From: Commissioner Jerdonek
Date: February 8, 2023

Subject: DVSorder Confirmation and Observations

During the Elections Commission's last meeting on January 18, 2023, the discussion during the Director's Report touched briefly on a privacy flaw in the Dominion voting system used by San Francisco. The flaw was brought to the Commission's attention in a January 9, 2023 email from J. Alex Halderman to the Commission (see Attachments 1 and 2).

Since that meeting, I wanted to confirm for myself whether the flaw affects the Dominion data posted on the Department's website for the November 8, 2022 election and to see what that process involves. The purpose of this memo is to share information about my findings.

Table of Contents

1. DVSorder Background

The DVSorder privacy flaw was discovered sometime before August 23, 2022 by computer scientists J. Alex Halderman of the University of Michigan and Drew Springall of Auburn University and three University of Michigan student researchers: Braden Crimmins, Dhanya Narayanan, and Josiah Walker. After notifying Dominion Voting Systems, the U.S. Election Assistance Commission (EAC), the U.S. Cybersecurity & Infrastructure Security Agency (CISA), and the affected state election agencies (including the California Secretary of State's Office), the researchers announced DVSorder publicly on October 14, 2022 at the following website (see also Attachment 3): https://dvsorder.org/.

The privacy flaw affects ballots that are cast by inserting into a polling place precinct scanner (specifically Dominion's ImageCast Evolution). The flaw allows a member of the public to determine, from the electronic record of one of these ballots posted on the Department of Elections website, the exact time to the second that the ballot was inserted into the scanner.

As described in Attachment 2 and on the DVSorder website, here are some scenarios where someone could use this flaw to figure out how a voter at a polling place voted:

- If someone shared or posted a photo or video of themself voting (e.g. a ballot "selfie"), and that photo or video has timestamp metadata, then others could figure out how that person voted.
- If someone voted at a polling place and knows when they voted, they could figure out how the people before or after them voted (like one of their neighbors).
- If a security camera was filming the entrance to a polling place, that footage could be used to figure out how individuals voted, especially during times the polling place is less crowded.

The flaw was due to poor security coding in the Dominion voting system, namely the use of an inappropriate pseudorandom number generator (PRNG) in the code. Incidentally, this flaw is an example of the kind of thing that would be easy to spot if the voting system were open source. (San Francisco has been wanting to move towards open source since 2008.)

The existence of the flaw also shows that neither California's nor the federal government's certification process was sufficient to catch the security issue. Because the Dominion software is proprietary, or secret, these are the only public entities in a position to catch these issues. Thus, more scrutiny is needed.

2. Example: San Francisco's November 8, 2022 Election

As a software engineer, and by following the information the researchers posted publicly, I was able to confirm that DVSorder can be used to determine the time (and place) that the polling-place ballots posted on the Department of Elections' website were cast. I used Dr. Halderman's proof-of-concept code that is linked from the DVSorder website: https://github.com/research/dvsorder

Knowing this level of detail about the ballots allows certain types of research and analysis to be done that isn't normally possible after an election. For example, one can use this information to examine how voting patterns change based on the time of day the ballots were cast. To show that I was able to confirm DVSorder, I carried out a few simple examples of this kind of analysis and graphed them below.

The graphs include 24,831 voters, which is around the number of voters affected by the privacy flaw. This amounts to 69.8% of the 35,595 ballots cast in person and 8.0% of the 310,071 ballots cast in all.

2.1. Illustration of Findings

For my analysis, I focused on the 14 local measures, Measures A through O. First, I graphed for each local measure the percent of voters that undervoted (left the contest blank), rounded to the nearest hour. This is Figure 1 below. Incidentally, you can see that the percent of undervoting is correlated across measures (though one doesn't need DVSorder to confirm this).

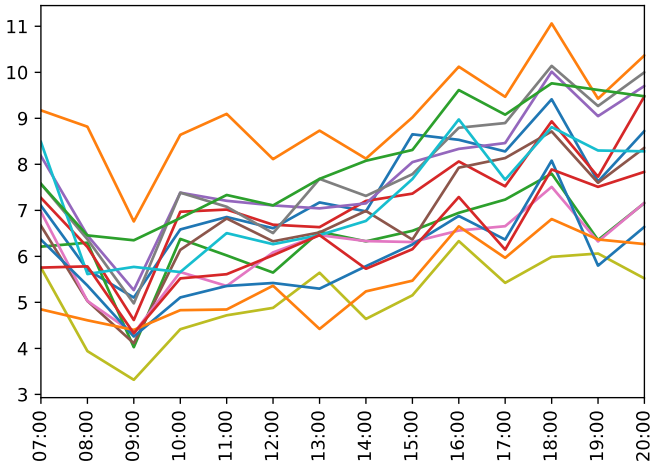

Figure 1. *Of the ballots cast and counted at a precinct polling place, the percent of undervotes on each local measure over the course of Election Day, grouped by the nearest hour.*

I also graphed the percent undervoting averaged across all 14 of these measures (Figure 2 below). Here you can see more clearly than in Figure 1 that the least amount of undervoting happens around 9am (perhaps before people are going to work), while undervoting peaks at around 6pm (e.g. when people are coming home from work).

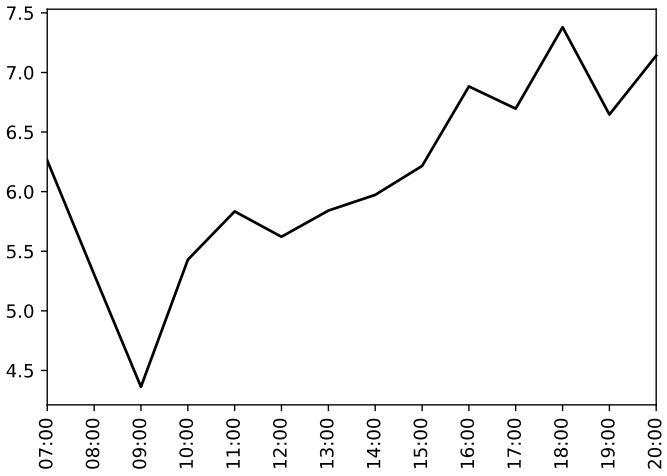

Figure 2. *Of the ballots cast and counted at a precinct polling place, the (average) percent of undervotes on the local measures over the course of Election Day, grouped by the nearest hour.*

Next, I graphed the percent of "Yes" votes on the local measures, averaged over the 14 local measures (Figure 3 below). Here you can see that the percent of voters voting "Yes" is lowest both at around 11am (an hour before lunchtime), as well as in the mid-afternoon around 2pm or 3pm. The highest percentage of "Yes" votes occur at the end of the day starting around 5pm.

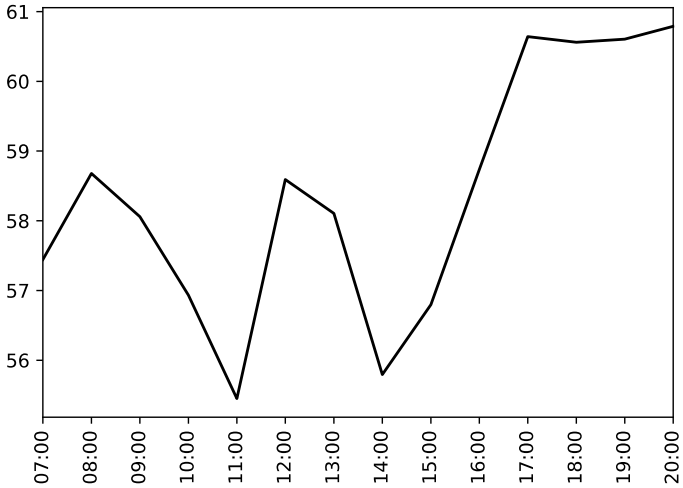

Figure 3. *Of the ballots cast and counted at a precinct polling place, the (average) percent of "Yes" votes on the local measures over the course of Election Day, grouped by the nearest hour.*

The last graph (Figure 4 below) provides more context for the graphs above. It shows the number of voters voting at a precinct polling place rounded to the nearest half-hour. For this graph, DVSorder isn't needed, as it can be constructed from the Department's scanner log files alone. This graph shows that the number of voters steadily increases throughout the day and then peaks at 6pm, before dropping off sharply (the polls close at 8pm).

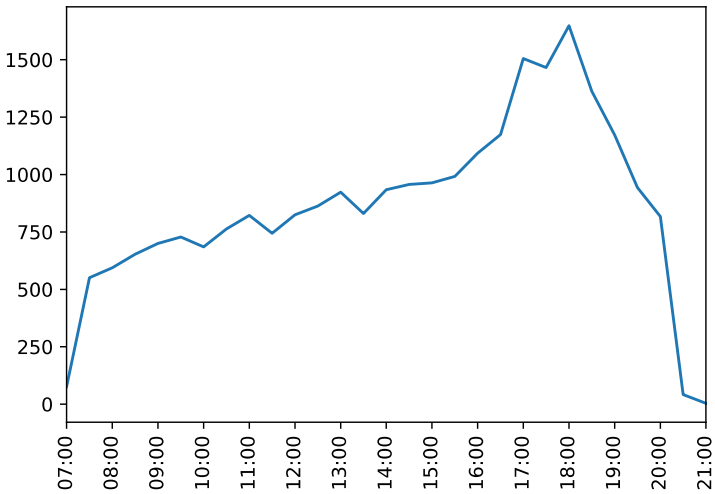

Figure 4. *The number of voters whose ballot was cast and counted at a precinct polling place over the course of Election Day, grouped by the nearest half-hour.*

2.2. Notes on Data Analysis

This section contains notes of a more technical nature. It contains various observations I made during my analysis, some of which can be viewed as questions.

2.2.1. There were 514 precincts in the PrecinctManifest.json file. 13 of these were mail precincts by the name (e.g. “PCT 9147 MB”). This left 501 non-mail precincts.

2.2.2. There were 1,049 tabulators in the TabulatorManifest.json file. Of these, 501 had a “Type” of ImagecastEvolution, which is the type corresponding to a polling-place precinct scanner. Of the 501 ImagecastEvolution scanners, four of them didn’t have CVR files from the polling place posted on the Department’s website, so I excluded these from the analysis. These were (I haven’t asked yet why these were missing)—

1. Tabulator 555 (Precinct 1108 - 218 Granada Avenue)
2. Tabulator 562 (Precinct 1117)
3. Tabulator 713 (Precinct 7601)
4. Tabulator 903 (Precinct 9143 - 651 6th Avenue)

I was later able to figure out which central scanners scanned most of these ballots.

2.2.3. In addition, 16 of the 501 scanners didn’t seem to have a log file in the “Ballot Scanning Machine Logs” download (there were 485 precinct log files), so I also excluded these scanners. (I also haven’t asked yet why these log files weren’t present.) The 16 scanners were—

1. Tabulator 569 (Precinct 1125 - 209 Ocean Ave)
2. Tabulator 570 (Precinct 1126 - 37 Rudden Ave)
3. Tabulator 574 (Precinct 1131 - 263 Maynard St)
4. Tabulator 626 (Precinct 7038 - 2111 Jennings St)
5. Tabulator 628 (Precinct 7041 - 1654 Sunnydale Ave)
6. Tabulator 694 (Precinct 7514 - 1399 Mc Allister St)
7. Tabulator 722 (Precinct 7611 - No Poll Found At This Time)
8. Tabulator 758 (Precinct 7706 - 261 Wawona St)
9. Tabulator 776 (Precinct 7806 - 25 Sanchez St)
10. Tabulator 783 (Precinct 7814 - 170 Valencia St)
11. Tabulator 796 (Precinct 7828 - 100 Hoffman Avenue)
12. Tabulator 806 (Precinct 7839 - 829 Duncan St)
13. Tabulator 820 (Precinct 7901 - 390 Valencia St)
14. Tabulator 895 (Precinct 9134 - 734 46th Ave)
15. Tabulator 898 (Precinct 9137 - 855 27th Ave)
16. Tabulator 993 (Precinct 9446 - 20 Crestlake Drive)

2.2.4. The log file for Tabulator 588 (Precinct 1146 - 494 Pope St) showed its ballots being cast between 10:30:37 PM on November 8, 2022 to 7:11:01 AM on November 9 (this is an 8 hour, 40 minute window). It looks like the clock may have been set incorrectly on this scanner, so I also excluded this scanner.

2.2.5. Excluding the 4 + 16 + 1 = 21 scanners mentioned above from the 501 ImagecastEvolution scanners left 480 scanners. These are the scanners included in my graphs in Section 2.1. This amounted to 24,831 voters, which is 69.8% of the 35,595 ballots cast in person and 8.0% of the 310,071 ballots cast in all as I said above.

2.2.6. The ballot scanner log files were encoded using latin1. This didn't seem to be documented in the Dominion PDF documenting the format of the log files. Also, curiously, all the precinct log files contained the same data repeated multiple times. Specifically, 416 of the files contained the same data repeated twice, 66 were repeated 3 times, 2 of them 4 times, and 1 of them 5 times. It took a bit of extra work to figure out that this was happening and then to divide up the data into an appropriate number of pieces, confirm the pieces were the same, and then extract just the first copy.

2.2.7. Finally, I noticed that for Tabulator 1038 (Precinct 9732 - 50 Frida Kahlo Way), one of the record ID's in the pseudo-random number generation (PRNG) sequence was skipped in the precinct's CVR file (what would have been the 235th ballot in the precinct). After thinking about this and investigating, I believe it's because the corresponding record ID would have been 0 (as it corresponded to sequence number 557837 for the Imagecast Evolution PRNG posted on the DVSorder website). Perhaps Dominion wanted to exclude this as an ID and so skipped over it.