An Electronic Election System

Robert L. Walton[1]

October 24, 2021

Table of Contents

1 Introduction

Our purpose is to specify an all-electronic election (E-Election) System suitable for replacing the current optical paper system used in Massachusetts. Our system is based on the notion of 'tamper-proof software', which is software that is known to do what it should do. With the aid of tamper-proof software, we can make a verifiably secure election system without using paper ballots.

Tamper-proof software may have other applications. It could be of use with the existing optical paper system, for example, to ensure that the software within the ballot boxes does what it is supposed to do.

2 Tamper-Proof Software

Tamper-Proof Software is software written in a 'tamper-proof' programming language that has the following property:

> The translation of source code to binary files is so precisely specified that several independent groups of programmers can write compilers that will produce identical binaries from any given source.

Clearly a sufficiently simple assembly language qualifies, but we need a somewhat higher level language. Still, such a language should be possible. Two aids to this end are:

1. Computer efficiency is not required in this application.
2. If two compilers disagree on some source, the matter can be investigated, and either one or both compilers fixed, or the language specification improved.[2]

The real goal is obtaining binaries that can be trusted to do what the source code says they should do. The other requirement on tamper-proof software is that the source code be published, so the world can see what it does and look for flaws.

The simplest tamper-proof programming language is a straight forward simple assembly language. Given this, one can write an interpreter for a simple interpreted language in the assembly language. The interpreted language should suffice for most of the code, and can be machine independent. A few functions can be optimized by being written in assembly language.

[2]This method has been used to bring Ada compilers into agreement so all produce code that does the same thing.

3 E-Election System Overview

A Private Voting Machine (PVM) is constructed from a single thumb drive and a registration card. Part of the thumb drive contains the tamper-proof software of the PVM, part contains an encrypted data set called a Voter Lock Box (V-Box), and the registration card contains a passkey that is entered on the keyboard in order to decrypt/encrypt the V-Box. The software runs on a (fairly) arbitrary computer, is bootstrapped directly, runs without using any operating system or BIOS, uses only RAM memory during operation, writes the encrypted V-Box back to the thumb drive, and sends a very few encrypted messages to other 'machines' in the system.

The other machines are the Registration Machines, which make ready-to-mark ballots available and time-stamp and certify finished ballot cryptographic signatures; the Collection Machines that collect finished ballots; and the Counting Machines that count the ballots. Each machine has the same design as the PVM, except these other machines have different software, are each run by a single election official instead of a voter, and are replicated with different officials running different replicas. Counting Machines also use separate thumb drives containing master lock boxes (M-Boxes), and a Counting Machine requires any 5 out of 10 election officials to cooperatively provide decrypted M-Boxes in order to decrypt ballots and produce final counts.

Each machine and each M-Box has a public key and a private key. The E-Election machines that participate in a particular election for a particular precinct form a virtual private network that is disconnected from the rest of the world, the internet in particular. This E-Election network works by encrypting all messages with the private key of the sender and the public key of the receiver, and distributing keys only by directly copying them from thumb drives, with the exception of public keys of Registration and Collection Machines. Messages to and from Private Voting Machines do pass through the internet, but as they are encrypted by PVM keys that never pass through the internet, the E-Election network is informationally disconnected from the internet.

Each precinct has its own separate set of Counting Machines and M-Boxes, but precincts may share Collection Machines and Registration Machines. Collection Machines, Counting Machines, and M-Boxes are specific to a particular election, but Registration Machines and Private Voting Machines and their keys are shared across multiple elections.

The genesis of this E-Election System is simply the author's realization that tamper-proof software could be made and combined with the notions of thumb drives as machines and a virtual private network whose keys were distributed by hand, and not over the internet, to make a secure E-

Election System[3,4]

4 Advantages

The main advantages of the E-Election System are:

1. Substantially reduced election staff manpower.
2. Substantially decreased cost of running frequent elections.
3. Substantially increased ease of voting.
4. If the voter votes more than once, the last vote counts, as opposed to current election systems in which the first vote counts. Thus a voter is free to change their mind as more information becomes available.

Hopefully these advantages will eventually lead to greater voter turnout and more frequent use of elections such as referendums that let the voters decide issues directly (as is done in Switzerland).

5 Disadvantages

The main disadvantages of the E-Election System are:

1. Requires voters to have basic computer literacy.
2. Requires computer manufacturers to provide unhacked basic bootstrap.
3. Requires voters to report by election date if both their thumb drive and registration card are lost or stolen together.

[3]More specifically, the area of electronic elections is not a specialty of the author, the author has not read the work of others in this area, and the author makes no guarantee that others have not had this idea first. The author merely observed that a secure E-Election System is possible and produced this preliminary design as a proof-of-concept. The author does not intend to do any more work in this area.

[4]On the other hand, the author has been both a voter and an ordinary election worker in a Massachusetts suburb, and has drawn upon this experience to compare the E-Election System to the current Massachusetts system.

6 Software Distribution

Since the software is published and compilers for its tamper-proof language are to be readily available, it is easy to make deviant versions of the software.

Thus the binaries that can be trusted must be reliably distributed.

The software for one of the E-Election System machines will be a set of read-only files, one of which will be the boot file that runs during startup.[5]

At least three different and independent organizations should provide the E-Election files. Each organization starts with the same public source files. Each organization writes its own compiler for the tamper-proof software. Each organization publishes on its web site the source files and binary files, and cryptographic hashes that can be used to verify that the software on an E-Election machine matches that provided by the organization.

All the organizations should end up with identical files and cryptographic hashes. The cryptographic hashes (e.g., SHA-256) of software on the E-Election machine thumb drives can be checked by widely available commodity software running under commodity operating systems. Results can be compared with the hashes on all the organization web sites.

Only the cryptographic hash of the boot file needs to be checked, as it can contain cryptographic hashes of the other files and it will check the other files before they are used in running the machine. To avoid hacking by commodity software during a cryptographic hash check, the E-Election machine thumb drive should have a physical read-only switch.

The software files will be on the same thumb drive as their associated machine's locked box, and therefore cannot be substituted by switching thumb drives.

Private Voting Machines are created by the town clerk[6].

Software updates for the Private Voting Machines are distributed by the Registration Machines. Specifically, the Private Voting Machines contain in their locked boxes the public keys of the organizations which provide their software. Registration Machines can inform a Private Voting Machine when an update is available, and then the Private Voting Machine can get and compare three copies of the update from three different independent software providers obtained from three different Registration Machines.

At any time a voter can check their Private Voting Machine software file cryptographic hashes, but

[5] For example, all the software but device drivers may be in the boot file, with each device driver being in a separate file. With this scheme no file need be more than several megabytes in size, and files can be requested individually if only some have been upgraded.

[6] Or other official in charge of local elections

this should not be necessary unless the voter suspects that someone has physically tampered with the machine's thumb drive.

7 Keys and Names

Each locked box in the system has a passkey, which is written on a registration card associated with the locked box. Passkeys are generated by the system, and not by people. Typical passkeys consist of 2 or 3 three-letter pseudo-words and 4 to 6 digits, chosen at random. An example is 'vox64bam45zip01'. A three-letter pseudo-word is equivalent to at least 3 digits, as there are over 1,000 pseudo-words to randomly choose from, so this example is the equivalent of 15 random digits, but the pseudo-words, if chosen to be pronounceable, tend to make the passkey easier to type.

Each machine has a locked box. Each precinct has 10 additional locked M-Boxes.

Each machine and each M-Box has a public key and a private key. These are long, e.g. 4,096 bit, values constructed randomly in public/private pairs. These are not the same as the passkey of any locked box. Passkeys are symmetric encryption keys.

Each machine and each M-Box has a Unique ID, or UID, which is a 1024 bit random number used to identify the machine or M-Box. This UID serves as the machine's or M-Box's name.

Each Registration and Collection Machine has a URL.

The correspondence between Private Voting Machine UIDs and voter names and addresses is not part of the E-Election system described here. The E-Election system only knows about Private Voting Machines identified by their UIDs. Generally the correspondence between UIDs and voter names and addresses would only be made available to organizations and people authorized by state law. For example, Massachusetts currently makes voter registration lists available to registered political parties.

When one machine sends a message to another, the sender encrypts the message proper with the sender's private key, and attaches the unencrypted sender's UID to the encrypted message. Then the sender encrypts this whole message with the receiver's public key. The sender must have the receiver's public key, and the receiver must have the sender's public key. Public keys are not published, though the system is designed so that publishing public keys will not compromise security or essential privacy, and a voter can find out the public keys of Registration Machines and Collection Machines by examining his/her locked box using deviant software (see page 5) to decrypt and display the locked box.

Note that Counting Machines are never given Private Voting Machine UIDs or public keys, and can

never associate an encrypted ballot with a particular Private Voting Machine. And only Counting Machines are given decrypted M-Boxes, so only they can decrypt ballots.

8 Registration Machine Details

A Registration Machine is run by a single election official who holds the registration card containing the the passkey to the machine's locked box. The Registration Machine in this sense is specific to a particular election, but the URL of the Registration Machine and the machine's public/private key pair and UID will typically remain the same across several elections.

The locked box of a Registration Machine contains, for the current election:

1. The public/private key pair of the Registration Machine.
2. Public keys and UIDs of all Registration Machines, all Collection Machines, and all M-Boxes.
3. URLs of all Registration Machines and all Collection Machines.
4. The unmarked, unencrypted ballot.
5. UIDs and associated public keys of all Private Voting Machines. This includes de-registered and test Private Voting Machines, as Registration Machines do not know which Private Voting Machines are de-registered or test machines.

   During the election the Registration Machines will receive updates that add to this list of UIDs and public keys of Private Voting Machines.

   Registration Machines do not know the names and addresses that correspond to Private Voting Machine UIDs. To a Registration Machine, a voter is just a Private Voting Machine UID.
6. A log of any transactions the Registration Machine has been involved in, including Certificates issued by the Registration Machine and IP addresses used by the Private Voting Machines to request transactions.

   During the election the portion of the log involving transactions with a particular Private Voting Machine is visible to that Private Voting Machine. After the election the entire log will be available to the public under the same terms as voter registration lists are available (e.g., in Massachusetts they are available to registered political parties).

7. Information about Private Voting Machine Software, including complete copies of recent versions, and version numbers and cryptographic hashes of both recent and deprecated versions.

8. UIDs, URLs, and public keys of Registration Machines that will be used in future elections and Registration Machines that were used in previous elections and have been removed from service.

A Registration Machine receives four kinds of requests from Private Voting Machines:

1. Request for Ballot. The reply includes:

   (a) The unmarked ballot.

   (b) Public keys, UIDs, and URLs of Registration Machines and Collection Machines being used in the election.

   (c) Public keys and UIDs of M-Boxes being used in the election.

   (d) Version numbers and cryptographic hashes of Private Voting Machine software known to the Registration Machine, and the status of each version, active or deprecated.

   (e) UIDs, URLs, public keys, and status of Registration Machines not being used in the current election because they have been removed from service or are to be used in a future election.

   (f) If requested, a long truly random bit string for use when Private Voting Machine hardware does not have its own hardware random number generator (most recent computer processors have such). If this has to be used, it will be exclusively OR'ed with the random bit strings of at least 2 other Registration Machines so that no one Registration Machine will know the random numbers being used.

2. Request for Certificate. The request includes the cryptographic hash of an encrypted ballot prepared by the requesting Private Voting Machine. The reply is a certificate containing the cryptographic hash of the encrypted ballot, the Registration Machine and Private Voting Machine UIDs, the date and time according to the Registration Machine, the IP address of requester, and a digital signature of the certificate by the Registration Machine (see page 12).

3. Request for Log. The log entries created by requests from the requesting Private Voting Machine are returned in the reply.

4. Request for Software. The version of Private Voting Machine software requested is returned in the reply.

During voting, Registration Machines communicate only with Private Voting Machines, which make requests of the Registration Machines and receive replies. All requests are logged.

A Registration Machine may serve requests from more than one precinct, but the election information flow remains specific to the precinct. Each precinct should have at least 3 Registration Machines available.

Note that Private Voting Machines have no concept of time. If they cannot contact Registration Machines used in the last election, they will try to contact Registration Machines slated for use in future elections, and if successful, will update their locked box according to the Request for Ballot replies of these Registration Machines.

When a PVM makes a request of a Registration Machine for a particular election, the request must contain a copy of the last certificate received by the PVM for that election from that Registration Machine, if any. If it does not, requests for information will be allowed, with a warning message, but requests for a new certificate will be dis-allowed. Furthermore, in order to make a ballot official, a PVM must get a quorum (over half) of all Registration Machines to certify the ballot.

Thus if a PVM is duplicated and then one copy votes and receives certificates from a quorum of Registration Machines, the other copy will be notified when it attempts to access any Registration Machine in that quorum. This other copy will not be able to change the vote, because it would need a quorum of Registration Machines to certify the change, and that is not possible given the rules of the last paragraph.

Of course, if this situation is not in the voter's interest, the voter may get the registrar to de-register the PVM and give the voter a new PVM. Voters may make copies of their PVM without issues as long as they stick to just one of these copies for each separate election.

9 Private Voting Machine Details

A Private Voting Machine is run by a single voter who holds the registration card containing the passkey to the machine's locked box. Usually a Private Voting Machine is used by its voter for multiple elections over a period of several years, after which the Private Voting Machine is de-registered and a new Private Voting Machine is given to the voter.

A voter gets her/his Private Voting Machine as part of voter registration at the Town Clerk's Office. If a Private Voting Machine (i.e., its thumb drive and/or registration card) is lost or stolen, the voter may have the machine de-registered and replaced by a new Private Voting Machine.

Initially the locked box of a Private Voting Machine contains:

1. The UID and public/private key pair of the Private Voting Machine.
2. Public keys, UIDs, URLs, and current status of some Registration Machines.
3. Version number and cryptographic hash of Private Voting Machine's current software.

The Private Voting Machine makes requests of Registration Machines. In every case at least 3 Registration Machines should be used and the results compared to avoid single point failures or corruption.

The first thing a Private Voting Machine does, when prompted by its voter, is make a Request for Ballot and use the replies to add to its locked box the following information:

4. URLs, public keys, and UIDs of all Registration Machines and all Collection Machines used in the current election.
5. Public keys and UIDs of the M-Boxes used in the current election.
6. The unmarked ballot.
7. Public keys, UIDs, URLs, and status of Registration Machines that are obsolete or will be used in future elections.
8. Version numbers and cryptographic hashes of different Private Voting Machine software versions known to the Registration Machines.
9. The exclusive OR of the random bit strings provided by the Registration Machines, if this is needed.
10. A log of all transactions between the Private Voting Machine and Registration or Collection Machines for this election.

This can be done days or just minutes before the voter votes.

At this point the voter may or may not choose to update Private Voting Machine software using a Request for Software made to the Registration Machines.

The next step is for the voter to vote by marking the ballot.

After it is marked the ballot is encrypted and certified. Encryption of the ballot is described below. To certify the ballot, a cryptographic hash is computed of the encrypted ballot, and the cryptographic hash is sent in a Request to Certify message to each of at least three Registration Machines. These in turn send back certificates that are appended to the encrypted ballot to form the submitted

ballot. Each certificate contains the cryptographic hash of the encrypted ballot, the UIDs of the Private Voting Machine and Registration Machine, and a time stamp. The certificates are signed by the Registration Machines but are not encrypted. The certificates are copied to Registration Machine and Private Voting Machine logs.

The submitted ballot is now sent by the Private Voting Machine to at least three Collection Machines. The submitted ballot contains:

1. The encrypted ballot.
2. Certificates received for the encrypted ballot. Each contains the UIDs of the Private Voting Machine and certifying Registration Machine.
3. A copy of the unencrypted unmarked ballot, for diagnostic purposes.
4. Version number of the Private Voting Machine software, for diagnostic purposes.
5. Digital signature of the submitted ballot, computed by the Private Voting Machine using its private key.

A voter may change her/his mind and re-vote. The last ballot he/she casts counts. This may be important in an election, as on the one hand a voter may receive new information and want to change their vote, and on the other hand to avoid system congestion it is convenient to have votes strung out over time.

10 Ballot Encryption Details

The ballot is encrypted as follows.

First, the ballot is embedded in a long random bit string. This is done because in a typical election most marked ballots will have only a few possible values. For example, if there are only 6 contested positions in an election with only 2 candidates each, there would only be 64 possible marked ballots, exclusive of write-ins.

One method would be to generate a long random bit string and then use the first part of this as instructions to copy the ballot voting bits into various random places in the remainder of the random bit string.[7]

[7] This is one of many ways of 'adding entropy' to the message.

Second, a random symmetric key S is generated and used to encrypt the randomized ballot. Then Shamir's 5-Way Secret Sharing encryption scheme is used to generate 10 more keys K1, ..., K10 that encode S. The 5-way scheme is such that any 5 of the 10 K-keys can be used to compute S, but fewer then 5 of the K-keys gives no information about S (the K-keys are not keys in the normal sense of encryption but are technically called 'shares' of the secret S). Then each of the 10 K-keys is encrypted by the public key of a different one of the 10 M-boxes.

The encrypted ballot consists of:

1. The random bit string containing the embedded ballot, encrypted with the S key.
2. The 10 K-keys, each encrypted with a different one of the 10 M-Box public keys. Each encrypted K-key is paired with the unencrypted UID of the M-Box whose public key encrypts the K-key.

The encrypted ballot can be decrypted by any 5 M-box private keys. It does not contain the voter's UID.

11 Certificate Details

For each encrypted ballot, a certificate is generated by each of several Registration Machines acting independently. Each certificate contains:

1. UID of the Private Voting Machine requesting the certificate.
2. UID of the Registration Machine certifying the ballot.
3. Date and time the certificate is issued, as computed by the Registration Machine.
4. IP address used by the Private Voting Machine to make the request for certificate.
5. Cryptographic hash of the encrypted ballot as computed by the Private Voting Machine. The encrypted ballot itself is never sent to the Registration Machine.
6. Digital signature of the certificate computed by the Registration Machine using its private key.

12 Collection Machine Details

The Collection Machines receive the submitted ballots, each containing an encrypted ballot and at least 3 certificates for the encrypted ballot.

The locked box of a Collection Machine initially contains, for the current election:

1. The public/private key pair of the Collection Machine.
2. UIDs and public keys of all Registration Machines.
3. The unmarked, unencrypted ballot for diagnostic purposes.
4. UIDs and associated public keys of all Private Voting Machines. This includes de-registered and test Private Voting Machines, as before the election ends the Collection Machines do not know which Private Voting Machines are de-registered or test machines.

   During the election the Collection Machines will receive updates that add to this list of UIDs and public keys of Private Voting Machines.

During voting the collection machines will receive submitted ballots. These cause the following to be added to the Collection Machine locked box:

5. The set of all accepted submitted ballots sent to the collection machine by Private Voting Machines.
6. A log of all transactions between the Collection Machine and Private Voting Machines, including the IP addresses used by the Private Voting Machines to request transactions.

During voting, Collection Machines communicate only with Private Voting Machines, which make requests of the Collection Machines and receive replies. Collection Machines process two kinds of requests from Private Voting Machines:

1. Ballot Submission. The Private Voting Machine uses this to send its submitted ballot to the Collection Machine. The latter accepts the submitted ballot if its digital certificate signatures check, its certificate cryptographic ballot hashes check, and the copy of the unmarked ballot checks. Otherwise the Collection Machine rejects the submitted ballot. Note that a message

not correctly encrypted with the private key of the submitting Private Voting Machine and the public key of the receiving Collection Machine is completely ignored.[8]

Accepted submitted ballots are added to the locked box of the Collection Machine.

2. Request for Log. The log entries created by requests from the requesting Private Voting Machine are returned in the reply. This includes the certificates of accepted submitted ballots that were sent to the Collection Machine by the Private Voting Machine.

Note that messages received that cannot be decrypted using the private key of the Collection Machine are discarded, and so cannot contain ballots that are either valid or invalid. Collection Machine public/private key pairs are election specific, so ballots for the wrong election will not be considered.

At the end of voting Collection Machines are disconnected from the internet. They then

1. Receive the UIDs and public keys of all Collection Machines.
2. Transmit to each other their sets of accepted submitted ballots.
3. Merge into their own set of accepted submitted ballots any received from other Collection Machines that they did not have previously. At this point, the Collection Machines all have the same set of accepted submitted ballots.

Lastly, Collection Machines participate with Counting Machines to produce vote counts. More than one vote count will be produced: e.g., a count of initial valid votes, a re-count that includes allowed provisional votes, a re-count that includes test votes. The Collection Machine's job is to prepare for each vote count the set of valid ballots and submit the encrypted ballots in this set to the Counting Machines. To do this the Collection Machines:

4. Receive from outside the E-Election System a list of UIDs of Private Voting Machines that are valid for the vote count. E.g., for the initial vote count a Private Voting Machine would be valid unless it is de-registered, provisional, or a test machine.
5. Create a set of valid ballots (defined below) whose encrypted ballot parts are to be submitted to the Counting Machines.

[8]However, if the message is correctly encrypted with the public key of the receiving Collection Machine, the IP address of the message and the message itself may be recorded for forensic use. But the security of the E-Election system cannot be compromised by such messages.

6. Separated this set of valid ballots into a set of ballot headers and a set of encrypted ballots. The header of a submitted ballot is everything but the encrypted ballot contained therein. In particular the certificates are in the header.

7. Receive the UIDs and public keys of all Counting Machines.

8. Transfer the set of encrypted ballots to each Counting Machine. Only the encrypted ballots, and not the ballot headers, are transferred.

Since the Collection Machines are disconnected from the internet after the election ends, and Counting Machines are never connected to the internet, data transfers are made by copying to and from thumb drives, or by a network physically separate from the internet.

A ballot is valid unless:

1. It is from a Private Voting Machine that is not valid for this count (e.g., it is from a test machine and test machines are not being included in the current count).

2. It has been superseded by a later ballot from the same Private Voting Machine.

Information specific to a particular vote count, e.g., the list of valid Private Voting Machines and set of valid ballots, is not added to the Collection Machine locked box, so that a Collection Machine has no memory of past vote counts. This is an aid to testing (see page 16).

13 Counting Machine Details

Counting is done separately for each vote count.

During counting, the Counting Machines may communicate with each other using a network that is physically disconnected from any other network. The information they receive over this network is used to check the counting process, but is not used by any Counting Machine to modify its counting process.

Each Counting Machine receives a complete sorted set of encrypted ballots from one of the Collection Machines (assuming there are the same number of Counting Machines as Collection Machines, and using an arbitrary 1-1 correspondence). Using cryptographic hashes of the sets, the Counting Machines check that each Counting Machine has received the same set of encrypted ballots.

Each Counting Machine is then given 5 M-boxes with passkeys, and the Counting Machine proceeds to decrypt the M-boxes and decrypt and count the ballots. The Counting Machines then use cryptographic hashes to check that all Counting Machines produced the same counts.

An M-box contains nothing but its UID and public/private key pair.

The Counting Machine locked box contains nothing but its UID and public/private key pair and the UIDs and public keys of all Collection Machines and all other Counting Machines. It does not contain any information specific to a vote count, as an aid to testing (see page 16).

Write-ins can be tabulated by the Counting Machines in a manner generally similar to regular votes. A write-in is a piece of text, and counts can be given each piece of text and for groups of similar pieces of text. Normally all write-in votes can be lumped together as there are not enough to change the outcome of an election regardless of which candidates they name.

14 Testing

All election machines are in triplicate copies which check up on each other, except the Private Voting Machines. Therefore testing for security and other problems is centered on the Private Voting Machines.

The basic E-Election System philosophy of testing is to run tests that appear to hardware just like real election runs. Then if some hardware is hacked and corrupts the result, the corruption will be found out. For this to work, it is important that the hardware not be able to distinguish a test run from a real election run.

Some Private Voting Machines are designated as ‘test machines’. These are used by election officials to cast known votes that can be checked to ensure that the system and hardware is working as desired. These test votes are cast during the normal election period, using the same Registration and Collection Machines as official votes, and in fact the Registration and Collection Machines do not know the difference between a test vote and a normal vote.

To check test votes, a re-count is done with both valid Private Voting Machines and Test Private Voting Machines presented as valid to the Collection Machines. The difference between the official count and this re-count should be exactly the votes from the test machines.

None of the E-Election Computers will be able to distinguish test voting or test counting from real election voting and counting. In particular, the Collection and Counting Machines cannot use the number of valid Private Voting Machines in a count as an indicator of the presence of test votes, as the numbers will be similar in the official count and the test re-count, and these machines retain no records of previous vote counts in their locked boxes.

Test Private Voting Machines should be run on the same kinds of hardware that real voters use, so if computers manufactured by a particular vendor are hacked, the hack will be detected.

15 Vote Buying

No voting system prevents a person who knows a voter well from 'buying' the voter's vote with some favor.

The E-Election System has the property that a voter can vote multiple times, with only the last vote counting. If buyer X buys the vote of voter Y, Y may vote differently later, effectively deleting the bought vote. However after the election is over, election logs specifying when the voter voted are made public, so X can check then that Y did not vote later.

The E-Election system has somewhat different vulnerabilities than the current election system to a Vote Market Business[9](VMB) that buys votes and resells them. Such a business would be a foreign company accessed via the internet and not subject to prosecution. To understand this, we consider the current voting systems and the E-Election systems in more detail:

- **In Person Voting**: The voter who sells his/her vote agrees to simply not show up at the polls. The VMB provides a replacement voter. The replacement must be of the same age and sex as the selling voter, else the replacement can be detected by election workers during checkin and checkout. The replacement will be challenged for ID if the voter did not register at town hall or has not voted recently or if election officials have become suspicious of the selling voter. In this case, the replacement is liable to be arrested promptly, as there is a policeman on site at every polling place.

  It appears that no such scheme has been realized as of the current time. However, were such as scheme realized, some replacements would be jailed driving the price of a vote up to unsustainable levels.

- **US Mail Balloting**: The voter who sells his/her vote obtains a mail-in ballot legally, signs the secrecy envelop, and then mails the ballot and secrecy envelop to the VMB with the sold election contest unmarked. The VMB marks that contest, xeroxes everything so it can

[9]To my knowledge VMBs do not exist. One of the dangers of talking about VMBs is that they are likely to become the bogey of a non-existent-bogey-scam (NEBS). This is a political scam in which a bogey is targeted, such as Communists-in-government, weapons-of-mass-destruction-in-Iraq, pedophiles-in-government, voter-fraud, or VMBs. But the targeted bogey does not actually exist. A big advantage to the scammer is that negatives are hard to prove convincingly; one is left with lines like 'the FBI spends a lot of money tracking down pedophiles and has not been finding them in government'. Another advantage is that the bogey, being non-existent, will not counter-attack.

prove to the election purchaser that the ballot was 'delivered', seals the ballot in the secrecy envelop, and returns this to the seller who submits it to the local election authorities.

It appears that no such scheme has been realized as of the current time. However, were such as scheme realized, if very many voters sell their votes some will be discovered and prosecuted and the cost of a vote will be driven upward to unsustainable levels.

- **E-Election System**: This is actually similar to US Mail Balloting with the internet replacing the US Mail.

  When the voter sells his/her vote, the voter installs a VMB provided version of deviant software (see page 5). The VMB uses this to cast the sold vote from the voter's computer.

  When the deviant software is installed, the voter's PVM thumb drive becomes incriminating evidence. The only way to fix this is for the VMB to clean up the thumb drive after casting the sold vote. If it does this before the end of the election, the voter can re-vote, or resell their vote to a different VMB, wiping out the sold vote. However the VMB will be able to check when the voter submitted a vote after the election is over, and so will know if the sold vote counted. Then the VMB cannot pay the vote seller until after the election.

  Like the US Mail Ballot situation, if very many voters sell their votes some will be discovered and prosecuted and the cost of a vote will be driven upward to unsustainable levels.

In all these scenarios the VMB will have extreme difficulties in advertising and in paying vote sellers. The only people likely to sell their votes are poor people, who will be ill equipped to deal with the VMB in a covert way, and will be found out. If there are local facilitators for the VMB, then as they must have many clients to make it worthwhile, they will likely be detected and prosecuted.

16 Voter Suppression

Attempts to suppress votes are more common than attempts to buy votes.

Because voters can check the logs of Registration and Collection Machines for summary information about the votes that the voter's Private Voting Machine has cast, using nothing but the Private Voting Machine itself, suppressing a significant number of votes without knowledge of the voters is likely to be detected.[10].

[10]In the current Massachusetts mail voting system, a voter can find out from a state web site if the Town Clerk has received the voter's application, has mailed the ballot, and has received the ballot.

Another way to suppress E-Election votes is to refuse registration of legitimate voters, but the E-Election system does not differ from the current election systems in this respect.

A last way to suppress votes is to pay voters to not vote. This could be done by a Voter Suppression Business located in a foreign country and not subject to prosecution. It is possible under both the E-Election system and the current voting systems for third parties to determine whether a voter voted in a given election, so the E-Election system is no more vulnerable to this scheme than the current election system.

17 Hacking, Hardware, and Profit

So how does one hack this e-election system?

The software, Registration Machines, Collection Machines, and Counting Machines are all in independent triplets, so hacking these without being detected is very difficult.

Loading deviant software into a PVM thumb drive is easy for the voter to detect: they just use standard software, widely available from multiple sources, to compute the cryptographic hash of the software's boot file, and compare it with the hashes published by the software organizations.

However, modern hardware contains a lot of embedded micro-code and BIOS code, so maybe this can be hacked.

You cannot hack the keyboard, as keyboard output is displayed on the screen and keyboard mistakes are immediately obvious. You cannot hack the net interface, as the e-election system does not trust the net and the only thing a hacked net can do is deny service, which is immediately obvious. You cannot hack the display computer,[11] for it has no access to an outside source of information and could not be adapted to a particular election or screen format after the display was manufactured.

This leaves the computer BIOS and the micro-code in the thumb drive itself.

Perhaps the only good chance one has of having an undetected, election effecting hack is to hack the BIOS of a popular computer so that when it is asked to boot PVM software it instead goes out over the web and gets a deviant copy. The deviant copy then maintains two locked boxes, one reflecting the voter's intent, and one reflecting the hacker's intent. To avoid detection, the thumb drive must be kept as it would be for the voter, so the hacker's locked box must be kept on a deviant web server.

The testing described in section 14 (page 16) will detect computers hacked in this manner.

[11] Yes, your display has a (very small) computer inside it, for it can present a menu and accept commands.

If some time after voting the voter manages to load non-deviant software, the voter might use it to check with the Collection Machines to see if a vote matching that in the voter's locked box was properly collected, and thereby discover the hack.

The micro-code inside the PVM thumb drive can be hacked in a manner similar to the computer's BIOS. As the thumb drives are purchased by the election officials, its likely that there will be fewer thumb drive vendors than there are computer vendors involved in an e-election. The thumb drive situation is similar to the current Massachusetts voting system, where the paper vote counting ballot boxes are purchased by election officials and might be hacked.

In any case hacks of this kind are likely to be detected. Of course finding that some popular computer or a thumb drive vendor has been hacked in this manner will be a scandal, but then finding that your paper election ballot boxes have been similarly hacked would be a similar scandal.

In the case of a hacked paper ballot election, one counts the ballots by hand. In the case of a hacked e-election, one simply repeats the vote with somewhat different hardware. It may be cheaper to repeat the vote than to count the election by hand, but it requires more participation by voters. In either case these events are likely to be very annoying but very rare.

Small computers big enough to run PVM software can be integrated into the thumb drive resulting in a voting unit with special shape and markings.[12] This eliminates issues with BIOS and computer vendors and produces the same situation as we now have with Massachusetts paper ballots, namely that there is computerized hardware purchased by election officials that can be hacked but the hacks can be detected.

In the long run all election hardware should be manufactured in an inspected facility by a consortium of manufacturers and governments, and should run only tamper-proof software.

So how can a private company profit from such election systems? In the current Massachusetts system, only the maker of paper ballot boxes profits from the election proper, while in the e-election system only the maker of the thumb drives or integrated voting units would profit from the election proper. Here we are ignoring ancillary actions, such as voter registration accounting, or in the current election system, voter checkin/checkout, which may involve hardware or software provided by private companies.

18 Election Preparation

There are many ways to prepare the various machines needed in an election, and the one we outline here is but an example.

[12] Probably costing at most $25.

The town holding the election has 3 preparation computers that are connected to each other by a local private network but which are not physically connected to the internet.[13] To detect hacked hardware produced by a single vendor, the 3 computers are from different computer vendors. Each has a good hardware random number generator (as is built into most modern processors).

Assume that the election is going to be carried out by 3 election officials, each responsible for 1 Registration Machine, 1 Collection Machine, and 1 Counting Machine. They each use a different one of the 3 preparation computers.

Registration Machines pre-exist the election. A new Registration Machine can be made from time to time by means similar to those described here. Assume each election official has one Registration Machine.

Initially the software for each election machine is copied from the internet to the machine's thumb drive, the cryptographic hash of its boot file is obtained by examining and crosschecking multiple independent public web pages, the thumb drive is set read-only with a physical switch and the cryptographic hash of its boot file is checked using commodity software. Then the machine is run and makes its own initial locked box containing its UID and public/private key pair and its own symmetric passkey that is copied to a paper card and used to encrypt the locked box.

The input to the preparation process consists of the pre-existing Registration Machines and three thumb drives:

1. A Registration Information Drive that contains the public keys, UIDs, URLs, and status of all Registration Machines.

2. An Election Information Drive that contains the the URLs of all Collection Machines and the unmarked ballot.

3. A Private Voting Information Drive that contains the public keys and UIDs of all Private Voting Machines.

The Election Information Drive is specific to the current election, but the other two Information Drives contain information that spans elections and is updated by addition and subtraction from one election to the next.

First the 3 election officials each make a Counting Machine, and each uses their Counting Machine to make 3 or 4 of the M-Boxes. There is no input to making these Counting Machines, with each generating the election specific random numbers that it needs.

[13] If there are 3 Registration Machines, a PVM would have to contact a quorum of 2 Registration Machines to certify a vote. If there were 4 or 5 Registration Machines, a quorum would be 3 Registration Machines.

Then the Counting Machines exchange their UIDs and public keys and M-Box UIDs and public keys over the local private net.

Lastly each Counting Machine writes one thumb drive for each of its M-boxes plus two more thumb drives:

1. An M-Box Information Drive that contains the public keys and UIDs of all the M-Boxes. This is input to the Registration Machines before the election.

2. A Counting Information Drive that contains the public keys and UIDs of all the Counting Machines. This is input after the election during counting to all Collection Machines and all Counting Machines.

The M-Box Information Drives written by different Counting Machines should all be identical, and similarly for the Counting Information Drives.

Then the 3 election officials each make a Collection Machine that takes as input the Registration and Election Information Drives and makes a Collection Information Drive containing the URLs, UIDs, and public keys of all the Collection Machines, and the unmarked ballot.

Then the 3 election officials each update their Registration Machine using as input the Registration, M-Box, and Collection Information Drives.

Note that Registration Machines never communicate with each other or with Collection Machines, except via the Registration and Collection Information Drives before the election starts.

Note that Collection Machines never communicate with each other before the end of the election except when making the Collection Information Drive before the election starts.

At any time the Collection Machines and Registration Machines may input the Private Voting Information Drive to update their locked boxes.

Registration and Collection Machines may each publish on its web page a cryptographic hash of its list of Private Voting Machines. When all machines are up-to-date these hashes should all be the same, and this sameness can be checked by the public.

In between elections the preparation machines can be used to create Private Voting Machines that are not yet assigned to a voter, and record their UIDs and public keys in a separate data base. Each of these can then be given to a voter and its UID can be recorded in the separate database that records the name, address, and birthday of the voter.

Preparation machines can also be used to maintain the data base of Private Voting Machine UIDs and public keys, the voter data base, and the data base of Registration Machine URLs and public/private keys. These data bases can have triplicate copies, each maintained by one of the prepa-

ration machines. These data bases should be log based: each has an underlying log of all changes that is only added to and never otherwise modified, and the database can be remade at any point by executing the operations in the log. From these data bases the Private Voting Machine Information Drives and drives containing lists of valid Private Voting Machine UIDs for each type of count can be prepared.

Note that most information is specific to a particular election and a particular precinct. The maximum size of a precinct in the United States is currently less than 3,000 voters, so there is no need to provide for large data bases.

Election officials only need to know the information above and be able to follow a check list that tells when to re-boot machine thumb drives, when to create a new writable thumb drive with a particular label, when to insert a thumb drive with a particular label as read-only or writable, and when to remove a thumb drive with a particular label. A small amount of information may be input when the programs run, e.g., the current time.

19 Other Considerations

There are other considerations that need to be taken into account when producing an E-Election System. Some of these, with suggestions for dealing with them, are as follows.

19.1 Thumb Drive Identification

Although any thumb drives can be used, preferably ones with a physical read-only switch, it would be better to have special thumb drives with distinctive markings or even a larger than usual size. This would make it less likely that a voter's PVM would be lost among the voter's other non-election thumb drives.

19.2 Denial of Service

Standard methods need to be employed to prevent Registration and Collection Machines from being overloaded by spurious internet packets. However as the voting period will be several weeks, such overloading is unlikely to be more than a nuisance.

To prevent overloading the memory used to hold the locked boxes, Registration Machines and Collection Machines need to limit the number of requests they receive from a particular Private

Voting Machine. If these limits need to be legitimately expanded for a particular voter, that voter may have his/her Private Voting Machine de-registered and get a new Private Voting Machine.

19.3 Meta-Data

ISPs will be able to collect meta-data that will allow them to make good guesses as to which IP addresses were used to cast votes. However, better information of this kind is made publicly available (e.g., to registered political parties) after the election anyway, and ISPs are unlikely to give this meta-data to anyone trying to guess voter IP addresses.

20 Some Thoughts on the Current Election System

The E-Election System gets its integrity by doing all critical functions using triplicate independent groups of people whose actions can be checked against each other. The current Massachusetts voting system gets some of its integrity by different means that are worth explicating here.

The Massachusetts voting system is based on paper ballots marked at a polling place (leaving aside mail-in for the moment) and counted by an electronic ballot box. The polling place is staffed during polling by 5 election workers per precinct, [14] two for checkin, two for checkout, and one to supervise the precinct's ballot box. Most election workers are in effect volunteers: they are officially town employees (and thus bound by regulations and background checks ensuring proper behavior), but they only earn $100-$400 per year.

Much of the security of the current system comes from the large number of election workers in a polling place who have a ring-side seat for watching each other. In general these workers are volunteers dedicated to making sure the system works as it is supposed to.

The checkin workers have a list of registered voters, in which they mark each voter that is given a ballot. The checkout workers have an identical list, in which they mark each voter that puts a ballot in the electronic ballot box. The result is two independent lists of who voted. Note the use of independent redundancy. Should, for example, a dead person 'vote', it would be easy to prove by checking these lists and the death certificate.

The lists also give counts of the number of votes for each precinct. The electronic ballot box produces its own count, which should match. After the election is over, the ballots are removed

[14] A precinct has at most 3,000 registered voters, and it is not unusual for 85% of these to vote in a contested election. Also note that election workers work in shifts, so 5 election workers per precinct are required for each shift, though many workers work several shifts.

from the ballot box and counted by hand, and this fourth count should match the others.

Each electronic ballot box prints two copies of its vote counts at the end of the election, and one of these is immediately posted in a public place (e.g., school building door).

If there were voting fraud it would be easy to prove, as there is a dense paper trail. If there were significant fraud, it would be easy to both find and prove. This is why claims of extensive fraud not accompanied by extensive proof are routinely rejected by courts.

What goes on inside the ballot boxes when they count the ballots is, however, not visible to anyone in the polling place. To ensure they are counting correctly, after each election some precincts should be selected at random, recounted by hand, and the hand and ballot box counts should be compared to see if they match. Lawyers representing the major candidates should monitor the recount. Each precinct is independent, so each can be counted separately by itself. It is very important that this be done, even though year after year will go by with no significant mis-matches.

In Massachusetts mail-in voting is done as follows. The voter obtains an application form (e.g., from the web) and submits it to the town clerk. It is signed. The clerk sends the voter a ballot and a privacy envelop. The voter marks the ballot, puts it into the privacy envelop which has the voter's name and address on it, signs and seals the privacy envelop, and returns that to the town clerk. The privacy envelops are tabulated and the registration lists at checkin/checkout have the voter marked as having voted by mail, so the voter cannot vote in person. At the beginning of the election proper, privacy envelops are slit but not opened, and are sorted to match the checkin/out registration list order. Then an election worker takes a batch of privacy envelops to checkin where they are checked in, takes them on to checkout where they are checked out, and then standing next to the ballot box opens each envelop and puts the ballot directly into the ballot box. The set of empty privacy envelops constitutes a list of who voted by mail, and this list has the voters' signatures. Again, if there were significant fraud, it would be easy to find and prove.

Also, in Massachusetts a voter can go online and check the status of her/his mail-in ballot. The voter can see if the town clerk has received the application, if the clerk has mailed the ballot, and if the clerk has received the privacy envelop containing the ballot.

The only likely way to commit voting fraud is to prevent people from voting at all. A good way to defend against this is to establish independent monitors that monitor the election and answer the following questions statistically:

1. How many legitimate voters were not permitted to register.

2. Were legitimate registered voters unable to vote. Importantly, were there excessively long voter lines due to polling place under-staffing. Additionally, are there registered voters who are <u>not</u> on the lists of people who actually voted who credibly claim to have submitted a

mail-in ballot.

3. Do the lists of people who voted contain names of people who can credibly be shown to have not voted. This may be hard to properly determine, as a few percent of potential voters may not remember accurately whether or not they voted. Video and audio recording of the checkin and checkout process might be used to resolve discrepancies.[15]

4. Have the ballot boxes been tested by re-counting a few randomly chosen precincts by hand and matching the hand count to the ballot box's count, with lawyers representing major candidates present.

[15]Computerized checkin and checkout have not been allowed in Massachusetts, because you never know what a computer is doing. However, video and audio recording might be synchronized with computerized checkin and checkout, allowing the latter, provided the recording system itself is tamper-proof.