEB votes retrieved for P0137152
11-02  09:07 pm  SPP file record created for P0137152
11-02  09:07 pm  PEB votes retrieved for P0147256
11-02  09:07 pm  SPP file record created for P0147256
11-02  09:08 pm  PEB votes retrieved for P0147585
11-02  09:08 pm  SPP file record created for P0147585
11-02  09:08 pm  STOP PROCESS PEBS
11-02  09:09 pm  iVotronic GROUP 3 SELECTED FOR UPDATE
                 EQUIPMENT TYPE VTR - UPDATE PRECINCTS COUNTED:Y
```