Comments on California Secretary of State's Draft Criteria for TTB Review

Joseph Lorenzo Hall, UC Berkeley School of Information[1]

1 Introduction

The 2007 Top-To-Bottom (TTB) voting systems review is a major step forward in ensuring the reliability and integrity of the state's voting systems. While the draft criteria provide an outline of procedures that will help to improve election integrity, the criteria need additional specificity to prevent time pressure and resource constraints from adversely affecting the quality of the evaluation.

1.1 General Comments

Given the brief period allowed for public comment on the draft criteria, some discussion of general principles may be helpful to provide the Secretary with additional context for evaluating more specific comments. We discuss these general issues before going into specific, line-by-line comments.

1.2 Under-specification

At a substantive length of about four pages, the draft criteria have little choice but to leave many aspects of this review process under-specified. High-level goals such as security, reliability and usability are systems-specific issues for which general guidelines and benchmarks are difficult to write. Under-specification of these goals can lead to charges of arbitrary treatment and can lead to confusion as to the scope and nature of specific testing. Combined with the resource and time constraints discussed below, under-specification may result in no systems being about to meet the general goals of the TTB review. We go into specific areas where items could be better specified in Section 2 below.

1.3 Constraints

Successful completion of the TTB review of all systems used in California will involve completing the review of all 16 systems from 6 vendors[2] in less than four months, an average of one week per system. Source code review, only one component of the TTB review, can easily last an entire month for a single system. To accomplish the TTB review process it will undoubtedly be necessary to run multiple evaluations in parallel. The CA SoS will need to devote the resources necessary to ensure that the TTB review

---

[1] Contact: joehall@berkeley.edu. The authors' affiliations are provided for identification purposes only. The views expressed in this document are the authors' personal views. The authors do not purport to represent the views of their respective institutions.

[2] According to the CA SoS, for November 6, 2006, this included: Sequoia Optech 400C, Sequoia Optech Insight, Sequoia AVC Edge I, Sequoia AVC Edge II, Diebold AccuVote-OS, Diebold AccuVote-TSX, ES&S M100, ES&S Optech IV-C, ES&S Optech Eagle III-P, ES&S M650, ES&S AutoMARK, HART BallotNow, HART eSlate, DFM Mark-A-Vote, Ink-A-Vote OS and Ink-A-Vote Plus. *See:* http://ss.ca.gov/elections/voting_systems/systemsinuse_110606.pdf (last visited 26 March 2007).

for each voting system is thorough and comprehensive, and such a review must begin as soon as possible.

While we commend the Secretary for planning to conduct a comprehensive top-to-bottom review, the aggressive timeline might require setting some priorities. In particular, the Secretary should consider whether there are items that are of acute concern in the 2008 election year. Some requirements contemplated by the TTB draft criteria will clearly be met by only a few systems; to the Secretary might consider devoting resources to procedure and technology development aimed at bridging these known gaps, rather than a formal process of which some findings are a foregone conclusion.

For example, there are a number of criteria in the TTB draft criteria that are unrealistic given the state of technology available in the California voting system market. While we will go into this in more detail in Section 2, we'll briefly point out, for example, that only one voting system certified for use in California, the AutoMARK, meets the requirement that the system read back the contents of the verifiable paper record (§II(2)(f) of TTB draft criteria). This will result in counties that use other systems having to both procure and use the AutoMARK in each precinct.

1.4 Publication Requirements

Throughout the TTB draft criteria document – §I(3), §II(3) and §IV – there are references to the effect that the CA SoS "may" make written findings based on the results of testing. Just as the public has an interest in the testing criteria, it also has a substantial interest in the outcome of testing conducted to measure compliance with these criteria. All systems that do not meet one or more of the final adopted criteria should have a written report and findings issued publicly.

1.5 Blended Systems Evaluation

The draft criteria also don't contemplate the evaluation of blended systems – systems where different goals are accomplished by voting systems manufactured by different vendors. Will the red team exercise involve the blended system? Is the most accessible element of a blended system what will be tested in the accessibility evaluation?

2 Specific Comments

In this section we discuss specific comments we have on sections of the TTB draft criteria document.

2.1 Security

2.1.1 "untraceable vote tampering" is too narrow

Sections §I(1), §I(1)(a)-(c) and §I(3) should encompass all kinds of tampering, not just "untraceable vote tampering," for two reasons. First, if a voting system allows "traceable" or readily-visible "altering of the record of votes" or "chang[ing] the result of an election", there are cases in which that system could not be considered to be secure.

Second, the phrase "result of an election" can mean a variety of things and is unclear. We suggest this be clarified to involve changing "vote records" and "aggregate vote records" or "tallies" to reduce ambiguity.

In general, we would advise having a separate itemized section or glossary for definitions.

2.1.2 Denial of Service Attack

In the preamble to §I(1), the term "sheer physical destruction" is unclear. It seems to imply that any physical destruction is out of scope. However, some types of physical destruction are only possible through poor design of certain elements of the voting system. For example, DESI's AccuVote-TSx is known to have problems with the electrical cord easily falling out and exposing people to the risk of electrical shock. This could happen in the course of normal use and could render the system "inoperable" but might not directly affect vote records or tallies.

2.1.3 "effectively secure" is Too Vague

In §I(1)(a)-(c), voting systems are required to have features that "effectively secure" against tampering and denial of service attacks. Defining security in general is a difficult if not impossible task. However, it could be defined in the draft criteria as the systems robustness with respect to specific types of attacks. We would recommend an itemized robustness definition that includes common voting-system security concerns such as changes to vote records, changes to event logs, denial of service attacks, attacks against privacy and ballot secrecy, etc.

2.1.4 Qualifications of Security Testers

The preamble to §I(2) discusses "qualified industry and academic experts". This section should set forth more detailed qualifications for these experts. As written, it is unclear what qualifications industry and academic experts must possess. This groups might also be unduly exclusive; the Secretary should consider including elections officials with relevant expertise.

2.1.5 Red Teaming

In §I(2)(a) the red teaming procedure is very vague. It needs to be clear in what environment the testing will take place, how the red team will be constructed, what information they will have (not "might" have), how the "blue team" will operate and what will constitute a successful breach.

Another consideration is that red teaming might fall victim to a learning bias. That is, the first red team/blue team exercise will contribute to general knowledge that would be useful by both teams in the second such exercise. Essentially, the red and blue teams will "learn" additional techniques and information during each exercise. This means that a vendor's system will naturally be evaluated differently depending on if it is evaluated near the beginning of a series of red team exercises compared to the end of such a series of exercises. A natural way to eliminate a learning bias is to use different red teams and blue teams for each exercise, but that will inevitably increase expenses and reduce consistency from one test to another. We are unaware of a general method for

conducting a large number of red team exercises. It would be next to impossible for one single red team and blue team to conduct all of these exercises in the time period available.

To maximize the amount of information that the red team exercise will generate under the top-to-bottom review's timeline, the red team will need a high level of knowledge about these systems – comparable to those available to "insiders" that might be positioned to attack these systems. This should include as much information as possible and at least all the information that the CA SoS holds in escrow for forensic analysis, including source code, operational documentation, use procedures, etc. Testers can be required to sign a standard non-disclosure agreement and reports can be produced without proprietary or confidential information.

Also, the objective of the red team exercise shouldn't be as narrow as defined in the last sentences of §I(2)(a). The objective scope should include impairment of elections equipment and software as well as attacks that compromise voter privacy and ballot secrecy.

2.1.6 Source Code Review

Source code review is a time-intensive, laborious, and highly specialized process; for example, the recent thorough source code review of the ES&S iVotronic voting system conducted by the State of Florida took a team of nationally renowned computer security experts approximately one month to complete. It will be a considerable challenge to complete source code review on all 16 current voting systems currently in use in California. Outside of these time constraints, there are other concerns with the language of the source code review. The Secretary should be aware that the top-to-bottom review's time constraints will probably not allow comprehensive source code review of any system.

The word "maliciously" in §I(2)(b) should be stricken. Requiring malicious might lead the review to overlook certain classes of vulnerabilities. Some of the most serious security vulnerabilities discovered in voting systems in recent years involved designed-in "features" that could easily have been misused. Also, if there is a possibility of an operator with innocent intentions making a given mistake, it is crucial that this class of potential vulnerability is noted and that procedures are put in place to minimize the likelihood of such a mistake. Though a malicious attacker might amplify the consequences of these vulnerabilities, an attacker's intent should not be used to define the vulnerability itself.

The word "risk assessment" at the end of this section doesn't seem to make much sense. There is no other reference to a "risk assessment" in the draft criteria document and it should be made clear what part of the process is "the risk assessment" or if such a risk assessment was left out of the details (or if that reference in this section is a mistake). Typically, a source code review can be one part of a risk assessment.

2.2 Access for Voters with Disabilities

2.2.1 Disability Access Testing

The preamble to §II(2), makes it clear that "assistance of persons from the disabled community" will be relied upon to help conduct disability access testing. However, it is unclear in what role this assistance will be provided. There are two possibilities (which are not mutually exclusive): persons from the disabled community might help to design tests, or they might participate in conducting the tests. In the former case, it is important that participants have backgrounds in accessible systems development, needs assessment and/or accessibility evaluation. Human factors and usability experts should play a role in the design and evaluation of the specific tests conducted; not all of these individuals will be persons from the disabled community. Thus, this section should provide that, "The examination will be conducted with the assistance *of experts in human factors and usability as well as* persons from the disabled community."

2.2.2 Dual-switch Inputs are Only One Class of Solution

The requirement in §II(2)(a) that a dual-switch input control interface be available in every polling place will eliminate voting systems that cannot provide this kind of interface or that cannot be upgraded to provide a dual-switch input. While some voting systems (such as the AutoMARK, Hart eSlate, Sequoia AVC Edge II) provide dual-switch input, some cannot (and some of these can only provide dual-switch capability in audio ballot mode). This requirement could require that some counties procure entirely new voting systems for their precincts. Note that some voters with paralysis or manual dexterity disabilities can use other means of voting, for example, using a head wand[3].

2.2.3 Voting Systems Don't Typically Allow Changing Color Settings

While most, if not all, voting systems have a high-contrast mode and magnification capabilities, the ability to change "color settings" (§II(2)(c)) outside of contrast settings is not widely supported. In many cases, the high-contrast mode removes color from the interface entirely to display a black and white interface (although some do not).

2.2.4 Audio is Useful for the Sight-Impaired *and* Hearing-Impaired

In §II(2)(d) "variable output levels and playback speed" are associated in the draft criteria with improving accessibility for hearing impairments. However, adjustments in playback speed are more often used by sight-impaired or non-sighted voters who are accustomed to using high-speed playback on other types of accessible information technologies. This should say, "sight and hearing impairments".

---

[3] A head wand allows an individual to use a touch screen by using a wand attached to their heads. See: "Adaptive technology." *Wikipedia, The Free Encyclopedia.* 23 Nov 2006, 06:08 UTC. Wikimedia Foundation, Inc. *available at:* http://en.wikipedia.org/w/index.php?title=Adaptive_technology&oldid=89603954. (last visited on 26 March 2007)

2.2.5 Audio Playback of Paper Records is Currently Not Possible

§II(2)(f) implements a requirement currently present in the California Election Code that has not been enforced to date. Make no mistake, this is an essential aspect of voter verification that is not being provided by the voting systems market: all voters should be able to and encouraged to verify their votes. If a certain class of voters does not or cannot verify their vote, they would be a natural target for a malicious attacker, especially if there is a high probability that the machine could algorithmically determine that a given voter would be likely not to verify her vote.

However, there is only one machine certified in California that could perform this operation currently: the AutoMARK. It is non-trivial to upgrade or adapt other voting systems to allow audio playback of the paper records. This requirement could result in all counties having to procure one AutoMARK per precinct.[4]

Another concern is that of printer malfunctions. Some voting systems with paper record attachments, such as the Hart eSlate, can detect when certain types of errors occur. However, other paper record attachments will allow electronic recording of a vote record *without* the corresponding paper record. If it is not apparent to a disabled voter that nothing is being printed, and if they are provided with audio feedback that relies on printer signals instead of scanning the paper record, they will falsely be lead to believe that they had verified their vote. Here, the criteria should specify that reading back the content of the paper record by interpreting signals sent to the printer is only valid on voting systems that can detect critical, non-printing errors with the paper trail feed.

2.3 Access for Minority Language Voters

2.3.1 Recording and Playback of Minority Language Paper Records

In §III, the criteria provided do not specify if paper records should be recorded in the voter's ballot language. If they are not, voters may have a difficult time verifying that the paper record contents correspond to the summary screen presented on the machine. Of course, paper records that are not in English might be difficult to recount or manually tally. One solution to this is to provide both the ballot language and English on the paper record, which will double the length of paper records and half the capacity of paper rolls.

Also, the criteria do not specify if the mechanism for reading back the contents of a paper record would be required to playback the voter's ballot in their ballot language. Naturally, if a disabled voter chose a non-English ballot language, their audio verification should also be required to be provided in that language. Requiring Optical Character Recognition (OCR) of non-English languages could be very difficult for vendors to support, especially in languages that are written in non-Latin (non-Roman) alphabets.

[4] Note: now that the AutoMARK is no longer exclusively licensed to Election Systems and Software, this isn't as much of a burden as it would have been in the past where an entire elections system solution would have to be procured (absentee and polling place voting and election management system).

2.4 Usability for Elections Officials and Poll Workers

2.4.1 Current Work on Training and Documentation Heuristics

We commend the focus of the final section of the draft criteria on the usability of these systems from the perspective of poll workers and election officials. As a polling inspector in Alameda County, author Hall has experienced first hand how complicated the intersection of elections and technology can be.

In addition, our research team at UC Berkeley is currently involved in a project developing heuristics for poll worker documentation and training. It has become clear, in the course of this work, that poll worker documentation is a very low priority for voting systems vendors. Often the customer jurisdiction has no choice but to develop their documentation largely from scratch. Jurisdictions have mixed success on this front. We aim to provide jurisdictions with a set of heuristics based on document design principles that they could use to improve their pollworker documentation. We hope to also extend this to poll worker training. We can provide the CA SoS with preliminary heuristics that could be used as the basis for an evaluation instrument.

3 Conclusion

These draft criteria, beyond a shadow of a doubt, constitute the single most important and forward-thinking move to increase the quality of our voting systems. The timing and resource constraints involved with the TTB review are substantial and it will be a challenge to complete the evaluation process by August. We hope that our comments were helpful in refining the criteria for the TTB evaluation process. Thank you for considering public comments and we are available to expand upon these comments in public or private.