ACCURATE
A CENTER FOR CORRECT, USABLE, RELIABLE,
AUDITABLE, AND TRANSPARENT ELECTIONS

PUBLIC COMMENT ON THE VOTING SYSTEM TESTING & CERTIFICATION PROGRAM MANUAL, V2.0*

Submitted to
The United States Election Assistance Commission

January 31, 2011

*This material is based upon work supported by the National Science Foundation under A Center for Correct, Usable, Reliable, Auditable and Transparent Elections (ACCURATE), Grant Number CNS-0524745. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. This public comment was prepared by Joseph Lorenzo Hall in consultation with the ACCURATE Principal Investigators.

ACCURATE Principal Investigators

Aviel D. Rubin
ACCURATE Director
Department of Computer Science
Johns Hopkins University
rubin@cs.jhu.edu
http://www.cs.jhu.edu/~rubin/

Dan S. Wallach
ACCURATE Associate Director
Department of Computer Science
Rice University
dwallach@cs.rice.edu
http://www.cs.rice.edu/~dwallach/

Dan Boneh
Department of Computer Science
Stanford University
dabo@cs.stanford.edu
http://crypto.stanford.edu/~dabo/

Michael D. Byrne
Department of Psychology
Rice University
byrne@rice.edu
http://chil.rice.edu/byrne/

David L. Dill
Department of Computer Science
Stanford University
dill@cs.stanford.edu
http://verify.stanford.edu/dill/

Jeremy Epstein
Computer Science Laboratory
SRI International
jepstein@csl.sri.com
http://www.csl.sri.com/people/epstein/

Douglas W. Jones
Department of Computer Science
University of Iowa
jones@cs.uiowa.edu
http://www.cs.uiowa.edu/~jones/

Deirdre K. Mulligan
School of Information
University of California, Berkeley
dkm@ischool.berkeley.edu
http://www.ischool.berkeley.edu/
people/faculty/deirdremulligan

Peter G. Neumann
Computer Science Laboratory
SRI International
neumann@csl.sri.com
http://www.csl.sri.com/users/neumann/

Natarajan Shankar
Computer Science Laboratory
SRI International
shankar@csl.sri.com
http://www.csl.sri.com/people/shankar/

David A. Wagner
Department of Computer Science
University of California, Berkeley
daw@cs.berkeley.edu
http://www.cs.berkeley.edu/~daw/

1 Introduction

1.1 ACCURATE Background

A Center for Correct, Usable, Reliable, Auditable and Transparent Elections (ACCURATE),[1] a multi-institution, interdisciplinary, academic research center funded by the National Science Foundation, appreciates the opportunity to provide these comments on the draft *Voting System Testing & Certification Manual, v2.0* to the Election Assistance Commission.

ACCURATE was established in 2005 to research methods for improving voting technology. ACCURATE's Principal Investigators direct research into software architecture, tamper-resistant hardware, cryptographic protocols and verification systems as applied to electronic voting systems. Additionally, ACCURATE evaluates voting system usability and how public policy, in combination with technology, can better support elections.

Since 2005, ACCURATE has made many important contributions to the science and policy of electronic voting.[2] With experts in computer science, systems, security, usability, and technology policy, and knowledge of election technology, procedure, law and practice, ACCURATE is uniquely positioned to provide helpful guidance to the EAC as it seeks to improve the administration of the federal voting system certification program.

1.2 Overview

The draft *Voting System Testing & Certification Manual, v2.0*[3] ("VSTCP", or "the Draft Manual") would modify the requirements for vendors seeking federal certification of voting systems based on the existing testing and certification program currently governed by the *Testing and Certification Program Manual, v1.0*[4] ("VSTCPv1"). ACCURATE submitted a public comment on the VSTCPv1 in 2006,[5] and much of that commentary still stands, particularly regarding how to handle trade secrets in materials submitted by manufacturers and in VSTL materials.[6] This document focuses on the proposed changes in the Draft Manual.

Overall, ACCURATE welcomes the proposed changes to the VSTCP Manual and finds that the changes help move the EAC's testing and certification program in a number of positive directions. With these changes, the certification program will be more flexible, comprehensive, rigorous and understandable to a broad set of constituencies, including manufacturers, testing laboratories, state and local election officials, advocates, academics and voters.

Our comments in this narrative cluster around four areas: establishing a feedback loop between use of voting systems in the field and certification testing (§ 2), the introduction of the Test Readiness Review (TRR) (§ 3), changes to the requirements for test plans and test reports to make them more detailed and usable (§ 4) and proposed enhancements to procedures for the trusted build (§ 5). We feel the proposed changes could be further improved:

---

[1] *See:* `http://www.accurate-voting.org/`.

[2] See ACCURATE's list of publications (`http://accurate-voting.org/pubs/`), reports & commentary (`http://accurate-voting.org/pubs/reports/`) and testimony (`http://accurate-voting.org/pubs/testimony/`).

[3] U.S. Election Assistance Commission. *Testing and Certification Program Manual, v2.0*. Nov. 2010. URL: `http://www.eac.gov/assets/1/Documents/Test_Cert_Prog_Manual_V2.0_60_Day_Public_Comment-DRAFT-11.18.10.pdf`.

[4] U.S. Election Assistance Commission. *Voting System Testing and Certification Program Manual*. Dec. 2006. URL: `http://www.eac.gov/assets/1/Page/Voting%20System%20Testing%20and%20Certification%20Program%20Manual.pdf`.

[5] Aaron J. Burstein, Joseph Lorenzo Hall, and Deirdre K. Mulligan. *Public Comment on the Manual for Voting System Testing & Certification Program (submitted on behalf of ACCURATE to the U.S. Election Assistance Commission)*. Oct. 2006. URL: `http://accurate-voting.org/wp-content/uploads/2006/11/ACCURATE_VSTCP_comment.pdf`.

[6] Ibid., 17-19.

- The term "malfunction", referenced in § 2.3.2.7, should be defined in the VSTCP Manual and the conditions for triggering manufacturer reporting to EAC specified along with a more detailed set of reporting requirements.
- The requirement for source code review of 1% of Lines of Code (LOC) during the TRR needs to be better specified to be effective; we propose a few ways this could be improved.
- There should be explicit recognition that an important goal of the test plan and test report is to facilitate reproducibility of certification testing.
- The procedure for dealing with modifications to software in relation to the trusted build process needs to be better specified to handle each possibility of availability/unavailability of the original build environment and/or file signatures.

Finally, we include some section-specific comments in the Appendix where the issues involved are relatively minor and don't need to be explicitly mentioned in our comment narrative.

2 Closing the Loop to Support Quality Monitoring

ACCURATE researchers have called since 2001 for a closed "feedback loop" between reports of voting system anomalies that occur in the field and federal certification testing.[7] Recently, the Brennan Center for Justice at New York University School of Law released an extensive report highlighting the need for more comprehensive quality monitoring of anomalies in fielded voting systems, as well as the need for a clearinghouse that would notify election officials anomalies with systems they administer or are planning to purchase.[8]

Certification testing performed in a laboratory cannot replicate conditions that voting systems face during actual elections. Accordingly, it is important that Voting System Testing Laboratories (VSTLs) use information about voting system performance in the field for a number of reasons. First, VSTLs should use this information to better understand the conditions in which voting systems are actually used, and modify their test suites and plans to make testing more realistic. Second, VSTLs can assess whether or not critical assumptions and configuration choices made by the vendor during certification testing about the reliability, usability and security of their systems hold when they are used in the field. This will allow VSTLs to catch unrealistic assumptions during early stages of testing and require manufacturers to eliminate such assumptions or provide alternative mechanisms that render critical assumptions less critical if they are more likely to be unrealistic. Finally, for technical glitches that are seen in the field but not seen during certification testing, the VSTL can add regression tests and test cases specifically

[7] Douglas W. Jones. *Problems with Voting Systems and the Applicable Standards*. Testimony before the U.S. House of Representatives' Committee on Science. May 2001. URL: http://www.cs.uiowa.edu/~jones/voting/congress.html, (remarking "Someone should be going over those incident reports probably nationwide in consolidating the big picture of which systems are proving troublesome, what kinds of problems are showing up."), Deirdre K. Mulligan and Joseph Lorenzo Hall. *Preliminary Analysis of E-Voting Problems Highlights Need for Heightened Standards and Testing*. National Research Council's Committee on Electronic Voting. Dec. 2004. URL: http://www.law.berkeley.edu/files/evoting_standards.pdf, at 23; Erica Brand, Cecilia Walsh, Joseph Lorenzo Hall, and Deirdre K. Mulligan. *Public Comment on the 2005 Voluntary Voting System Guidelines*. A Center for Correct, Usable, Reliable, Auditable and Transparent Elections (ACCURATE). Sept. 2005. URL: http://accurate-voting.org/accurate/docs/2005_vvsg_comment.pdf, at 30–33; Aaron Burstein and Joseph Lorenzo Hall. *Public Comment on the Voluntary Voting System Guidelines, Version II (First Round)*. A Center for Correct, Usable, Reliable, Auditable and Transparent Elections (ACCURATE). May 2008. URL: http://accurate-voting.org/wp-content/uploads/2008/05/accurate_vvsg2_comment_final.pdf, at 17–18.

[8] Lawrence Norden. *Voting System Failures: A Database Solution*. Brennan Center for Justice at NYU School of Law. Sept. 2010. URL: http://www.brennancenter.org/content/resource/voting_system_failures_a_database_solution/.

meant to exercise such glitches and ensure that no future version of the voting system is certified that will display the observed anomalous behavior.

The proposed changes to § 2.3.2.7–2.3.2.8 and § 8 of the VSTCP Manual work together to better facilitate quality monitoring. § 2.3.2.7 and 2.3.2.8 together require manufacturers of EAC-certified voting systems to report to the EAC, within 30 days, a number of crucial pieces of data concerning malfunctions in voting systems and the name of each jurisdiction to which the manufacturer delivers EAC-certified voting systems. It is essential that the EAC keep a current database of which jurisdictions use particular certified voting systems in order to alert them to possible anomalies and malfunctions with this equipment. Similarly, to begin any investigation or request to a local jurisdiction to examine fielded voting systems that may have experienced malfunctions, the EAC will need data such as the "location, nature, date, impact, and resolution (if any)" (§ 2.3.2.7) of any anomaly or malfunction.

However, we feel the VSTCP Manual should go further, to better define what a "malfunction" is that would trigger reporting by the manufacturer to the EAC and what pieces of data need be reported by the manufacture to the EAC. ACCURATE PIs and researchers consulted closely with the Brennan Center in the creation of their 2010 anomaly database report, and the guidelines outlined in that report on these issues are instructive. The term "malfunction" is not defined in the VSTCP Manual so it will be left to the vendors to decide when to report a potentially problematic event to the EAC. To avoid this problem, the conditions from the Brennan Center report that trigger manufacturer reporting about an anomaly are instructive.[9] These conditions include: when the vendor receives a complaint from an election official about the performance of their products; when they receive a warranty claim or take warranty-fulfilling action; when they conduct an investigation of a reported anomaly; and, when they are sued in relation to an anomaly. Similarly, the reporting requirements suggested by the Brennan Center report seem more useful and comprehensive than the minimal set outlined in § 2.3.2.7 of the VSTCP Manual, including: the location where the malfunction occurred; a description of the malfunction; the vendor or vendors of the voting system; all versions of hardware and software and relevant configurations affected; a list of each jurisdiction where this voting system is used; the manufacturer's assessment of any risk to election outcomes due to the malfunction; the vendor's suggested fix or workaround; and how the vendor intends to remedy or fix the malfunction in the future.[10] These are suggested changes we support that will make quality monitoring and outreach to affected jurisdictions more effective. At a minimum, the EAC must consider whether the important, although fairly simple, changes proposed in these reporting elements will best facilitate quality monitoring.

Further, § 8, "Quality Monitoring Program", has two crucial changes in v2.0 that will create more of a closed loop between certification testing and issues reported from use in the field. § 8.6 outlines the possibility of fielded voting system reviews. While we feel that the EAC should be a bit more aggressive in terms of requesting permission to review reported anomalies in fielded voting systems, we welcome the addition of language that clarifies any such review will compare the state and use of a fielded voting system to the expected configuration and use represented by the vendor and tested by the VSTL during certification testing.

[9] Norden, see n. 8, at 32.

[10] We note that the data requested on the proposed Anomaly Reporting Form used by election officials as part of the Quality Monitoring Program (§ 8) is more extensive than what is required of vendors in reporting data about malfunctions in § 2.3.2.7 the VSTCP Manual. It seems the data required in both of these cases should be better harmonized. (The Anomaly Reporting Form is available at: `http://www.eac.gov/assets/1/Page/Anomaly%20Reporting%20Form.pdf`.)

3 Test Readiness Review

The Draft Manual adds § 4.4, the "Test Readiness Review" (TRR), in order to add an initial step to certification testing that conditions further testing on a set of tests that attempt to assess if a voting system is ready for VSTL testing. This is a laudable step as it has the potential to make time-to-certification shorter and lower the cost of certification due to systems unprepared for testing.

However, there are a few issues that need to be clarified or enhanced. Preliminary source code review in § 4.4 requires 1% of total lines of code (LOC) to be examined. There is no guidance in the VSTCP Manual as to how these 1% of source code lines should be chosen by the VSTL. If they are chosen from the same or similar modules, the VSTL runs the risk of choosing source code that is particularly better or worse than the rest of the source code. Especially given that the definition of Lines of Code[11] includes "formatting (e.g. Blank lines) and comments" this provision also risks an evaluation based solely on comments in the code.

We would like to suggest a more effective procedure where a VSTL randomly chooses coherent pieces of source code until the aggregate amount reaches 1% of the total LOC. In addition to random selection of source code to ensure that all the relevant code has similar probability of selection, it is critical to have context from the surrounding code to understand a given line of source code. The VSTL should randomly choose a subset of functions (i.e., methods, procedures, etc.), chosen so that the aggregate count of LOC in the bodies of those functions is at least 1% of the source code. The VSTL would then review all of the bodies of those functions. That is, the VSTL could repeatedly select a random function, mark it "to be reviewed", count the number of lines in that function, add it to a running total, and continue random selection until the total LOC "to be reviewed" exceeds 1% of the lines of code. Alternatively, modules, components or source code files—instead of functions/methods—are coarser-grained units than functions, so randomly selection a this level of granularity might be less burdensome for reviewers than it would be to have them select a random subset of modules/components/files. Finally, a minimal alternative that will work to address our concerns here would be to require the VSTL to document what process they used to randomly select 1% of the code their in their TRR notice to the EAC and then EAC experts that review this document can evaluate this scheme as part of the TRR notice and acknowledgment process.

While it may seem minor, we feel that § 4.4.1 should clarify that the VSTL must affirmatively state in the Test Readiness Notification that they've performed the TRR and that the system has passed and appears ready for further testing.

4 Changes to Test Plan and Test Reports

In addition to the incorporation of the TRR and an explicit coding convention declaration (§ 4.3.1.6.4), we welcome the proposed modifications to test plans (§ 4.4) and test reports (§ 4.5).

ACCURATE appreciates the clarifications in what a test plan must do, must contain and must be useful for. ACCURATE PIs and researchers participated in both the California Top-To-Bottom Review[12] (TTBR) and the Ohio EVEREST Review[13], where document review made extensive use of TDP materials, test plans and test reports. The document review reports from the TTBR discuss difficulties we faced in ascertaining, even with privileged access to all the materials a VSTL would have, exactly

[11] U.S. Election Assistance Commission, *Testing and Certification Program Manual, v2.0*, see n. 3, at 9.

[12] California Secretary of State. *Top-To-Bottom Review of California's Voting Systems*. Aug. 2007. URL: `http://www.sos.ca.gov/voting-systems/oversight/top-to-bottom-review.htm`.

[13] Patrick McDaniel, Matt Blaze, Giovanni Vigna, et al. *EVEREST: Evaluation and Validation of Election-Related Equipment, Standards and Testing (Academic Final Report)*. Ohio Secretary of State. Dec. 2007. URL: `http://www.sos.state.oh.us/SOS/upload/everest/14-AcademicFinalEVERESTReport.pdf`.

how tests were conducted and under what conditions and configurations.[14] Accordingly, we welcome the new level of detail for test plans outlined in § 4.5. However, we would like to see more emphasis placed on reproducibility of testing. One important constituency for test plans are state-level voting system testing authorities as well as independent evaluators hired by voting system customers and/or voting system vendors. There are good analogs in other areas of Federal government regulatory testing; for example, the test plans submitted by vendors to the Federal Communications Commission[15] capture very detailed aspects testing data including detailed images and descriptions of testing configuration and set-up, performance in the test and detailed descriptions of the test method, useful if the FCC or another entity must reproduce the results of the test.

The clarifications in § 4.5 and § 4.7.2 that specify test plans and test reports should be usable to a wide range of readers are also very welcome, especially as they go further to encompass the voting public as members who might find these documents useful. This is an important part of including election officials in the certification process as election officials do not generally come from technical disciplinary backgrounds.

5 Enhancements to Trusted Build Procedures

We are very encouraged by the changes to the procedures for creating, disseminating and maintaining chain of custody over the trusted build.

Requiring that VSTLs independently procure COTS software products in § 5.5.5 will ensure that COTS software is unmodified and not designed for voting-specific applications. We see one wrinkle that should be clarified: the phrase "The *source* of each COTS product. . . " (emphasis added) in § 5.5.3 clearly is not intended to refer to source code, as obtaining source code availability of COTS software has been a difficult proposition in past certification and evaluation efforts. An alternative wording for this phrase that made it clear that the TDP should specify how the VSTL obtained COTS software would be, "The vendor or source from which each COTS product was procured must be included in the TDP."

§ 5.6.4 addresses how a trusted build can be created for modifications, where some of the source code has changed but only in a subset of files. We agree that if the original build environment and file signatures are available, the trusted build for the modified system can be created by verifying the signatures of unmodified files, including the modified files in the build environment (after source code review and other testing) and then building new executables.

§ 5.6.4.3 addresses the case where the original build environment is unavailable and/or file signatures cannot be verified. This subsection includes the statement: "Further source code review may be required of unmodified files to validate that they are unmodified from their originally certified versions."[16] This should be clarified somewhat; there are three cases involved here:

- If the original build environment is unavailable, but file signatures are available, unmodified files that pass signature verification should be acceptable.[17]

[14] Aaron J. Burstein, Nathan S. Good, and Deirdre K. Mulligan. *Review of the Documentation of the Sequoia Voting System*. July 2007. URL: `http://www.sos.ca.gov/voting-systems/oversight/ttbr/sequoia-doc-final.pdf`, § 3 (at 10–18), § 4.2.2 (at 30–31); Joseph Lorenzo Hall and Laura Quilter. *Documentation Review of The Hart Intercivic System 6.2.1 Voting System*. July 2007. URL: `http://www.sos.ca.gov/voting-systems/oversight/ttbr/hart-doc-final.pdf`, § 4.6 (at 24–34); Candice Hoke and Dave Kettyle. *Documentation Assessment of the Diebold Voting Systems*. July 2007. URL: `http://www.sos.ca.gov/voting-systems/oversight/ttbr/diebold-doc-final.pdf`, § 4.1 (at 17–21).

[15] To examine an example of a test report submitted to the FCC, visit `https://fjallfoss.fcc.gov/oetcf/eas/reports/GenericSearch.cfm` and enter a recent date range. Click on "Detail" and then read a corresponding test report.

[16] U.S. Election Assistance Commission, *Testing and Certification Program Manual, v2.0*, see n. 3, at 43.

[17] We assume here that there has been a well-maintain chain of custody over the file signatures such that they could not be compromised.

- If the original build environment is available and unmodified files do not pass signature verification, the unmodified files will have to be compared against their counterparts in the original build environment, using tools like `diff`[18], to ensure they are exactly the same. In fact, it should be possible to simply copy the original build files that are claimed to be unmodified to the new build environment such that the new (copied) file passes signature verification.
- Finally, if the original build environment is unavailable and either unmodified files do not pass signature verification or file signatures are unavailable, these supposedly unmodified files will have to be treated as modified files as there is no basis to verify that they have not been modified.

6 Conclusion

ACCURATE appreciates the opportunity to comment on the draft VSTCP Manual. The EAC's Voting System Certification Program has come far since the adoption of VSTCPv1 in 2006 and ACCURATE welcomes the proposed changes to the VSTCP Manual outlined in the draft. We offer these constructive suggestions in the hope that voting system certification can continue to evolve and better serve our country and the principles for which the certification program was designed to test. We would be happy to answer any questions the EAC has about our comments and engage in further dialog with EAC, VSTLs, vendors or technical experts.

[18] `diff` and similar programs compare lines of two files, one-by-one, and can identify lines of source code that differ between two files.

Appendix

In this Appendix we include an additional list of our section-specific comments, where a given issue did not rise to the level of importance of the discussion in the body of our comment. Where applicable we include references to pertinent sections of the VSTCP Manual.

1. § 1.16: The VSTCP Manual references "COTS" ("Commercial off-the-shelf") in a number of places[19] but it does not define this term. The list of definitions in this section must define COTS or include a reference to the COTS definition in the Glossary of the 2005 VVSG.[20]

2. § 2.3.1.6: The three paragraphs following the initial paragraph in this section are not justified to align with the 2.3.1.* list. As such, it is unclear if these three paragraphs are part of § 2.3.1.6 or perhaps part of the parent element, § 2.3.1. They should be justified to be part of § 2.3.1.6 or moved and/or set apart to make clear that they are part of some other requirement or commentary of some sort.

3. § 4.5.4.3: The forth bullet in this subsection (p. 37) includes an incomplete sentence. It currently reads (emphasis added):

   > The test plan is a living document and is expected to change and be updated during various phases of the testing life cycle. A final version that reflects all of the testing completed (TDP, S/W, Hardware, Software etc) should be submitted to the EAC at the completion of testing. *If this final "as run" test plan does not reflect all the testing required.* The EAC reserves the right to request further updates to the test plan and possibly additional testing.

   It seems most likely that the period after the incomplete sentence highlighted above should be a comma. However, we recommend that the incomplete sentence should be deleted to emphasize that the EAC reserves the right to request updates without qualification.

4. § 5.7.7: The abbreviation "HDD", presumably meaning "hard disk drive" is not defined elsewhere in the manual. This abbreviation should be defined when it is first used.

5. § 8.6 and 8.8.7: Both these sections use the term "correlation" to compare the use and configuration of fielded voting systems to the state of those systems during certification testing. Technically, "correlation" has a very specific meaning referring to quantitatively comparing the interdependence of two variables. To avoid any confusion—e.g., that the EAC will be comparing two lists of variables between fielded voting systems and those systems as tested at the VSTL—we suggest a more qualitative word be used, such as "correspondence".

[19] U.S. Election Assistance Commission, *Testing and Certification Program Manual, v2.0*, see n. 3, at 13 (§ 2.3.1.6), 32 (§ 4.5.1) and 41 (§ 5.5.3, 5.5.5).

[20] U.S. Election Assistance Commission. *2005 Voluntary Voting System Guidelines*. Dec. 2005. URL: `http://www.eac.gov/testing_and_certification/2005_vvsg.aspx`, at A-7.