EXAMPLE ATTACK DOCUMENTATION

Optical Scan Calibration

Douglas W. Jones
Sept 25, 2005

Taxonomy: Administrative, wholesale
Applicability: optical-scan voting systems (precinct-count and central-count)

Method:

Optical scan voting systems have a mark-sensing threshold. Marks that appear darker than this threshold, to the scanner, will be counted as votes. Marks that appear lighter than this threshold will not be counted as votes. (Some scanners can be configured to detect marks within an intermediate range as questionable.)

The threshold (or thresholds, for those scanners with an intermediate range) is generally variable. It can be set to reject all but very dark marks or it can be set to accept even the faintest of marks. Scanner calibration involves setting the thresholds of the various scanners being used so that they will, as nearly as possible, count ballots in conformance with the applicable law. Ideally, all scanners should be set so that they will apply the same standards, as nearly as possible, and so that these standards are comparable to the standards a person examining the ballot would apply in determining whether or not a mark is a vote. Overly sensitive scanners will sometimes detect overvotes as a result of counting dots, dust specks or printing defects as votes, while overly insensitive scanners will frequently fail to notice legitimate marks in the voting target, leading to undervotes.

Depending on the scanner, setting the threshold can be a matter of physical adjustment, for example, of trimmer potentiometers, or it can be a matter of setting the contents of configuration memory (possibly through a configuration file). In some cases, calibration changes can be made by substitution of different photosensors, for example, replacement of infra-red sensors with visible-light sensors or visa-versa. Scanner calibration is frequently done by vendor's representatives.

Errors in scanner calibration are probably more frequent than any deliberate manipulation of calibration. Manipulation of election results by deliberately improper scanner calibration is possible. For example, if the scanners used in precincts (or counties) that are favored by the perpetrator are calibrated reasonably, while scanners used in precincts that the perpetrator wishes to attack are set unreasonably (overly sensitive or overly insensitive), leading to a higher likelihood that ballots scanned on those machines will be scanned as containing overvotes or undervotes.

Resource requirements: The perpetrator must control the calibration of the scanners. Since calibration is typically done by the vendor's technicians, they will typically be involved.

Potential gain:

The reviews of optical mark-sense ballots cast in Florida in 2000, done by the Miami Herald group, include data showing that widely variable numbers of voters made such errors as marking an X or checkmark in the voting target instead of blacking it in. Reported percentages were as high 1 percent (Washington county) and as low as zero. The average rate, statewide, for circled voting targets, improper marks or use of the wrong type of marking implement came to about 1/2 percent. These figures, based on eyeball examination, should not be taken as better than a rough lower bound on the mismarking rate, since the methodology varied from county to county and did not necessarily involve inspecting all ballots for potential problems.

Nonetheless, it is reasonable to guess that deliberate moderate manipulation of the calibration depending on the precinct or depending on the county could lead to swings of on the order of 1/4 percent. Larger manipulations of the thresholds leading to larger swings in the election output may be feasible.

Likelihood of detection:

In the absence of countermeasures, such small tinkering is very likely to go undetected.

Countermeasures:

Preventative measures:

The standard for pre-election logic and accuracy testing of optical mark-sense scanners involves scanning a stack of perfectly marked ballots. This test does not check the scanner thresholds, but only checks whether the scanner can count accurately. Augmenting this basic test with a test of scanner calibration is not hard. Ideally, the test ballots used for this purpose should be marked using not only the recommended ballot markers (number 2 soft lead pencil and black felt-tipped marker are the two most common), but also with a variety of pens and pencils representative marking implements of the kinds of markers people actually use (at the very least, several makes of black and blue ballpoint pen should be included in these tests). Ideally, the calibration test ballots should include ink and pencil specks (hesitation marks) that should not be counted as well as X and checkmarks that should be counted.

If all ballots that scan as blank or overvoted are kicked back for inspection by the voter (at the precinct) or by the canvassing board (for centrally counted absentee ballots), then this attack will quickly become visible and most of the ballots that would otherwise have been mis-evaluated will either be re-marked or correctly evaluated by people. This measure will be least effective if just one sensor of a multi-sensor scanner is miscalibrated to be underly sensitive, so that only votes read by that sensor are likely to be misread as blank; this makes totally blank ballots unlikely.

Elimination of human involvement in scanner calibration is possible. Self calibrating scanners calibrate themselves by observing the brightness variations on each ballot.

Detection measures:

Hand recounts of randomly selected precincts do not check the scanner calibration with any precision, but they will quickly find scanners that have been calibrated in an unreasonable way. Of course, the probability of detection depends on the fraction of the precincts subject to a hand recount and the fraction of the scanners that are miscalibrated.

In a machine recount, scanning on a different scanner than the one used for the first count will expose differences in scanner calibration, while scanning twice on the same scanner (without recalibration between runs) will expose the uncertainty of the machine count -- such uncertainty can arise if some ballots are marked very close to the detection thresholds.

Citations:

For a tutorial on mark-sense ballot technology, see http://www.cs.uiowa.edu/~jones/voting/optical/

(particularly Figures 8 and 9).

For a discussion of pre-election testing of mark-sense scanner calibration, see http://www.cs.uiowa.edu/~jones/voting/miamitest.pdf (section 8, pages 15 and 16).

Retrospective:

The complete lack of discussion of this issue in the 2002 voting system standards is strange. Yes, it is a matter of human factors, and the 2002 standards did not discuss human factors, but without discussion of this issue, most of the accuracy requirements of the standards as applied to optical mark-sense ballots are trivial and meaningless. What matters is how well the system captures the intent of real voters, not how well it counts perfectly marked test ballots.

The fact that very few jurisdictions properly test scanner calibration is also a serious problem. When I began doing these tests for the state of Iowa in the mid 1990s, we failed one vendor outright when their absentee ballot scanner could not come within ten percent from trial to trial on a stack of 100 ballots marked by real people. When I tried to perform these tests in Miami (cited above), I met with significant resistance.