Draft Recommendations for Software Distribution: Supplement to the 2002 Voting Systems Standard

Draft Version March 2, 2005

National Institute of Standards and Technology (NIST)

Draft

Provided for consideration by the Technical Guidelines Development Committee (TGDC) and the Election Assistance Commission under the requirements of the Help America Vote Act (HAVA) of 2002.

Acknowledgements

The National Institute of Standards and Technology (NIST) would like to acknowledge the individuals and groups who helped contribute to the preparation of this document. Members of Technical Guidelines Development Committee (TGDC) and the NIST voting team that provided substantial assistance. NIST would also like to acknowledge the IEEE organization for permission to use excerpts from the IEEE P1583 Draft Standard for the Evaluation of Voting Equipment.

Draft

Authority

This document has been provided for consideration by the Technical Guidelines Development Committee (TGDC) and the Election Assistance Commission under the requirements of the Help America Vote Act (HAVA) of 2002.

Disclaimer

This document is a work in progress, provided solely as draft input to the TGDC. Portions of this document may change substantially. This document references some material from an unapproved draft of a proposed IEEE Standard, P1583. As such, the material is subject to change in the final standard. Because this material is from an unapproved draft, the IEEE recommends that it not be utilized for any conformance/compliance purposes. It is used at your own risk.

Draft

1. Introduction

The Technical Guidelines Development Committee (TGDC) of the Election Assistance Commission (EAC) has tasked the National Institute of Standards and Technology (NIST) via TGDC resolution 15-05 “Software Distribution” to research and draft standards for the distribution of voting software to increase the assurance of voting systems. This document represents NIST’s response to the tasking called for in TGDC resolution 15-05 “Software Distribution.”

1.1 Scope

This section analyzes and scopes the issues raised in TGDC resolution 15-05 “Software Distribution” to provide a framework by which the issues can be addressed systematically. The complete text of TGDC resolution 15-05 “Software Distribution” has been included in Appendix A as a reference.

The TGDC resolution 15-05 “Software Distribution” identifies the importance of knowing what software is installed on a voting system, when the software was installed, and the origin/source of the software. However, the resolution does not include text related to the correct operation of the software. Although the correct operation/behavior of the software is of critical importance in providing a high level of assurance for the overall voting system, this issue is beyond the scope of TGDC resolution 15-05 “Software Distribution” so will not be addressed in this document. Additionally, the resolution does not include text related to the assurance of the computer based platform on which the voting software is to be installed. Although the process used to distribute and install software leads to enhance the assurance of the platform on which it is installed and executed, the general issue of the assurance of the computer based platform is within the purview of TGDC resolution 16-05 “Setup Validation” and beyond the scope of this resolution so will not be addressed by this document.

For the purposes of this document, software will be considered executable code and associated configuration files (such as ballot formats developed by jurisdictions) used by the voting system. As the focus of this resolution is to determine whether the set of specifically identified software for a voting system has been distributed without modification, the proposed requirements are not limited to the software developed by the vendor of the voting system but include third party software such as operating systems, drivers, etc. that are critical for the proper operation of the voting system. The final location within the voting system (firmware, read-write media, write-one media, etc.) of the installed software does not affect the applicability of the proposed requirements. Third party development tools (such as compilers) used by vendors are generally not distributed as parts of voting system so are beyond the scope of this document.

The text of TGDC resolution 15-05 “Software Distribution” is ambiguous with regard to the scope of the voting system and associated software to be covered by the resolution. In the general, a voting system is a large system composed of several other systems including polling place systems, central counting/aggregation systems, and election management systems. These systems reside on several different computer based platforms and at different locations. The introductory text of the resolution talks about voting systems in general implying the scope of the resolution covers all software used to support a complete voting system. However, item (2) of the resolution creates some ambiguity by specifically referencing the specific term “voting machine” which seems to limit the scope of the resolution to only the software used by systems deployed to polling places used to collect cast ballots. Item (3) of the resolution adds to the ambiguity by using the phrase “the process of loading the software” without identifying the systems that the software is being loaded. In addition, item (4) of the resolution identifies “a system audit log” again without identifying the systems. As a result, items (3) and (4) of the resolution seems to include all software used to support a complete voting system not just systems deployed to polling places to collect cast ballots. To be consistent with the focus of the resolution, which is to determine whether the set of specifically identified software for a voting system has been distributed without modification, this document will consider all software within scope of the resolution irregardless of the location of installation and functionality provided to support the overall voting system.

Item (1) of TGDC resolution 15-05 “Software Distribution” states the distribution of any software to voting systems shall be performed via physical distribution using a “read-only” or “write-once” media such as a CD-R, ROM, PROM (not EEPROM, CD-RW). Distribution of software used to support voting systems can originate from at least four different points: voting system vendors, third party vendors, Independent Testing Authorities (ITAs)/Voting System Testing Laboratories (VSTLs), and jurisdictions. Voting system vendors distribute the set of software they have created as well as the required third party to ITAs/VSTLs for testing. Voting system vendors or ITAs/VSTLs distribute the set of qualified software to jurisdictions for use. Finally, jurisdictions distribute/install/load the set of qualified software and customized configurations files to the computer based platforms for use in the voting process. The distribution of software by jurisdictions seems overly restrictive based on text of item (2) of the resolution that states “the electronic transmission of any software to voting machines” is acceptable under certain conditions. Therefore, this document assumes the requirement for physical distribution of “write once” media is limited to software distributed between vendors and ITAs/VSTLs as well as between ITAs/VSTLs/vendors and jurisdictions. Other methods may be used when distributing the software to computer based platforms.

Item (2) of TGDC resolution 15-05 “Software Distribution” makes a general statement about the risk introduced by “the electronic transmission of any software to voting machines”. Although the resolution qualifies the statement using the restrictive term “voting machines,” the alluded to risk is the same for any software distributed via an “electronic transmission” technology. Several different technologies exist that can be used for the “electronic transmission” of software with built-in and/or add-on mechanisms used to provide security. TGDC resolution 35-05 “Wireless” begins to address the security issues associated with one specific “electronic transmission” technology and sections 5 and 6.6 of volume I of the 2002 Voting System Standard, tries to address these security issues in general. Since it is unclear what security can and will

Draft

be provided by the specific “electronic transmission” technology, this document will take the conservative position of assuming the “electronic transmission” technology is unsecured which will affect how the requirements are developed to address this resolution.

Items (3) and (4) of TGDC resolution 15-05 “Software Distribution” references the phrase “loading of the software.” The term “loading” in the phrase can be interpretation in at least two ways when talking about computer based systems. The term “loading” in the phrase could mean when software is being “loaded” or installed onto the computer base platform. Another interpretation could be that the software already resides or is installed on the computer based platform and is “loaded” or transferred into the platform’s memory for execution. Although technology exists to support the requirements of items (3) and (4) when software is transferred into the computer based platform’s memory for execution, this interpretations seems to overly extend the focus of this resolution which is to determine whether the set of specifically identified software for a voting system has been distributed without modification. Therefore, this document assumes the phrase “loading of the software” to mean the installation of software on a computer based platform for the purpose of conducting voting activities but not loading of the software into the platform’s memory for execution.

Draft

The following is a summary of the assumptions and interpretations upon which the proposed recommended requirements for software distribution will be created:

- requirements for determining or validating the correct operation or behavior of the software will not be addressed in this document;
- the assurance of overall the computer based platform on which software is installed is not addressed in this document;
- not limited to the software developed by the vendor of the voting system but include third party software such as operating systems, drivers, etc. that are critical for the proper operation of the voting system.
- software will be considered the executable code and associated configuration files used by the voting system irregardless of the location of installation and functionality provided to support the overall voting system;
- physical distribution of “write once” media is limited to software distributed between vendors and ITAs/VSTLs as well as between ITAs/VSTLs/vendors and jurisdictions. Other methods may be used when distributing the software to computer based platforms;
- the “electronic transmission” technology used to distribute software is not secured; and
- the phrase “loading of the software” to mean the installation of software on a computer based platform used for conducting voting activities.

2. Authoritative Sources and Reference Information for Voting System Software

The focus of TGDC Resolution 15-05 “Software Distribution” is to determine whether the set of identified software for a voting system has been distributed without modification. As stated in the previous section, software is considered to be the executable code and associated configuration files of a voting system. In general, static software such as executable code does not change based on election being conducted or voting machine on which they are installed. Semi-static software contain configuration information for the voting system and are modified based the voting machine on which they are installed and the elections being conducted. Semi-static software is only modified during the installation of the voting system software on a voting machine and before a given election. In order to determine if modifications have occurred to specific static or semi-static software, reference information (hash value, digital signature, or binary image) needs to be created from a known, fixed binary image of the specified software. Once the reference information has been created, it can be compared with similar information (hash value, digital signature, or binary image) generated using the software being inspected. Two factors determine the level of assurance that can be provided by the software inspection: (a) the source that creates the reference information and (b) the type and quality of reference information.

Draft

Authoritative sources create reference information from known, fixed binary images of the software obtained from appropriate sources such as vendors or developers of the voting system software. In general, there are three main sources can be the authoritative sources for the reference information of the voting system software: vendors, ITAs, and jurisdictions. Voting system software vendors are the source for the known, fixed binary images of the executable code of their voting systems which are created by their development tools such as compilers. The vendors can create reference information as an authoritative source using the binary images of the executable code of their voting system. ITAs use the executable code[1] they receive from voting system vendors to perform qualification testing. Once the voting system software meets all of the requirements of the qualification tests, the ITAs are the main source for the known, fixed binary images of the executable code that have passed the qualification tests. The ITAs can create reference information as an authoritative source using the binary images of the executable code that passed the qualification tests. Voting system software vendors and ITAs cannot be the authoritative sources of reference information for the semi-static software containing configuration information dependant on specific elections information and voting machine installation details. However, jurisdictions that conduct elections and install voting system software on voting machines are the main source for the known, fixed binary images of the semi-static software that contain configuration information. Jurisdictions can create reference information as an authoritative source using the binary images of the semi-static software that contain configuration information.

An alternative authoritative source for reference information for voting system software could be the National Software Reference Library (NSRL) housed at NIST. The NSRL

[1] ITAs receive source code as well as executable code for qualification testing. However, this is beyond the scope of TGDC resolution 15-05.

was established to meet the needs of law enforcement community for court admissible digital evidence by providing an authoritative source of commercial software reference information. The NSRL generates reference information for commercial software from the official original CDs and floppy disks of the software. Currently, the NSRL contains over 6000 software products including some voting system software and has generated 10 million unique pieces of reference information for the software. NIST publishes the references information quarterly as the Reference Data Set (RDS). The NSRL could be leveraged as the authoritative source of reference information for voting systems. In order to be used as an authoritative source, the NSRL will need to obtain the known, fixed binary images of the executable code that have passed the qualification tests from the ITAs. The NSRL could be the authoritative source for reference information of the semi-static software containing configuration information dependant on specific elections information and voting machine installation details, if jurisdictions supply the appropriate known, fixed binary images.

Draft

In general, authoritative sources can create three types of reference information, which can be used to detect when voting system software has been modified: (a) complete binary images, (b) cryptographic hash values, and (c) digital signatures of the software. Complete binary images of voting system software are a simple non-cryptographic approach to form reference information. The complete binary images can be placed on a write once media such as a CD by the appropriate authoritative source. A chain of custody for the write once media containing the complete binary images is required to form the traceability back to an authoritative source. The write once media containing the complete binary images needs to have unique identifiers and markings such as holograms, watermarks, etc. that is difficult to recreate and deters forgery and substitution of the media. To verify the integrity of the software, a complete binary comparison between the set of voting system software and binary images of the reference information on the write once media needs to be performed. If the binary images of the reference information are copied off the write once media they can be undetectably modified; so it is important that the binary images used during the comparison come directly from the write once media. This technique of verifying the software integrity can yield very specific information on exactly where and the amount of software that has been altered or modified but may be very time consuming. However, this and the other techniques described cannot determine if the modification or alteration to the software was done maliciously or benignly. Finally, complete binary images of the voting system software are considered highly valuable property of vendors. When using complete binary images as reference information, vendors may be unnecessarily exposing their valuable property. The other techniques described in this document do not require complete binary images of the voting system to be distributed to verify voting system software integrity.

Cryptographic hash functions are an approach to creating reference information. In general, hash functions take an arbitrary length binary input and create a fix length binary output called a hash-value or hash. There are two types of hash functions: non-cryptographic and cryptographic. Non-cryptographic hash functions called checksums (such as cyclic redundancy codes (CRC)) provide protection from accidental or nonmalicious errors of binary information generally in noisy transmission channels. Checksums were not designed to detect intentional modifications, so do not posses the security properties needed to produce reference information for voting system software. Cryptographic hash functions make it difficult to determine the input string that generated the specific hash value and find two input strings that generate the same hash value. Given these properties, cryptographic hash functions can be used to generate hash values for reference information of voting system software. An authoritative source can generate hash values for known, fixed binary images it has purview over and places these hash values on a write once media such as a CD. As with the write once media containing complete binary images, a chain of custody for the write once media containing the hash values is required to form the traceability back to an authoritative source. The write once media containing hash values needs to have unique identifiers and markings such as holograms, watermarks, etc. that is difficult to recreate to deter the forgery and substitution of the media. To verify the integrity of the software, the software to be inspected will be used as input to the cryptographic hash function to produce a hash value that is compared to the hash value associated with that software on the write once media. Similar to the binary images if the hash values are copied off the write once media they can be undetectably modified; so it is important that the hash values used during the comparison come directly from the write once media. This technique for verifying software may be faster than a complete binary image comparison of the software but can only determine if the software has been altered or modified. This technique cannot provide specific information on exactly where and the amount of software that has been changed.

Draft

Digital signatures are another approach to creating reference information that builds on cryptographic hash values. Digital signatures are created by encrypting the hash value of a known, fixed binary image using the private key of a public key cryptographic algorithm.[2] Depending on the digital signature scheme used verification of a digital signature can be performed in two different ways. In the first verification method, a hash value is generated using the binary image of the software to be verified. Next, the digital signature from an authoritative source is decrypted using the authoritative source's public key. Finally, the results of the generated hash value and the decrypted digital signature are compared and if they are equal the signature is valid. Otherwise the signature is invalid due to one of two reasons: (1) the binary image of the software has been modified or (2) the public key was used was incorrect. In the second verification method, a hash value is generated using the binary image of the software to be verified. The digital signature, the calculated hash value, and the public key are given to a verification function that indicates whether the signature is valid or invalid. The digital signature approach for reference information has the additional requirement that public and private cryptographic key information be managed (generate, stored, distributed, and destroyed) properly which is not found when using the binary image and hash value approaches. However unlike the binary images and hash values approaches, if the digital signatures are copied off the write once media they cannot be modified without being detected.

[2] There are digital signature schemes in which symmetric cryptographic algorithms are used. However symmetric cryptographic algorithms require the use of shared secret keys calling into question the generator of the digital signature.

Similar to the hash value approach, digital signatures cannot provide specific information on exactly where and the amount of software that has been changed.

The different authoritative sources can use the techniques described to create reference information – binary images, hash values, or digital signature – for the voting system software they have purview over. By using these techniques, one can determined if a specific set of voting system software has been distributed without modification.

3. Proposed Recommendations

The following are NIST recommendations to the TGDC be considered for inclusion in voting system standards related to the distribution of software used by voting systems.

3. 1 General Requirements

1) The ITA shall witness the final build of the executable version of the qualified voting system performed by the vendor. [related to VSS Volume I, Section 9.4.1.4]

2) Complete binary images of voting system software including installation programs shall be distributed on a "write once" by authoritative sources (vendors, ITAs/VSTL, and jurisdictions) [NEW from Resolution item 1]

- CD-R, ROM, PROM, etc. satisfy this requirement
- EPROM, CD-RW, etc. do not satisfy this requirement

3) Authoritative sources that generate reference information shall document the list of voting system software covered and the type of reference information.

4) Hash values used as reference information to verify the integrity of qualified voting system software (including installation programs) shall be distributed on a "write once" media by authoritative sources (vendors, ITAs/VSTLs, NSRL, and jurisdictions). [NEW from Resolution item 3]

- CD-R, ROM, PROM, etc. satisfy this requirement
- EPROM, CD-RW, etc. do not satisfy this requirement

5) The "write once" media containing binary images and hash values of the voting system software shall be labeled by authoritative sources (vendors, ITAs/VSTLs, NSRL, and jurisdictions) so that is uniquely identifiable (including the authoritative source and date created.) [VSS Volume I, Section 3.4.6]

- Researching the possible ways to include security labels such as those with holograms, etc.

6) A chain of custody record shall be kept by vendors, ITAs/VSTLs, NSRL, and jurisdictions for "write once" media containing binary images and hash values of the voting system software indicating at minimum the dates and who have handled the media.

Draft

7) Authoritative sources (vendors, ITAs/VSTL, NSRLs, and jurisdictions) that generate sets of hash values and digital signatures as reference information shall include a hash value or digital signature covering the set of all hash values and digital signatures.

8) Hash values and digital signatures used for reference information shall be generated by authoritative sources (vendors, ITAs/VSTLs, NSRL, and jurisdictions) using a FIPS 140-2 level 1 validated cryptographic module.

9) Digital signatures shall be verified using a FIPS 140-2 level validated cryptographic module.

10) The authoritative sources (vendors, ITAs/VSTLs, NSRL, and jurisdictions) that generate hash value and digital signature reference information shall use a FIPS approved hash function

- FIPS 180-2, Secure Hash Standard, August 2002
- SHA-1, SHA-224, SHA-256, SHA-384, and SHA-512

11) The authoritative sources (vendors, ITAs/VSTLs, NSRL, and jurisdictions) that generate digital signature reference information shall used a FIPS approved digital signature scheme

- FIPS 186-2, Digital Signature Standard, February 2000 (DSA, RSA, ECDSA)
    - Appendix 6 contains recommended curves for ECDSA
- ANSI X9.31-1998, Digital Signatures Using Reversible Public Key Cryptography for the Financial Services Industry, 1998 (RSA)
- ANSI X9.62-1998, Public Key Cryptography for Financial Services Industry: The Elliptic Curve Digital Signature Algorithm, 1998 (ECDSA)

12) Public keys used to verify digital signatures shall be distributed using a "write once" media from an authoritative source (vendors, ITAs/VSTLs, NSRL, and jurisdictions) or a secure electronic mechanism such as digital certificates.

13) The "write once" media containing public keys to verify digital signatures shall be labeled by authoritative sources (vendors, ITAs/VSTLs, NSRL, and jurisdictions) so that is uniquely identifiable (including the authoritative source and date created.) [VSS Volume I, Section 3.4.6]

- Researching the possible ways to include security labels such as those with holograms, etc.

14) A chain of custody record shall be kept by vendors, ITAs/VSTLs, NSRL, and jurisdictions for "write once" media containing public keys to verify digital signatures indicating at minimum the dates and who have handled the media.

Draft

4. References:

[HAC] Alfred Meneze, Paul C. vanOorschot, and Scott A. Vanstone, Handbook of Applied Cryptography, CRC Press, New York, NY, 1997.

[FIPS 140-2] Federal Information Processing Standard 140-2: Security Requirements for Cryptographic Modules, May 25, 2001.

[FIPS 180-2] Federal Information Processing Standard 180-2: Secure Hash Standard, August 2002.

[FIPS 186-2] Federal Information Processing Standard 186-2: Digital Signature Standard, February 2000.

[ANSI X9.31] ANSI X9.31-1998, Digital Signatures Using Reversible Public Key Cryptography for the Financial Services Industry, 1998.

[ANSI X9.62] ANSI X9.62-1998, Public Key Cryptography for Financial Services Industry: The Elliptic Curve Digital Signature Algorithm, 1998.

[VSS2002] Federal Election Commission, Voting System Standard, 2002.

Draft

Draft

A.1 Appendix A: Text of TGDC Resolution 15-05: Software Distribution

The TGDC has concluded that, generally speaking, the manner in which software is loaded onto voting systems is not governed by existing standards and that it is a significant security issue that warrants more stringent controls. It is important to know which software has been installed on a voting system, when the software has been installed, and from what sources. Without strict controls on these processes, non-certified software could be loaded onto voting systems, with potentially disastrous results. The TGDC directs NIST to research and draft standards documents requiring:

1. That the distribution of any software to voting systems shall only be performed by means of physically distributed "read only" or "write once" media, including software such as:
   (a) operating system required software,
   (b) updates and patches,
   (c) data files, and
   (d) voting system software.
2. That the electronic transmission of any software to voting machines via networks or wireless introduces extreme risk and should be approached with extreme caution,
3. That the software will include an integrity check (such as a digital signature that positively authenticates its source) that must be verified as part of the process of loading the software, and
4. That the record of loading the software will be written permanently to a system audit log kept in write-once memory.