

The Scantegrity System

An Introductory Whitepaper and Example

The Scantegrity system can be overlaid on any conventional optical-scan voting system. It aims to allow voter checking and public audit that provide confirmation of election results with the highest level of indisputability. It also aims to maintain and even enhance the degree of ballot secrecy achieved by the underlying scan system. These goals can be reached without requiring changes in whatever underlying conventional optical-scan voting system or its procedures and at very little additional cost.

This whitepaper is intended to provide a walkthrough of the Scantegrity system and some of its variations. It gives a more thorough, detailed and necessarily somewhat more technical description than that in the companion summary document or video.

Much of the description here is keyed to a single overall system example diagram (see page 11). The diagram is divided into three sections. In the upper left corner of the diagram are the six ballots used in the toy example shown as they would be printed. Below these ballots is a sequence of steps tracing a single one of the ballots through all potential voter interactions with it. On the right side of the diagram is a snapshot of the Scantegrity bulletin board for each election phase.

The Scantegrity part of an election proceeds in four phases: (1) pre-voting, (2) voting, (3) pre-audit, and (4) audit. (These phase numbers refer to those shown in the diagram.) The ballots of the conventional scan system are printed preferably after the Scantegrity pre-voting phase, since posting well in advance of audit gives more opportunity for others to record the data and thereby enhances the effectiveness of the commitment. The printing on a ballot includes the serial number and the letters committed to by the Scantegrity software during the pre-voting phase. The voting phase is

common to both Scantegrity and the legacy system. In some cases a single scan, whether legacy “mark sense” or standard “pixel-based” scan, of the ballot is made by the legacy system and the positions marked or pixel images are later fed to the Scantegrity software. Other options include a batch scan, after the legacy scan, to provide images for processing by the Scantegrity software. After the voting phase and announcement of the election results, audit of the results includes the Scantegrity pre-audit and audit phases.

The bulletin board can be seen in its initial state in the upper right corner of the diagram. A first column, the leftmost, lists the serial number of each ballot. The last column is where the results will be posted and is filled with question marks at least until the results are announced. Each row of the results column does correspond to a single ballot. But so as not to compromise ballot secrecy, the order of rows of the results column is essentially random and independent of the serial number ordering.

The rest of the bulletin board is made up of two columns of envelopes. Each of these two columns has an envelope per row and includes a space in which a letter can be written to the left of its envelope. The first of the two columns of envelopes, on the left adjacent to the serial number column, is in the same order as the serial number column and the rows match up directly. The second column of envelopes, however, has row order that is independent both of the serial number column and the results column. It provides an extra level of indirection, which allows effective audits that can still preserve ballot secrecy.

Physical paper envelopes can be used as shown on the bulletin board to demonstrate and explain the system, such as in a classroom. In an actual election, instead of

paper envelopes, the words and numbers contained in envelopes are encrypted to form what are in effect “digital envelopes.” These encryptions are then posted in corresponding positions in a tabular structure of numbers, letters, and encryptions available to anyone on election websites. To in effect “open” such an encryption envelope, those running the election post the particular key that decrypts it. Because of the type of encryption that is preferably used, there is always exactly one key that decrypts any envelope. This key then reveals and proves exactly what must have been encrypted when the encrypted envelopes were posted during the pre-election phase.

(A related but somewhat more elaborate use of such bulletin boards and encrypted envelopes is well known and documented in the Punchscan voting system. See www.punchscan.org for more detailed background on such systems.)

Voter interaction—apart from marking the ballot just as in the conventional scan system, including any “write in” votes—entails up to three optional successive steps: (A) voter notes the letters marked and also detaches and retains the chit, (B) voter checks online that the letters are posted correctly under the serial number on the chit, and (C) voter conducts the resolution procedure with election officials in case the voter believes that letters are posted incorrectly. The process by which voters optionally spoil ballots during voting is reflected in the diagram under the phase (2) posting during voting of the bulletin board. The resolution procedure is shown using opaque envelopes with windows. Unlike the illustrative envelopes used to show the bulletin board operation, the envelopes used in the resolution procedure would always be actual physical envelopes.

All voters preferably obtain the detachable counterfoil chit before they leave their ballot either in a ballot box or insert it into a scanner that feeds into a ballot box or in a mailing envelope. It is, however, up to the voter as to whether they wish to make notes and again whether they later check them using telecommunications and, in the unlikely event of a discrepancy, ultimately whether they opt for a resolution procedure.

The next four sections walk through an election, phase by phase, following the example illustrated in the diagram. The section headings have the number-in-circle symbols that callout the four bulletin board

snapshots in the diagram. Some subsections similarly have in their title the letter-in-circle used to callout the voter audit steps. After these four sections, a final section details some variations.

❶ Pre-Voting Printing and Posting

Six example ballots are shown as they would be printed. The ballots have only a single contest of two candidates. (This is to keep the diagrams manageable, but natural extensions to more general ballots are explained after the full description of the example.) All ballots list the candidates in the same standard order, such would typically be used in a polling place and sometimes called a “ballot rotation” or “ballot style.” Each ballot is identified by its serial number, which is printed for readability in sans-serif Arabic numerals in the example. The same serial number is printed again directly below using instead a barcode symbology (so-called “3 of 9” in the example). The barcode spans across the perforation line so that it clearly appears the same on both physical parts, thereby visibly linking, through the device of what is visibly the same serial number, the counterfoil chit to the ballot proper.

Each ballot has one oval labeled with the letter “A” and another labeled “B.” Those ballots with “A” labeling the upper candidate and “B” the lower are called out in the diagram as “not swapped,” i.e. in alphabetical order; those ballots where the positions of the letters is interchanged are called out as “swapped.” A reason for the random choice of whether the labeling letters are swapped or not is that the letters labeling the ovals marked by voters will be posted next to the envelopes in the first column of the bulletin board and this column is in serial number order. So if there were no random swapping, the posted letters would reveal how each serial numbered ballot is voted. The light-blue printing on the ballots, and of course which are swapped, should not be revealed except to the voter receiving the particular ballot.

The way a letter that has been marked by a voter is processed to transform it into the corresponding letter of the results column differs depending on whether the ballot is printed swapped or not swapped. Those ballots that are printed swapped will have their posted letters swapped one more time in the processing they undergo leading to inclusion in the results column. This makes the resulting letter correctly correspond to the vote: “A” for the first candidate and “B” for the second.

Those ballots that are labeled “not swapped” will have their result letter appear as if there were no swapping in processing. There are two different ways, however, to process a letter as if there were no swapping in processing: passed straight through without any swapping or swapped twice in succession, one swap canceling the other. All four combinations, two label printing orders each combined with one of two possible processing sequences, are included in the illustration (as can be checked by the reader using the light gray arrows shown in the final bulletin board snapshot). Each of the four is ideally picked independently and with uniform probability of a quarter, just as the chance that a ballot is printed swapped or not swapped is also independently and uniformly a half.

Corresponding to each ballot that is printed are two “envelopes” that are posted. As mentioned, these can be thought of either as paper envelopes or as encryption envelopes, but they will be described as paper envelopes here for concreteness. When the envelopes are posted, their content is not revealed. It is relied on though that neither envelopes nor their content are changed before those envelopes selected are opened during spoiling or audit. Enforcing this “commitment” property is essential to the implementation of envelopes and efficacy of the audits that will be described is based on this.

The first envelope corresponding to a serial number will be posted in the left column of envelopes and in the same row as its serial number. The second envelope corresponding to a serial number is in a row that is essentially “random” or unpredictable. (The permutations are ideally chosen independently and uniformly from the set of all permutations.) Inside the envelope next to the serial number is written the row number that the second envelope appears in. Similarly, inside that envelope in the second column is a number that indicates which similarly independently “random” row will be used to post the corresponding vote result.

Every envelope also contains (in addition to the row number for the item in the next column, whether envelope or result) one of the words “same” or “differ. If the ballot has “B for the upper candidate, then the word “differ” appears exactly once among the two envelopes chained by row numbers in envelopes to that serial number. If “A” labels the top candidate, then the word “differ appears either exactly zero or exactly two times. As explained earlier, this ensures that voter

marks are translated to results correctly. The two steps, through the chained envelopes, instead of just one that would be enough to protect privacy, allow audits that don’t substantially compromise vote privacy.

② Voting

An example voting session by a voter is shown in which the voter receives the ballot with serial number “3403.” The voter then votes for the second candidate by filling the lower oval. After so marking, the voter detaches the serial number chit and keeps it.

In some examples, such as that illustrated, the letters are visible through the translucent ink of the markers provided for use by voters. In other examples, no matter what shape is filled with opaque marks or indicated by voter-completed arrows or the like, the letters can be printed adjacent to the mark position. A partly- or even fully-hidden letter can, however, be readily inferred by voters from the other letter.

In some settings ballots are scanned in at the polling place, typically before the voter leaves, in so-called “precinct scan.” In other settings ballots cast at or delivered to polling places are scanned only after the ballot box containing them arrives at a satellite or central processing site. So-called “early voting,” “absentee” and “vote-by-mail” ballots are scanned in some jurisdictions as received and in others in a batch on, just before, or just after election day.

There are two types of scanning: mark-sense and pixel-based. Mark-sense only reads voter marks in fixed column positions and in row positions indicated by special bars printed along the sides of the form. It cannot read letters or numbers directly. Pixel scanning can of course be used to recognize vote position marks as well as stray marks, letters and numbers, and uses little or no extra printing for alignment.

In the most challenging pure retrofit setting, only a single mark-sense scan is made. Serial numbers, however, are then encoded as marks printed along with the Scantegrity letters, an option supported by current Scantegrity software. For instance, each digit of the serial number is printed encoded as a black mark filling an oval in a separate series of ovals, one for each possible value zero through nine of that digit. Data indicating the marked positions, including those pre-printed with the serial number, can then be supplied

from the legacy system to the Scantegrity system and are sufficient for its operation. The position in storage of the relevant ballots would presumably also be implied by the position of the records within the overall collection of data supplied.

Retrofit settings, in some cases, allow a second scan, using a production pixel-type scanner once the ballots are centralized. This obviates the need for Scantegrity to obtain data from legacy systems altogether. If the marked positions are recovered from the legacy system, a redundant read of the positions from the mark-sense readers are also obtained. For consistency in reported totals, the mark recognition decisions of the mark-sense scanner could be honored. However, gross inconsistency presumably should be cause for concern.

A non-retrofit setting can be conducted completely from a single pixel-based scan. This scan image can be obtained by a voting system and then shared with an otherwise separate Scantegrity audit system. The current Scantegrity software can alternatively provide the complete functionality of an election system.

Any pixel-scan contains both the voter marks and the corresponding letter labels, as well as the serial number from the portion of the barcode printing that extended over the perforation (unless invisible ink is used and appropriate lighting is not supplied, as mentioned in one extension below). Marks and their labeling letters are recognized from the same scan image, so that the number of marks and letters must match. There can be little doubt of the position a recognized mark is in on the ballot, but the so-called “OCR” (Optical Character Recognition) could conceivably error in recognizing the letter printed labeling the position marked. A double-check can be provided automatically between the OCR and serial-number barcode reading by consulting the data that controlled the printing of letters. Once the letters are recognized and checked they are posted on the bulletin board next to the envelopes in the first column, each in the row corresponding to its respective ballot serial number.

Voter Checking

Voter checking is something each voter should preferably be able to choose to do if they wish. Three levels of voter checking are shown: (A) voters keeping notes that voters made of the letters they marked, (B) voters checking online that the letters they noted are

posted correctly, and (C) an in-person resolution procedure. The first level of voter checking (A) is optional, but only if a voter performs that level does the second level (B) become a meaningful option for that voter. The third level (C) of checking makes sense only as an option for a voter who believes he or she has found an inconsistency while performing second level (B) checking.

The first level (A) for a voter is carried out within the voting phase (2) described in the current section; the subsequent step (B) and (C) are preferably carried out as soon as is practical after the letters are posted. In some examples the letters can be posted during the voting phase, such as during a vote-by-mail interval. In other examples the letters are posted only after the close of polls.

Ⓐ *Voter Notes the Letters*

Any voter should be able to make a note of the letter labeling the position that voter marked. In the example shown, the letter labeling the oval marked is the letter “A.” In the diagram, this letter is highlighted by a light-orange arrow in the various related places it occurs. The handwritten note of the letter is shown including along with the chit as being retained by the voter after voting. Some voters, such as those with visual disabilities, may make notes by means of their own audio recorder devices or for example by such features included in portable phones. Many voters may of course not choose to make notes. It is assumed, however, that those running the election are not able to readily determine which voters make mental or physical notes and will check.

Ⓑ *Voter Checks the Posting*

A voter wishing to check enters, through a browser or phone, the serial number from the chit he or she should have kept. The letters next to candidates marked, in this case the letter “A” for the example ballot “3403,” are rendered on the browser screen, audible on a self-vocalizing web page, or spoken by a recorded voice on a so-called “IVR” (Interactive Voice Response) 1-800 number system. The voter then checks that this letter matches that which the voter made a note of as described in the previous subsection.

The bulletin board is shown in snapshot (2) as updated from the scanned information obtained in the voting phase, either during the voting or after the close of polls. The letter “A” or “B” is shown posted corresponding to each serial number (apart from the spoiled ballot “3404” described later). If posted correctly, the letter should have labeled a voter mark on the ballot with that serial number. These same letters are served up by websites and audibly read out when the serial number is entered.

For maximum integrity, the consistency of the letters provided to voters are checked with the same data in bulk form (such as in XML format in the current software) that is also published online and authenticated by public key digital signature. One way to reduce the effort required for such maximum integrity checking is for each political party or candidate and even other entities to independently provide the web or IVR interface from the signed underlying bulk data they receive from the election officials.

Ⓢ *In-Person Voter Check of Ballot*

In the exceptional case where the voter believes that what was provided by the interactive system does not match the letter noted, an in-person dispute resolution can be conducted. All the voter should need to make such a check is the chit from the ballot and an idea of which letter should have been printed.

Such proceedings should be quite rare. Of course there will presumably always be some few people who for whatever reasons may wish to initiate such proceedings. Their number and maybe even which persons may become predictable through a series of elections. A small fee or other obstacle might reduce the number of nuisance procedures. In the event an election has exceptionally many disputed postings, for whatever reason, a significantly larger number of persons may wish to instigate proceedings. In such an eventuality, it may make sense to limit the number of proceedings, such as by a random sampling of applicants and/or assigning quotas to each political party or candidate and having them try to determine the most credible examples from those that are brought to their attention. Current practice typically entails election results being “certified” by a deadline expressed in days and such proceedings need only occur within that canvassing period.

The procedure will be described as if voters appear in person and only election officials are present. However, the voter could provide the chit and the letter in question to a friend, an attorney, a public-service organization, the media, a political party, or preferably the candidate or promoter of the ballot question. More generally various combination of such persons or groups may wish to be present for the resolution procedure and the proceedings could even be recorded on video by news media or others. Accordingly, how the voter voted should of course not be revealed.

The procedure begins with the voter providing the serial number from the chit. Then election officials in charge of the ballots use it to locate the corresponding ballot, first by finding the bundle, box, envelope or other group of ballots the database indicates that ballot was in during scanning. The database presumably also keeps various information about who processed the ballot, where the bundle is stored, and the like. The database is preferably developed when the ballots are scanned or re-scanned as already mentioned. The election official can simply count down to the corresponding ballot or run the bundle through a scanner programmed to kick out the particular number or search for it by leafing through sheets and using a barcode reader. Even these officials may be prevented from inspecting full ballots during such a search if observed, such as by a security camera in a storage facility or those present at the procedure.

Having located the ballot, the official places it in a special envelope that is opaque except for a cutout window that exposes the serial number part shown extending over the perforation line. The voter can then, without any votes being exposed, match the chit up with the ballot at the perforation line where they were separated. Such torn or cut paper matching can even be verified by known forensic techniques that rely on the pattern of fiber visible making up the paper surface. Such techniques are related to a system variation discussed later.

The second and final part of the voter’s in-person check of a ballot is aimed at verifying that the letter supplied by the system does in fact match the one labeling the mark on the physical ballot. In order to accomplish this, those in charge of the ballots transfer the selected ballot to an envelope that exposes both ovals as well as the letters labeling them. Even the candidate names can be exposed, but not the serial number. Those observing

can see that the original serial-number-verified ballot is transferred, without yet being able to see its mark. For example, making the transfer while the envelopes and ballot are face down on a table allows those present to see that a single piece of paper was in the one envelope. Another ballot is constructed, or borrowed from the actual ballot store, whose printing is, in this example, of the “not swapped” type, but that has a voter mark next to the same letter, “A,” and was thus voted for the other candidate. One simple finesse would be that if the election officials were to notice that the way the oval was filled was somehow distinctive, so that it might be recognized, they could try to make the mark on the other ballot substantially similar.

The two envelopes are then shuffled in a way that creates no doubt that the same two remain but effectively hides which is which. An example device to facilitate such shuffling is a cylindrical box, like a large hatbox or drum case, with a so-called “lazy Susan” type of rotating platform inside on the bottom, to which multiple vertical partitions are attached. It allows envelopes placed inside to be spun to essentially unpredictable locations but also for the box to be opened for complete inspection. Thus it can readily be established that the ballot must have had the particular letter next to the position marked because all ballots of the collection that includes that ballot have that same letter marked. Since multiple such positions are included among those displayed, however, which candidate the mark on the ballot with known serial number corresponds to is not revealed.

It might be more convenient to use a single type of envelope that has two windows and flaps that cover each, so that the ballot does not have to be moved between envelopes. A large window on the back of the envelope allows verification that it contains only a single integral piece of paper. Another example way to conduct such audits relies on letters being unique per contest (as in this example) and uses a picture frame that has a small hole exposing only the mark and letter.

Spoilt-Ballot Audit to Check Printing

A special check by voters is to establish that ballots were properly printed. This is shown only in snapshot (2) of the voting phase in order to simplify the subsequent illustrations and their description, though in practice it could be copied into the snapshots of the remaining two phases and can be conducted at any

point after phase (1) when the envelopes are committed.

In the example illustrated, a voter has spoilt ballot number “3404” and both envelopes on the bulletin board corresponding to it are opened. Everyone can then see that the form was printed “swapped,” that is “B” above “A,” and then verify that there is exactly one “differ” in the envelope chain for that serial number. If the form had been printed “not-swapped,” which case is not shown, then the chain would be verified to have either both “same” or both “differ.”

More specifically, the envelope in the first column labeled by the ballot serial number is opened and the row number, which happens to be “1” in the example, is thereby revealed. The row number “points to” the envelope in the first row of the second table of envelopes. The envelope so pointed to should then also be opened. Once both envelopes are opened they reveal whether the ballot should have been printed with letter order swapped or not swapped. Thus, if both envelopes contain the word “same” or both contain “differ,” then the ballot should have been printed not swapped, with letter “A” above “B.” But if either one envelope has the word “same” and the other the word “differ,” then the ballot should have been printed in a swapped order, “B” above “A.”

There are several other ways that voters can in effect make spot-check audits or more systematically audit that the letters printed on the forms do in fact correspond to what is in effect committed to by the information that is contained in the envelopes posted. For example, in some settings each voter can be provided with two forms and the voter chooses which to vote and which to spoil. In other examples, voters take ballots randomly from the stack of ballots at the polling place and all the ballots left unvoted are spoilt as observed by those remaining in the polling place at the time. As yet another example, a coin is flipped or die thrown to determine whether the ballot provided the voter should be voted or spoilt. For vote-by-mail voters, an option exists to spoil the ballot received through the mails and then optionally either request another or cast a provisional ballot at a polling place.

Exactly which ballots would not be voted was presumably unpredictable when the ballots were printed and distributed. Also, the envelopes are opened in place through a publicly-observable process

designed to make sure that their content cannot be changed. Thus, the spoilt ballot audit effectively establishes that the ballots are printed according to the secrets committed to when the bulletin board was posted. It also establishes with high probability that the pair of chained envelopes corresponding to each ballot serial number contains the proper number of swaps according to whether the ballot printing is swapped or not swapped.

③ Pre-Audit Posting of Results

The result of the present pre-audit phase (3) sets the stage for conducting the full public audit in the final phase (4). The main posting is of course the election results. A further posting, which will be needed in the final audit, is the second column of letters for all voted ballots.

The pre-audit values posted are determined by the Scantegrity system. This is done using knowledge of what is inside the unopened encryption envelopes. Thus the system is able to in effect trace a letter posted next to the envelope in the first column through that first envelope, which indicates whether or not it should be swapped in processing (that is changed “A” for “B” and vice versa) and also determines the row number in the second column of envelopes that the resulting letter should be posted on. This yields the letters of the second column that should be posted next to the corresponding second envelopes for all voted ballots.

The pre-audit posting is completed by repeating essentially this same process using as input, instead of the letters from the first column of envelopes, those letters that have just been posted for the second column of envelopes. The letters yielded by this process are placed in rows of the results column as called for by the row numbers in the envelopes of the second column. These resulting letters should be free of any swaps: any swaps in printing having been cancelled through the swaps forced by the envelopes (because there are an even number of swaps in total for each ballot when including its printing and chain of two envelopes). Thus, these letters indicate the candidates voted for. The letter “A” in the results column thus indicates a vote for the first candidate in the standard order in which the candidate names are printed on all the ballots and the letter “B” a vote for the second candidate.

④ Auditing the Bulletin Board Results

In the audit phase, which is the fourth and final phase of the Scantegrity system, certain envelopes on the bulletin board are selected in a public but unpredictable way and are opened for inspection. (The gray dotted arrows mentioned earlier are not made public.) The objective is to establish that the results already posted in the last column do in fact correspond, through the chaining of envelopes by the pointers they contain, to the letters already posted next to the first column of envelopes. Since these letters next to the envelopes of the first column have been given the opportunity to be vetted by voter checking, the posted results of the last column are in effect checked all the way back to what the voters saw when voting. The efficacy of the audit is further dependent on the ballots having been printed with letters and serial numbers that are consistent with the content of the envelopes on the bulletin board; but, this was checked by voters using the unvoted ballots in the spoilt-ballot audit explained at the end of the description of the voting phase.

The audit phase processing is determined by the unpredictable choice of a subset containing roughly half of the serial numbers. In practice, the particular subset is preferably arrived at in a pre-arranged manner from unpredictable and indisputable public data, such as closing prices of particular stocks on a particular day, all as committed to days earlier. For clarity here, however, it is shown as a public coin toss associated with each voted serial number. A coin showing “heads” means open the envelope and “tails” means leave it sealed. Particular rows in the independently-ordered second column of envelopes are pointed to by the row numbers contained in the envelopes opened according to the coins. That the letters posted pre-audit in these pointed to rows are consistent with the content of the opened envelopes of the first column is readily verified: the row number pointer is followed and the letter indicated should match that for the serial number if the envelope contains the word “same” and be the other letter if it contains “differ.”

None of the envelopes in these pointed-to rows of the second column should ever be opened, as each would provide a complete chain from a serial number to the corresponding vote. But all of the other envelopes in this column will be opened. Their consistency with the pre-audit postings is then checkable by anyone in a way

similar to that for the envelopes of the first column of envelopes: for each envelope opened in the second column of envelopes, two letters are indicated, the letter appearing with the envelope in the same row and the letter in the row of the results column that is pointed to by the row number in the envelope; for each such envelope opened, again, these two should match if the envelope contains “same” and not match if it contains “differ.”

Generalizing to Larger Ballots

The system naturally extends to incorporate any number of candidates, contests and ballot styles. For example, a vote-for-one contest of four candidates would use the letters “A,” “B,” “C” and “D” in that order except that each contest on each ballot starts with a random one of these letters and the letter sequence wraps around at the end, with “A” following “D.” Instead of just the words “same” and “differ,” a number is used to count the shifts forward around the cycle that should be performed. The total number of such shifts for the printing on a ballot and the two envelopes in its chain should always be a multiple of the number of letters, in this case four, so that votes are transferred correctly to the results column. Both numbers in an envelope chain could be chosen independently and uniformly and the printed shift amount adjusted accordingly.

When more than one mark is allowed per contest, random permutations of letters are used in a way similar to the cycles of letters used with vote-for-one. This extra randomization hides any patterns of marks in the letters posted.

When more than one contest is on the ballot, each preferably has its starting letters chosen independently and the single letters and same/differ indications of the bulletin board are replaced by lists of letters and amounts, respectively, with one element in the list for each contest. Each contest preferably uses a separate range of letters, so that a voter need only note the letters and not the contests they apply to. The assignment of letters to contests is preferably fixed and public and so the voter can use it to infer a letter that is rendered unreadable by marking.

A separate bulletin board is used for each polling place or other elementary unit such as a so-called “split precinct” with a unique combination of contests and

even including a printed “rotation” of candidate name orderings or other “ballot style.” The serial number of ballots is preferably extended so that distinct ranges identify each ballot style and any combination of ballot rotation.

In the resolution procedure, if there are multiple contests, then the voter should preferably choose one to dispute. The number of envelopes exposing the candidate names should be equal the number of candidates in the chosen contest. This prevents revealed anything about how the voter voted.

Some Variations

Three example variations are described below.

Non-individuated Ballots

Some jurisdictions require unique pre-printed serial numbers on ballot forms; other forbid any such marking, presumably for reasons of ballot secrecy.

Such prohibitions may provide ineffective protection and therefore false confidence. Latent fingerprints are of course always present on paper ballots handled by voters. These can be read without a trace using easily generated but non-staining iodine vapor or extremely sensitive infrared imaging used in chemical analysis. There are also many ways to deliberately mark a ballot uniquely that are either arguably innocent or not readily noticed. For instance, particular patterns of votes for “down ballot” contests or particular rank orderings are known to allow a voter to make his or her ballot recognizable to anyone who can see the votes per ballot. Without changing which ovals are filled, the way the ovals are filled can encode a unique and recognizable pattern. One example is by varying the direction of the zigzag motion used to fill with a pencil. Another example is filling beyond the oval boundary only on certain sides of each oval. Other examples include mechanical marks that can readily be seen with side lighting, such as writing made on a sheet above or indents made with a key or coin. Pinholes are only readily seen with backlighting. Folds or creases may appear innocent, but can be identifying. So-called “write-in” possibilities lend themselves to all manner of recognition. There are many kinds of so-called “invisible inks,” which are visible under certain type

of illumination or at certain temperatures or by application of certain chemicals.

Moreover, prohibiting such individuation means a much heavier reliance on audit of physical ballots, such as by representatives from all parties or candidates that typically scrutinize ballots in manual counting or in audits (not to mention so-called “chain of custody” with its extremely costly integrity and opportunities for ballot inspection). Providing visual inspection of ballots to partisans, even if only a fraction of ballots, provides substantial potential enforcement for typical improper influence schemes like vote buying and coercion. With Scantegrity, there is no need to allow partisans access to inspect ballots, except during the resolution procedure, where votes are not shown. In fact, even officials have no need to be able to see ballots apart from the serial number in resolution procedures. Thus, Scantegrity audit even with visible individuation of ballots allows an election system to avoid the otherwise substantial potential for improper influence needed for efforts at conventional optical-scan ballot obtain integrity without Scantegrity.

Nevertheless, if printing ballots in ink that can only be read using ultraviolet lamps is desired as a way to avoid prohibition of individuation, it is feasible. Many items are marked and numbered in such ink for various commercial and security purposes. Suitable inks are, for example, made by Hewlett Packard for industrial inkjet printers as well as their Indigo digital presses typically used to print ballots today. Inexpensive LED flashlights can be provided to voters for noting their letters at a polling place. Scanners or copy stands can be fitted with suitable lamps so that they can read the fluorescing inks without modification of the sensing elements or electronics. The serial number on the ballot form would be printed in visible ink on the counterfoil chit for the voter and in ultraviolet ink for the barcode spanning onto the ballot proper.

Ultra-High-Level Document Security

Fiber patterns form a well-known unique “fingerprint” of paper that easily makes it extremely difficult to copy such structure without detection. Such techniques were developed by national laboratories for use in international arms verification. Less than a mega-pixel image of roughly a twentieth of an inch square provides security of a very high level and is easily seen and understood. If such images are formed from, say, a

region next to the serial number and posted online with corresponding serial number before the election, then the paper used in the election is securely authenticated in a way the public can check. A valid such chit readily verifies the validity of the rest of the ballot, if the fiber patterns match across the separation line. Hand-held document inspection devices provide digital capture for under 1,000USD, such as the MiScope by Zarbeco, typically include oblique lighting to enhance visibility of the fibers on the paper surface.

Voters with Disabilities

Voters with disabilities who cannot readily see the letters on the ballot may choose to vote using audio as today, but a variant specific to Scantegrity has the advantage of allowing voters not to have to operate equipment other than the audio device. What they would hear using headphones is first the serial number. Then, for each contest the letters are read in alphabetical order, each followed by the corresponding candidate name. The voter then utters the letter, by saying it out loud, for each candidate that they wish to vote for. For instance, a voter would hear “Jefferson A, Madison B” for a swapped ballot and “Madison A, Jefferson B” for an un-swapped ballot. When the voter hears his or her candidate, the voter utters the corresponding letter heard. Poll workers then mark down the letters, preferably on a special ballot form that only shows letters in order. These forms can have pre-printed serial numbers or the numbers can be added by marks per digit.

These forms are then scanned along with other ballot forms. The special range of the serial number, however, indicates that the equivalent mark positions need to be determined from the data about which letter would have labeled which position had the ballot been ordered by candidate names. Uttering the letters for the candidates does not reveal the vote to someone listening who does not know which letter would have labeled what position. So a recording by the voter can serve much as the notes made on paper by other voters. The recording would include the voter repeating the serial number and could be made, for instance, using the memo feature on a cellular phone, music player, or cameras, or even by calling an answering machine or voicemail. Voters using audio ballots can thus gain the ability to autonomously verify the recording of their vote by dialing into an automated phone system or using a self-voicing web interface.

When ballots are printed, the computer that controls their printing could also produce the corresponding audio file. This can be accomplished using the same simple software techniques used to create so-called 1-800 number IVR systems: sound sample files are concatenated together under program control and played out using a sound card or the like.

Another way to read out letters relies on optical scan ballot marking devices. These first scan a ballot, take voter input, and then overprint on the ballot so that it can subsequently be processed as any other ballot. They could be modified to include the recognition of the letters and the reading of those letters aloud to the voter.

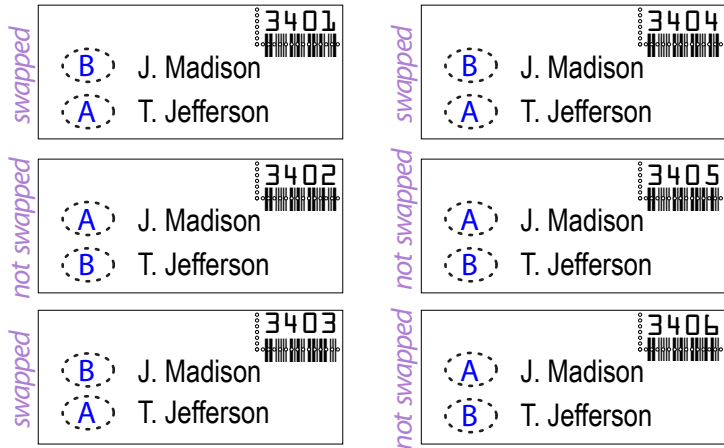
In order to ensure that the letters are correctly read in any case, voters should ideally be able to spoil such audio ballots and retain the unvoted copies. One example way to accomplish this is by providing the recording to the voter in a portable MP3 player that the voter would use to listen to the audio with the benefit of a familiar user interface. But if the voter wants to spoil that ballot, then the MP3 recording can be transferred to a player of the voter or an observer and the voter provided with another ballot.

Conclusions

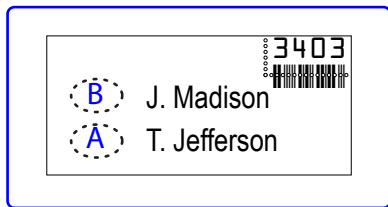
This whitepaper has taken a closer look at the particular technical and operational aspects of the Scantegrity system. Nothing beyond ordinary current commercial practice, from software to printing and document security, has been called for. Similarly, from an operational perspective the procedures required appear quite achievable in ordinary practice. Moreover, the simplicity in achieving combined integrity and ballot secrecy make the system of interest and readily explainable even for secondary school classrooms.

The unique printing on ballots can substantially improve ballot secrecy over that provided by traditional optical-scan paper ballot systems.

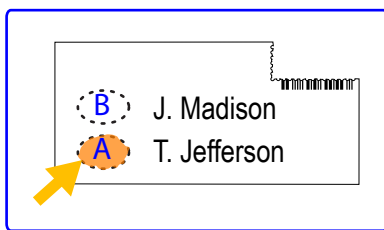
More fundamentally, voters receive for all practical purposes indisputable proof that their ballots are tallied as cast. Thus, without substantially changing existing optical scan voting systems or procedures, at very low additional cost, integrity can be raised to a level that is secure against a national adversary and that can give a maximal boost to voter confidence.



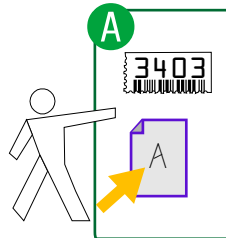
Ballots



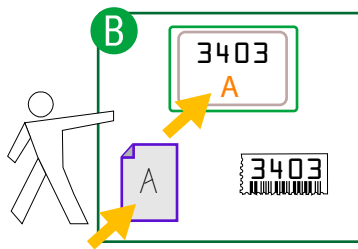
Ballot Provided to Voter



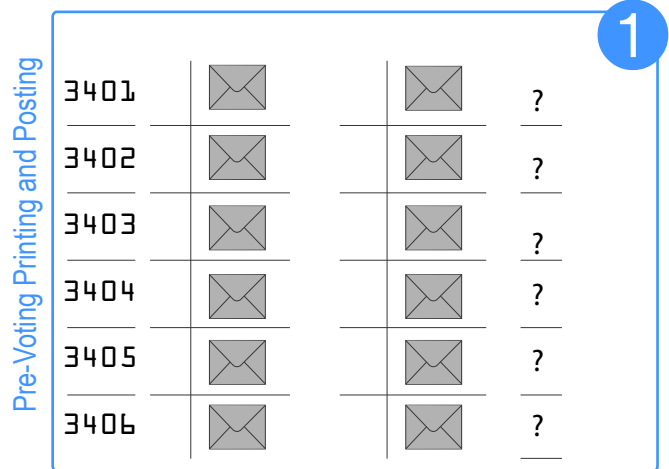
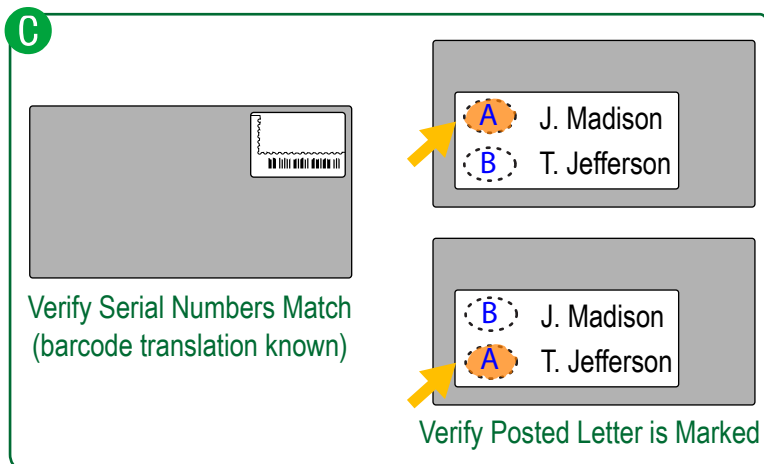
What Voter Leaves in the Ballot Box



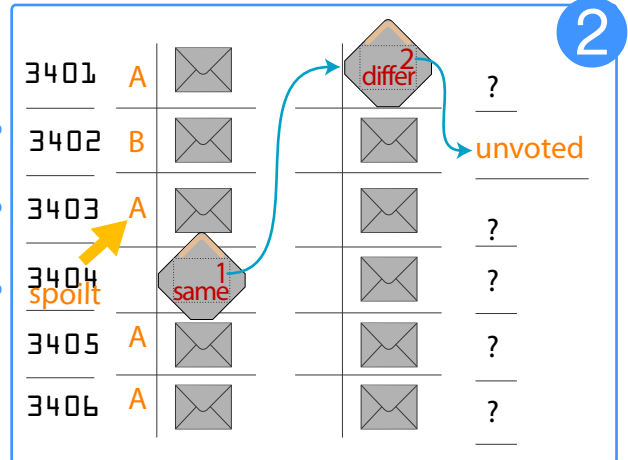
What Voter Takes Home



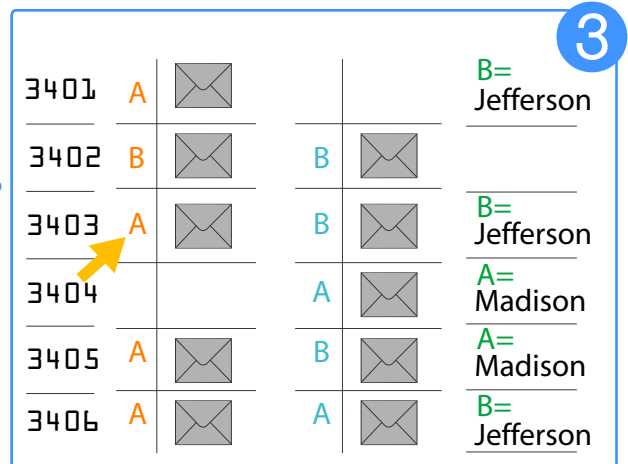
Voter Checking for Match Online



Posting During Voting



Pre-Audit Including Results



Audit Completion

