Auditing the PCOS-Based Automated Election System*(2nd draft – 6/11/2011)

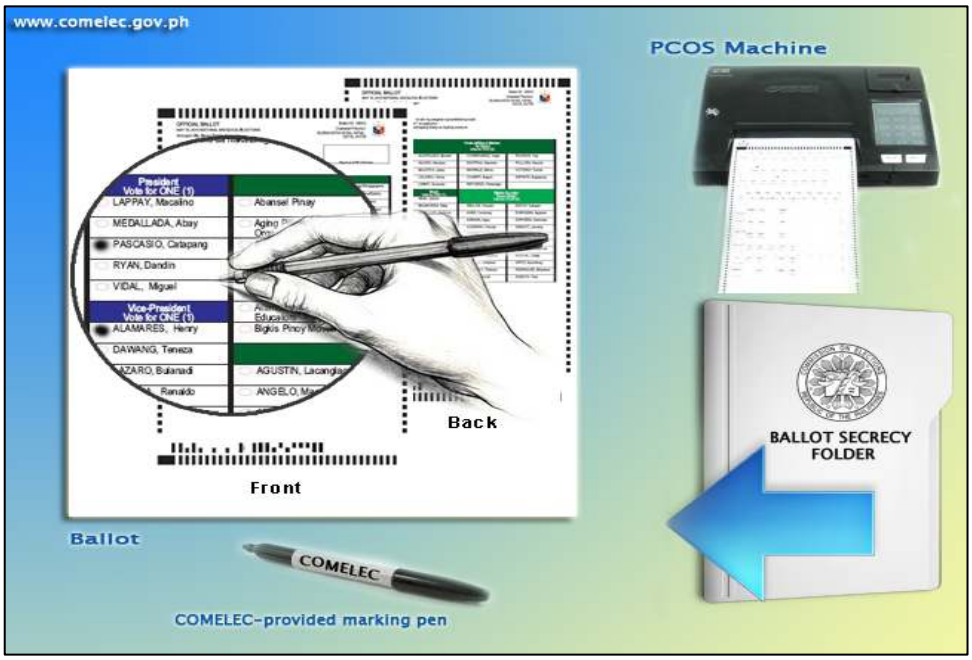
Felix P. Muga II, Ph.D.
Associate Professor, Mathematics Department, Ateneo de Manila University (ADMU)
Senior Fellow, Center for People Empowerment in Governance (CenPEG)

Introduction

OMR means ***Optical Mark Recognition***. Wikipedia defines it as "the process of capturing human-marked data from document forms".[1] According to Haag and others, "**OMR** is the scanning of paper to detect the presence or absence of a mark in a predetermined position".[2]

The OMR technology was used in the First Nationwide Automated Elections of May 2010 with the **PCOS** (**Precinct Count Optical System**) machines called **SAES-1800**[3] provided by **SMARTMATIC**. See Figure 1.

Figure 1. PCOS Machine, Ballot, Secrecy Folder and Marking Pen



http://www.comelec.gov.ph/modernization/2010_natl_local/how_to_accomplish_ballot.html

The *Omnibus Election Code of the Philippines* defines a *precinct,* thus[4] "the unit of territory for the purpose of voting is the election precinct, and every barangay as of the approval of this Act shall have at

least one such precinct." In the May 2010 Elections, the COMELEC grouped a number of precincts in the same locality into a ***clustered precinct*** with not more than 1,000 registered voters.

For each clustered precinct a PCOS machine was provided to appreciate the accomplished ballots and count the votes for the candidates of the different contests.

Project Specifications

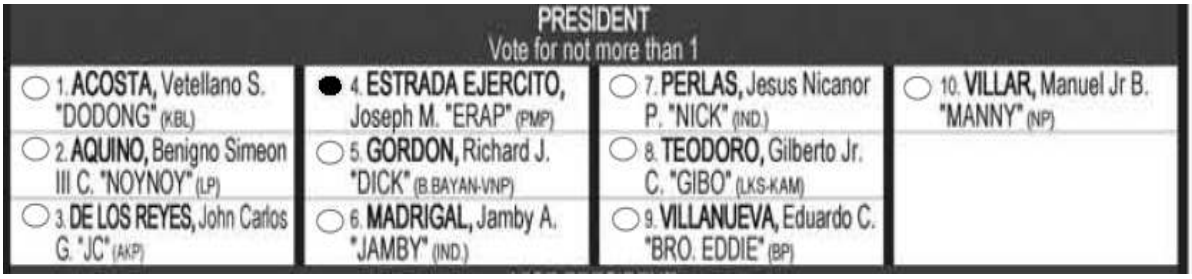
The Commission on Elections (Comelec) specified[5] that

1. the names of the candidates shall be pre-printed in the ballot,
2. one ballot sheet per voter shall be used to accommodate all the names of the candidates for all elective positions, with both sides of the ballot sheet may be utilized,
3. the system shall be able to recognize ***full shade***, ***partial shade***, ***check marks***, and ***x marks*** on the appropriate space on the ballot opposite the name of the candidate to be voted for, and
4. The system shall count the voter's vote as marked on the ballot with an accuracy rating of at least **99.995 %**.

Ballot Design

In the May 2010 elections, the front page of the ballot was devoted to the national contests consisting of presidential, vice-presidential, senatorial and the party-list contests. In Figure 2 the presidential contest is shown with 10 candidates where the oval that corresponds to the fourth candidate is shaded.

Figure 2: The Candidates of the May 2010 Presidential Contest



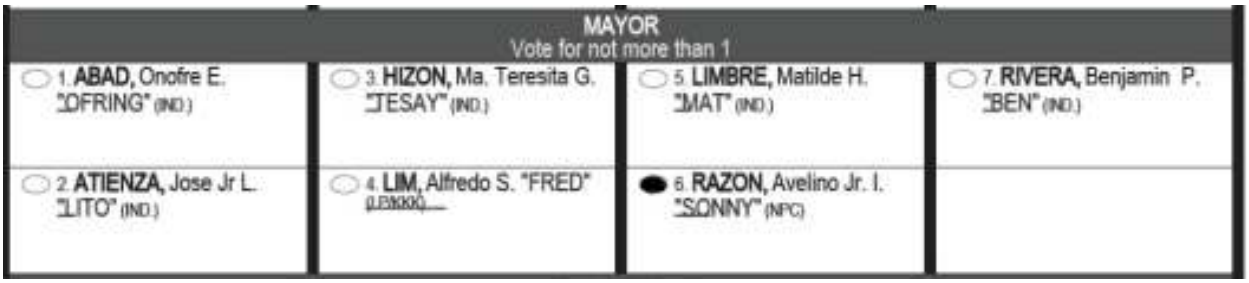
The local contests were assigned to the back page of the ballot. In the precincts of the municipalities and component cities, the contests were that of the Member for the House of Representatives, the Governor, the Vice-Governor, the Sangguniang Panglalawigan, Mayor, Vice-Mayor and the Sangguniang Bayan (for municipalities) or Sangguniang Panglunsod (for cities). In the precincts of the highly-urbanized cities and of the municipality of Pateros, the local contests did not include that of electing provincial officials.

Figure 3 provides the local contest for the Mayor in the City of Manila where there are seven (7) candidates and the oval corresponding to candidate number 6 is shaded.

All the candidates of a contest are arranged uniformly in two-dimension in the ballot. In the presidential contest, candidate number 8 is located in the second row and third column. The oval that corresponds

to a candidate is located to the left of a number and the name of the candidate. A PCOS machine shall scan these predetermined positions called as the "***vote positions***" in the ballot and shall determine the presence or absence of a mark in each oval. Each marked position shall be subjected to a set of criteria before it is declared a valid vote.

Figure 3: The Candidates of the May 2010 Manila City Mayor's Contest



If a ballot is designed in this manner, then a PCOS machine shall be able to interpret the decision of a voter to vote or not to vote a candidate of a contest. We shall call such ballot as ***PCOS-ready***.

Note that the front page of a ballot contains the national contests. Thus all the ballots have the same configuration in the front page. The configuration of the ballots differs in the back page. Two precincts may have different configurations if the precincts belong to different political jurisdictions like provinces, cities, districts, municipalities, or barangays.

Two precincts in a municipality will have the same ballot configuration. However, two precincts in a city may have different ballot configuration if each precinct belongs to different districts in the city.

The Mathematics of a Voter's Decision

The primarily function of a PCOS machine is not only to detect the presence of a mark in vote position but also the absence of such a mark. Thus, the decision of a voter in contest is a sequence of **YES** or **NO** for the vote positions in a contest. It is a **YES** for a vote position (candidate) if the voter chooses and marks the position and a **NO** if the voter does not choose such vote position (or candidate).

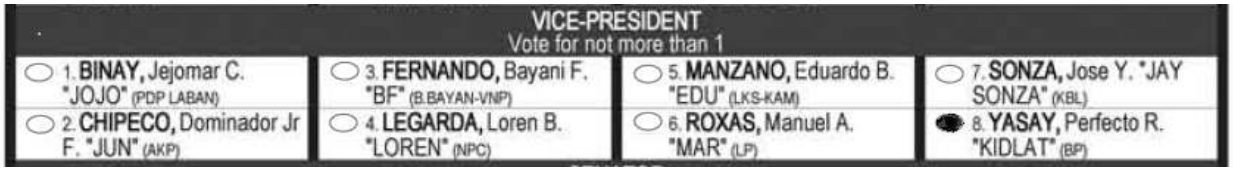
There is an order in the sequence of YES and NO since the candidates are arranged numerically. For example, in **Figure 2**, the decision of the voter for the presidential contest is ***NO, NO, NO, YES, NO, NO, NO, NO, NO, NO.*** This is a sequence of **YES** and **NO** of length ten (**10**). Since the voter marked the 4th vote positions, the candidate that corresponds to the vote position is the voter's choice for president.

In **Figure 3** the decision of the voter for the mayor's contest in the City of Manila is ***NO, NO, NO, NO, NO, YES, NO.*** There are seven candidates for the mayor's contest in the City of Manila. Thus, the voter's decision is a sequence of **YES** and **NO** of length seven (**7**). The voter marked the **6th** vote position hence the voter was voting for the 6th candidate in the list of candidates for the mayor's contest.

If we replace the **YES** decision for the one (**1)** bit or binary digit and the NO decision for the zero (0) bit, then the voter's decision for the presidential contest as shown in **Figure 2** becomes **0001000000** and the

voter's decision for the mayor's contest as shown in **Figure 3** becomes **0000010**. These are binary sequences and the length of each binary sequence is equal to the number of vote positions or the number of candidates of each corresponding contest.

Figure 4 The Voter's Decision for the Vice-President Contest



The voter's decision for the vice-president contest as shown in **Figure 4** is represented by the sequence **00000001.**This means that the voter voted for the 8th candidate of the vice-presidential contest.

Figure 5 The Voter's Decision for the Senatorial Contest with 12 Choices

The voter's decision for the senatorial contest is represented by the binary sequence **1001001001000000001100000000000010001100000000010000010001000** of length 61 with 12 one bits.

A change in the value of a term changes the decision of the voter. For example, if the first term from the left of the decision for the senatorial is changed from 1 to 0, and the next term to the right changes from 0 to 1, then the decision of the voter changes from the 1st candidate to the 2nd candidate as one of his or her choices for the senators.

This means that each decision of a voter for a given contest corresponds to one and only binary sequence.

There are **1**87 vote positions or candidates in the party list contest. Suppose that the voter chooses the **49th** candidate in the list. Then the decision of the voter for the party list contest is given by

00000000000000000000000000000000000000000000000010000000000000000000000000000000000
00000000000000000000000000000000000000000000000000000000000000000000000000000000000
000000000000000000000.

Therefore, the decision of the voter for the national contests as indicated in the front page of the ballot is given by the binary sequence of length **266**:

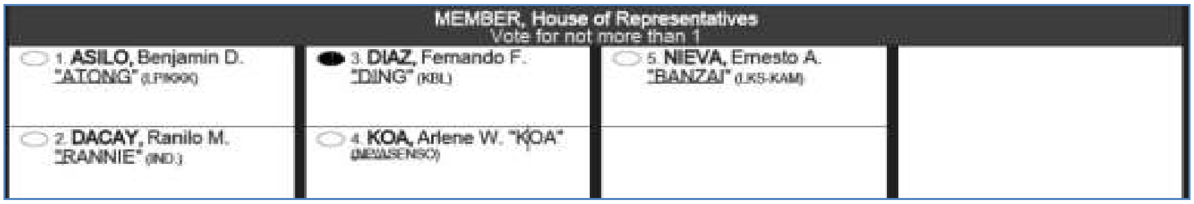
00010000000000000110010010010000000011000000000000100011000000000100000100010000000
00000000000000000000000000000000000000000000100000000000000000000000000000000000000
00000000000000000000000000000000000000000000000000000000000000000000000000000000000
00000000000000000.

The ballot configuration for the local contests differs in two types of precincts. The first type consists of the precincts from highly-urbanized cities and the Municipality of Pateros. The second type consists of the precincts from the municipalities and component cities of a province. The voters in the first type of precinct do not vote for the provincial contests while the voters in the second type are voting for the provincial contests.

The City of Manila is a highly-urbanized city. The local elections in this city consist of four contests: for the member of the House of Representatives representing one of the six legislative districts, for the city mayor, for the city vice-mayor, and for the Sangguniang Panlungsod with 6 members for each of the six districts.

In the first district of Manila, the contest for the member of the House of Representatives had five candidates. If our voter chooses for the third candidate, then the decision of the voter for the said contest is represented by the binary sequence **00100**.

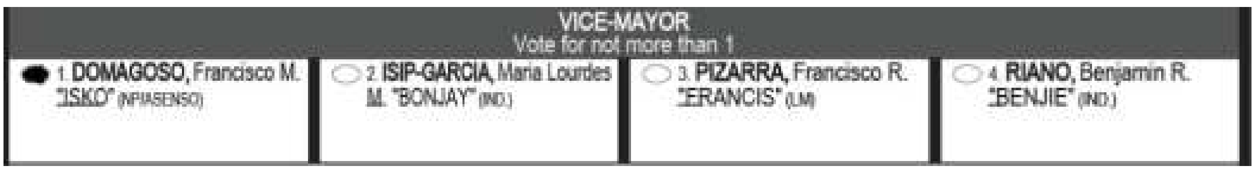
Figure 6The Contest for the Member, House of Representatives Representing the 1st District of Manila



In the mayor's contest the decision of our voter is given by **0000010** as shown in **Figure 3**.

The contest for the vice-mayor of Manila had four candidates. Suppose that our voter decides to choose the first candidate for the vice-mayor as shown in **Figure 6**, then the decision of the voter is represented by **1000**.

Figure 7 The Contest for the Vice-Mayor of Manila



The contest for the Sangguniang Panlungsod in the first district of Manila had 41 candidates. A voter is expected to vote for not more than 6 candidates. Suppose that our voter chooses for 2nd, 16th, 17th, 29th, 34th and 41st candidates. Then the binary sequence representing the decision of the voter for the Sangguniang Panlungsod contest is given by **01000000000000011000000000001000010000001.**

Therefore, the decision of our voter in the 1st district of Manila for the local contests is given by the binary sequence of length **57**:

001000000010100001000000000000011000000000001000010000001.

Therefore, the decision of the voter for the national and local contests as shown in the ballot is represented by the binary sequence of length **323**:

00010000000000000110010010010000000011000000000000010001100000000010000010001000000
00000000000000000000000000000000000000000100000000000000000000000000000000000000000
00000000000000000000000000000000000000000000000000000000000000000000000000000000000
00000000000000000001000000001010000100000000000000110000000000010000100000001

where the decision for

- the presidential contest is located in bits **1** to **10** (with yellow color),
- the vice-presidential contest is located in bits **11** to **18** (with bright green color),
- the senatorial contest is located in bits **19** to **79** (with turquoise color),
- the party list contest is located in bits **80** to **266** (with pink color),
- the contest of the Member, House of Representatives is located in bits **267** to **271**(with yellow color again),
- the mayor's contest is located in bits **272** to **278** (with bright green color again),

- the vice-mayor's contest is located in bits **279** to **282** (with turquoise color again), and
- the contest for Sangguniang Panlungsod is located in bits **283** to **323** (with pink color again)**.**

The second type of precincts is found in a municipality or in a component city of a province. Hence, voters in these precincts are expected to decide on **seven (7)** local contests. These are the contest for:

1. provincial governor,
2. vice-provincial governor,
3. Sangguniang Panglalawigan,
4. Member, House of Representatives,
5. Municipal or City Mayor,
6. Municipal or City Vice-Mayor, and
7. Sangguniang Bayan or Sangguniang Panlungsod.

Therefore, we can represent mathematically the decision of the voter on each of the vote positions in the ballot by the binary (ordered) sequence

$$v_{11}v_{12}\dots v_{1n_1}v_{21}v_{22}\dots v_{2n_2}\dots v_{k1}v_{k2}\dots v_{kn_k}$$

where the value of v_{ij} is either 0 or 1and $v_{i1}v_{i2}\dots v_{in_i}$ is the decision of the voter for contest *i,* for***i = 1, 2, …, k*** with ***k*** being the number of contests listed in the ballot.

We can also say that the properly accomplished PCOS-ready ballot is represented by one and only one binary sequence

$$v_{11}v_{12}\dots v_{1n_1}v_{21}v_{22}\dots v_{2n_2}\dots v_{k1}v_{k2}\dots v_{kn_k}.$$

The Accuracy of PCOS Machines

The primary function of the PCOS machine is to determine the presence or absence of a mark in a predetermined position. Since a vote of a voter is expressed as a marked vote position in the ballot, the vote counting accuracy of the PCOS machines is bound by the accuracy of the PCOS machines in reading the presence or absence of a mark in the vote positions and is dependent on the ballot configuration.

The ***error rate*** of a PCOS machine is equal to the number of errors the machine committed in reading the vote positions divided by the total number of vote positions under consideration.

The ***accuracy rate*** of the said PCOS machine is equal to the formula

$$accuracy\,rate = (1 - error\,rate) \times 100\%.$$

Hence, if a PCOS machine has at most one error in scanning 500,000 vote positions, then its accuracy rating is at least equal to $\left(1 - \frac{1}{500{,}000}\right) \times 100\% = 99.9998\%$.

Hence, the vote counting accuracy of the PCOS machine is below the value computed above. It can be precisely determined by the configuration of the ballots.

Suppose that the PCOS machine scans 1,000 ballots with 400 vote positions per ballot. Then the PCOS machine scans a total of 400,000 vote positions. Since its error rating is 1:500,000, it follows that it can be accurate (with high probability) in reading the 400,000 vote positions. Hence, no vote will be miscounted.

1. If the ballot contains 20 single-winner contests with 20 candidates per contest, then each ballot has at most 20 marked vote positions. Thus, the total number of marked vote positions scanned by the machine without an error is 20,000. Therefore, the vote counting accuracy of the PCOS machine with this ballot configuration is $\left(1-\frac{1}{20{,}000}\right)\times 100\% = 99.995\%$.
2. If the ballot contains 50 single-winner contests with 8 candidates per contest, then each ballot has at most 50 marked vote positions and the machine scans a total of 50,000 votes with no error. Therefore, the vote counting accuracy of the PCOS machine with this ballot configuration is $\left(1-\frac{1}{50{,}000}\right)\times 100\% = 99.998\%$.

Note that in the May 2010 elections there were **266** vote positions and at most **15** were marked in the national contests. For a ballot in a precinct of highly-urbanized cities (including Pateros) the number of marked vote positions for the local contests may have reached a total of **11** and for a ballot in a municipality or a component city, the local contests may have reached a total of **19** marked vote positions. Thus, the back page of the May 2010 elections may reach a total of **200** vote positions. Since a clustered precinct may reach a total of 1,000 ballots, the Comelec should have required an error rate of at most **one** out of **500,000** vote positions for each PCOS machine to avoid a vote miscount.

Determining the Accuracy of a PCOS Machine on a Ballot Configuration

Since a PCOS machine is primarily designed to determine the presence or absence of a mark in a vote position, the accuracy of the machine must be measured based on all the possible ways a vote position in a contest is marked.

For the May 2010 senatorial contest alone there were 1,742,058,970,275 ways of shading 12 vote positions exactly. Hence, it is not possible that we can test all the possible ways that a ballot will be accomplished. But since it is not practical to do it, we will choose a random sample of the ways that a vote position is marked or chosen.

The number of samples can be determined using the formula

$$\frac{z^2 p(1-p)}{c^2}$$

The value of *z* at 95% level of confidence is $z = 1.959963985$, c = .01 if the confidence interval is 1%, and *p* is the (worst) probability that a vote position is chosen or not which is $p = 0.5$.

Hence, the number of random ballot samples is **9,604.**

A ballot is chosen at random if the expected number of vote positions per contest is chosen at random.

Since a properly accomplished *PCOS*-ready ballot is represented by one and only one binary sequence $v_{11}v_{12}\dots v_{1n_1}v_{21}v_{22}\dots v_{2n_2}\dots v_{k1}v_{k2}\dots v_{kn_k}$, where each contest $c_i$is represented by the binary number $v_{i1}v_{i2}\dots v_{in_i}$, it follows that a randomly selected ballot is achieved by randomly selecting a binary number of length n_i for each contest c_i where the number of 1 bits in the binary number depends on the number of choices allowed for contest c_i, for ***i = 1,2, …, k*** where ***k*** is the total number of contests.

Thus for the presidential contest, one of the binary numbers given below is chosen at random:

1000000000, 0100000000, 0010000000, 0001000000, 0000100000,
0000010000, 0000001000, 0000000100, 0000000010, 0000000001

We shall denote the binary sequence representing each ballot by **A0001**, **A0002**, …, **A9604** where **A0001** corresponds to the first ballot numbered **0001**, **A0002** to the second ballot numbered **0002**, and so on, until **A9604**which corresponds to the last ballot numbered **9604**. These binary sequences shall be called as the **true binary representations** of the properly accomplished *PCOS*-ready ballots.

The ballots are fed manually into a PCOS machine. The appreciation of the PCOS machine on each ballot is represented by a binary sequence which we shall call as the ***PCOS*-interpreted binary sequence**. The binary sequence on ballot **0001** shall be denoted by **B0001.** The binary sequence on ballot **0002** shall be denoted by **B0002,** and so on, until the last ballot **9604,** and the PCOS-interpreted binary number shall be denoted by **B9604**.

The binary number **A0001** is compared to **B0001** bit by bit, the binary number **A0002** to **B0002** bit by bit also, and so on.

The total number of instances when there is a difference between an ***i***th bit of the ***true binary sequence*** and the ***i***th bit of the corresponding ***PCOS*-interpreted binary sequence** is equal to the number of errors in the appreciation of the vote positions by the *PCOS* machine for *i* = 1,2, ..,up to $n_1 + n_2 + \cdots + n_k$, which is the length of the binary sequence.

If the number of errors is equal to **m** and the number of vote positions of the 9,604 ballots is **r**; and suppose that **n** is equal to the floor of **r/m.** Then the error rating of the *PCOS* is ***one*** out ***n*** vote positions.

For example, if each of the **9,604** ballots has **350** vote positions, then the total number of vote positions is **3,361,400**. If the ***PCOS*** machine committed **9 errors** in interpreting the decision of each vote positions, then **3,361,400/9 = 373,488.8889.** Hence, the error rating of the *PCOS* is at least **1:373,488** or **1/373,488**. The Accuracy rating of the ***PCOS*** machine is **99.99973225%**.

If the ARMM regional elections will be synchronized with the May 2013 national and local elections, then there may be 3 types of ballot configurations for the local contests:

1. for the highly urbanized cities and the Municipality of Pateros,
2. for municipalities and component cities of non-***ARMM*** provinces, and
3. for municipalities and component cities of ***ARMM*** provinces.

Hence, three different sets of ballot samples with **9,604** ballots per set may be prepared to determine the accuracy of the *PCOS* machine in each ballot configuration.

Auditing the May 2010 Elections

Since a 100%-accuracy rate of the voting, counting and canvassing machines is difficult to establish, an audit on the counting and canvassing of the election results shall be conducted in every canvassing level before the results are certified and the winners are proclaimed.

The audit shall be transparent such that "the public must be allowed to observe, verify and point out procedural mistakes in all phases of the audit without interfering with the process."

The audit shall be risk-limiting such that "it must reduce the risk of confirming an incorrect outcome." The error shall be negligible such that it shall not affect the outcome of the contests.

The audit shall be mandatory such that "it is required at all levels of canvassing and consolidating the votes."

The audit shall be nation-wide such that "it must be done in all precincts and in all municipalities, cities, provinces, and at the national levels."

The provision in Republic Act 7936 (Sec 24) shall be amended since it leads to over-sampling in legislative districts with small voting population and under-sampling in districts with very large voting population and it does not address the problem of the errors committed by an *ACM*. The provision states that "A new Section 29 is hereby provided to read as follows[6]:

> Sec 29. *Random Manual Audit*. – Where the AES is used, there shall be a random manual audit in one precinct per congressional district randomly chosen by the COMELEC in each province and city. Any difference between the automated and manual count will result in the determination of root cause and initiate a manual count for these precincts affected by the computer or procedural error."

In the **Minute Resolution No. 10-0453**, the COMELEC decided to increase the number of precincts to be audited per district. The resolution specified 5 clustered precincts.

The oversampling problem was worsened. In BATANES, the probability of success of a clustered precinct to be chosen is now 5/30 or 16.667%, while in ANTIQUE, the probability of success of a clustered precinct in the province is 5/656 or 0.762%.

In a random manual audit, every clustered precinct must have an equally likely chance of being chosen to be in the sample.

The Random Manual Audit that was commissioned by the Comelec examined only the positions for president and vice-president in the national level and the positions for Mayor, Member of the House of Representatives, and the Governor in the local level. The Party-list Representatives and the Senators were not examined as well as the positions for the Vice-Mayor, Vice-Governor, etc.

A PCOS with Voter's Verification Feature

Post-election audit checks the performance of the automated election system in all of the contests. It is done after all the voters had cast their votes.

Right after the voter cast his or her votes, the **PCOS** generates a mathematical representation of the voter's decision which is a binary number. This is called as the ***PCOS-interpreted binary representation of the voter's decision***.

The *PCOS* shall have a **binary to vote-display program** that converts the ***PCOS-interpreted binary number*** into the decision of the voter and display the names of the candidates that the voter had chosen in the display monitor. If the voter accepts the displayed votes then the ***PCOS-interpreted binary representation*** becomes the ***accepted binary representation of the voter's decision.*** Otherwise, it shall be called as the ***rejected binary representation of the voter's decision.***

The counting of the votes by the *PCOS* must be based on the accepted binary representation of the voter's decision.

The addition of the votes shall be done on the corresponding terms of all the binary representations in the precinct. The sum in each term is a decimal value. The terms are separated by a comma (comma-separated values).

We shall denote the mathematical expression as

$$V_{11}V_{12}\ldots V_{1n_1}V_{21}V_{22}\ldots V_{2n_2}\ldots V_{k1}V_{k2}\ldots V_{kn_k}$$

where V_{ij} represents the sum of the votes for the ***j***th candidates in the ***i***th contest expressed as a 3-digit decimal. This decimal number is the mathematical representation of the decision of the voters who actually voted in the precinct. Note that any change in the value of a term s_{ij}represents another decision of the actual voters in the precinct.

Suppose that a precinct consists of 7 voters each of whom confirms her or his decisions as interpreted by the *PCOS* machine are given as:

Figure 8: The Decision of Seven Voters as a Binary Sequence
01010000011001
10010001010001
10000100000111
00101001101000
00010001001001
10010001001001
01001000001101

The precinct has three contests. The first contest has 3 candidates with 1 winner, the second contest has 4 candidates with 1 winner, and the last contest has 7 candidates with 3 winners.

The sum of the corresponding terms of the accepted binary numbers is given in the last row:

Figure 9: Computing the Votes For Each Candidate													
0	1	0	1	0	0	0	0	0	1	1	0	0	1
1	0	0	1	0	0	0	1	0	1	0	0	0	1
1	0	0	0	0	1	0	0	0	0	0	1	1	1
0	0	1	0	1	0	0	1	1	0	1	0	0	0
0	0	0	1	0	0	0	1	0	0	1	0	0	1
1	0	0	1	0	0	0	1	0	0	1	0	0	1
0	1	0	0	1	0	0	0	0	0	1	1	0	1
3	2	1	4	2	1	0	4	1	2	5	2	1	6

Note that the fifth voter (5th row) did not vote for any candidate in the first contest. The leading candidates in the precinct are:

a. The first candidate of the first contest with votes $V_{11} = 3$,

b. The first candidate of the second contest with votes $V_{21} = 4$, and

c. The first, fourth and last candidates of the third contest with votes $V_{31} = 4$, $V_{34} = 5$, and $V_{37} = 6$.

On Nullified Votes

A nullified vote (or NULL VOTE) is an invalid vote of a voter in a contest. A vote in a contest is not a valid vote if it does not pass a set of criteria for a valid vote. The votes in a ballot are not valid if the ballot does not pass a set of criteria for a valid accomplished ballot.

It is a possible that in a multiple-winner contest some of the votes are valid and some are not.

The nullified votes of a voter in an accomplished ballot may be represented by a binary sequence $n_{11}n_{12}n_{13}n_{14}n_{21}n_{22}n_{23}n_{24}\ldots n_{k1}n_{k2}n_{k3}n_{k4}$ where $n_{i1}n_{i2}n_{i3}n_{i4}$ is the number of nullified votes expressed as a binary number of length **4** for each contest c_i and for ***i*** = **1, 2**,…, ***k*** where ***k*** is the number of contest.

For a single-winner contest c_i there are just two possible binary sequences **00000** for a valid vote and the vote is not nullified and **0001** where the vote is invalid.

For a multi-winner contest like the senatorial contest, the possible 4-bit binary sequences representing the number of nullified votes is **0000** for no nullified vote, **0001** for one nullified vote, **0010** for two

nullified votes, **0011** for three nullified votes, and so on until **1100** where all the twelve votes are nullified.

From here on we shall denote $n_{i1}n_{i2}n_{i3}n_{i4}$ by n_i as a 4-bit binary number.

The number of nullified votes of a contest per precinct is the sum of the corresponding terms of the sequence of nullified votes per ballot. The sum is denoted by N_i. Hence, in a precinct we have $N_1\ N_2 \ldots N_k$ representing the number of nullified votes of each contest in the precinct.

On Empty Votes

The voter may not choose the maximum of allowable choices in a contest.

In a single-winner contest, a voter may decide not vote for any candidate. Hence, we have **1** empty vote for the contest.

In a multiple-winner contest a voter may decide to vote for a lesser number of candidates instead of the number of votes allowed in the contest. For example, if the voter is allowed to choose **12** candidates, but chooses only **7** candidates, then the number of empty votes in this contest is **5**.

The empty votes of a voter in an accomplished ballot may be represented by a binary sequence $e_1 e_2 \ldots e_k$ where e_i is the number of empty votes for each contest c_i and k the number of contest where each $e_i = e_{i1}e_{i2}e_{i3}e_{i4}$ similar to the 4-bit binary representation of the nullified votes.

The number of empty votes per contest in a precinct is the sum of the corresponding terms of all the sequence of empty votes. We shall denote the sum by E_i. Hence, in a precinct we have $E_1\ E_2 \ldots E_k$ representing the number of empty votes per contest in the precinct.

Auditing the Election Results at the Precinct Level

We can associate every accomplished ballot three parts: the decision of the voter in binary number, the sequence of nullified votes, and the sequence of empty votes.

Hence each accomplished ballot is represented by the extended binary sequence

$$v_{11}v_{12}\ldots v_{1n_1}v_{21}v_{22}\ldots v_{2n_2}\ldots v_{k1}v_{k2}\ldots v_{kn_k}n_1n_2\ldots n_ke_1e_2\ldots e_k$$

and the decision of the voters in a precinct is represented by the decimal sequence

$$V_{11}V_{12}\ldots V_{1n_1}V_{21}V_{22}\ldots V_{2n_2}\ldots V_{k1}V_{k2}\ldots V_{kn_k}N_1N_2\ldots N_kE_1E_2\ldots E_k$$

where

- V_{ij} is the sum of all the v_{ij}s in the precinct for candidate *j* in contest c_i,
- N_i is the number of nullified votes in contest *i*, and

- E_i is the number of empty votes for contest *i*.

Note that $V_{i1} + V_{i2} + \cdots + V_{in_i} + N_i + E_i$ is the expected total number of votes in the precinct for contest c_i with $V_{i1} + V_{i2} + \cdots + V_{in_i}$ the total number of valid votes, N_i the total number of invalid votes and E_i the total number of empty votes in the precinct for contest c_i.

ElectionAudits.Org defines an **audit unit** as[7]

> "In post-election audits, each ballot (or paper record) is assigned to an audit unit – a group of paper records from a precinct, counting machine, or batch of ballots."
>
> a precinct, counting machine, or batch of ballots."

If the voter is allowed to verify his or her votes and accept the votes that are displayed, then the PCOS-interpreted binary sequence of the accomplished ballot of the voter becomes the accepted binary representation and represents also the true decision of the voter.

Thus, instead of using the actual ballots in a precinct for election audit, the audit can be performed on the accepted binary sequences and the sequences of nullified votes and empty votes.

An audit unit shall consist of these sequences.

Complexity of the Philippine Canvassing Structure

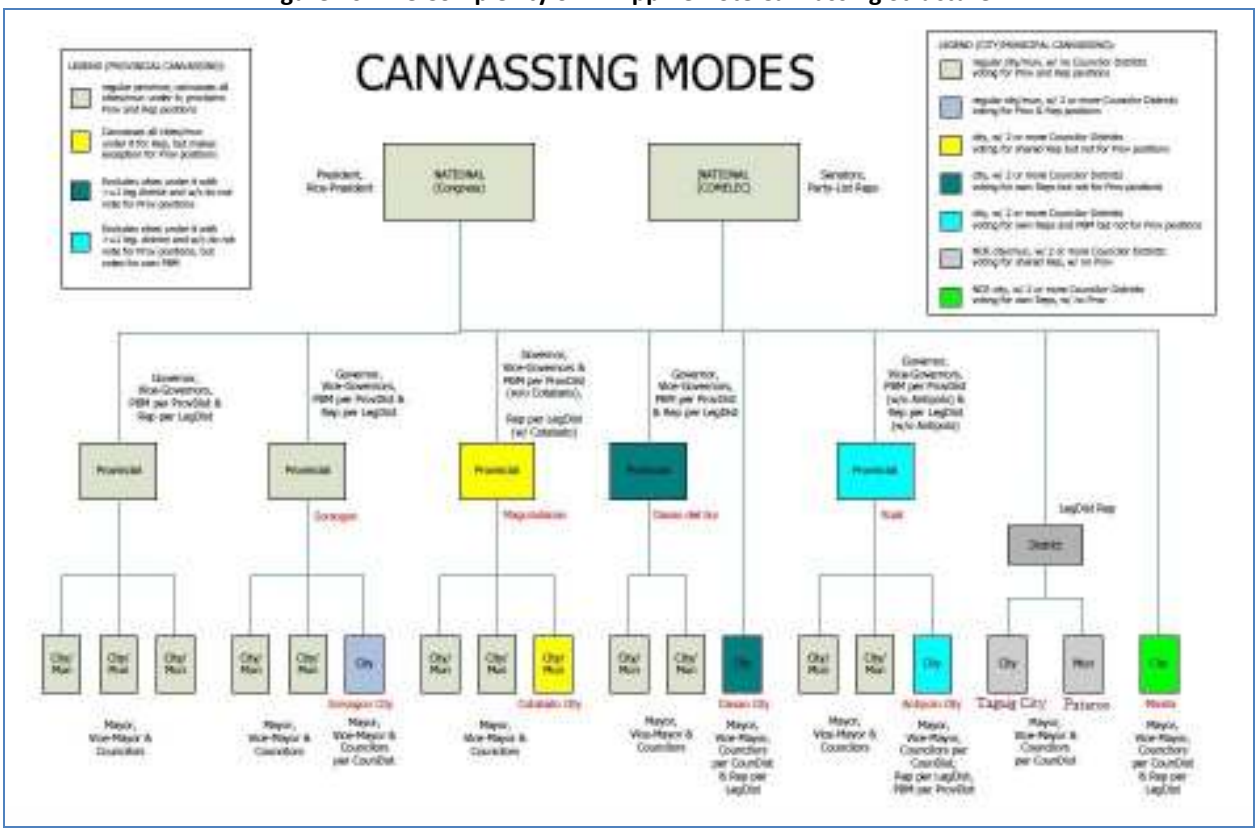
The Philippine elections have about 12 canvassing modes.

- There are two (2) different modes in the national canvassing level.
  - The first one is conducted by Congress to canvass the votes for president and vice-president and proclaims the winners.
  - The second mode is conducted by the COMELEC Commissioners as the National Board of Canvassers (**NBOC**) to canvass the votes for the senators and the party-list organizations and proclaims the winners.
- The canvassing in the provincial level is conducted by the provincial board of canvassers (**PBOC**). The PBOC canvasses the votes coming from all the municipalities and cities of the province for the positions of provincial governor, provincial vice-governor, provincial board members, and the representatives from the legislative districts of the province. The votes for the national positions are also canvassed at this level.
  - The first mode is the regular canvassing of the votes for these provincial positions and the representatives.
  - In the second mode there are some positions for the provincial board members and district representatives where the votes for these positions are no longer canvassed in the provincial level since the canvassing was done in a city with 1 or more legislative district. For example, the City of Antipolo in the province of Rizal has two legislative districts and the voters voted for their provincial board members and district representatives. Although they also vote for the positions of governor and vice-governor.

- In the third mode there are cities that are located inside the jurisdiction of a province, but the voters in these cities do not vote for the positions of governors and vice-governors. However, they voted for their own district representatives representing the city. Examples are Lapu-Lapu City and Cebu City in the province of Cebu.
- In the fourth mode the voters of a province share with the voters of a city in voting for a district representative. But the voters in the city do not vote for the provincial positions. Examples: Maguindanao and Cotabato City.

- There are seven canvassing modes for the city and/or municipal level. In this level, the canvassing and proclamation of winners for the positions of mayor, vice-mayor, members of Sangguniang Bayan (SB) or Sangguniang Panlungsod (SP) are conducted. The canvassing of votes for the district representatives and the national positions are also being done. There are some cities that do not vote for the provincial positions.
  - The first mode is the regular canvassing of votes of a city or municipality with 1 district for SP or SB voting for provincial and district representative positions.
  - The second mode is the canvassing of votes of a city or municipality with 2 or more districts for SP or SB voting for provincial and district representative positions.
  - The third mode is the canvassing of votes of a city with 2 or more districts for SP or SB voting for shared district representative but not for provincial positions.
  - The fourth mode is the canvassing of votes of a city with 2 or more districts for SP or SB voting for own district representatives but not for provincial positions.
  - The fifth mode is the canvassing of votes of a city with 2 or more districts for SP or SB voting for own district representatives and provincial board members.
  - The sixth mode is the canvassing of votes of an NCR city or municipality with 2 or more districts for SP or SB voting for shared district representatives. For example, Taguig City and Pateros.

Figure 10: The Complexity of Philippine Vote Canvassing Structure

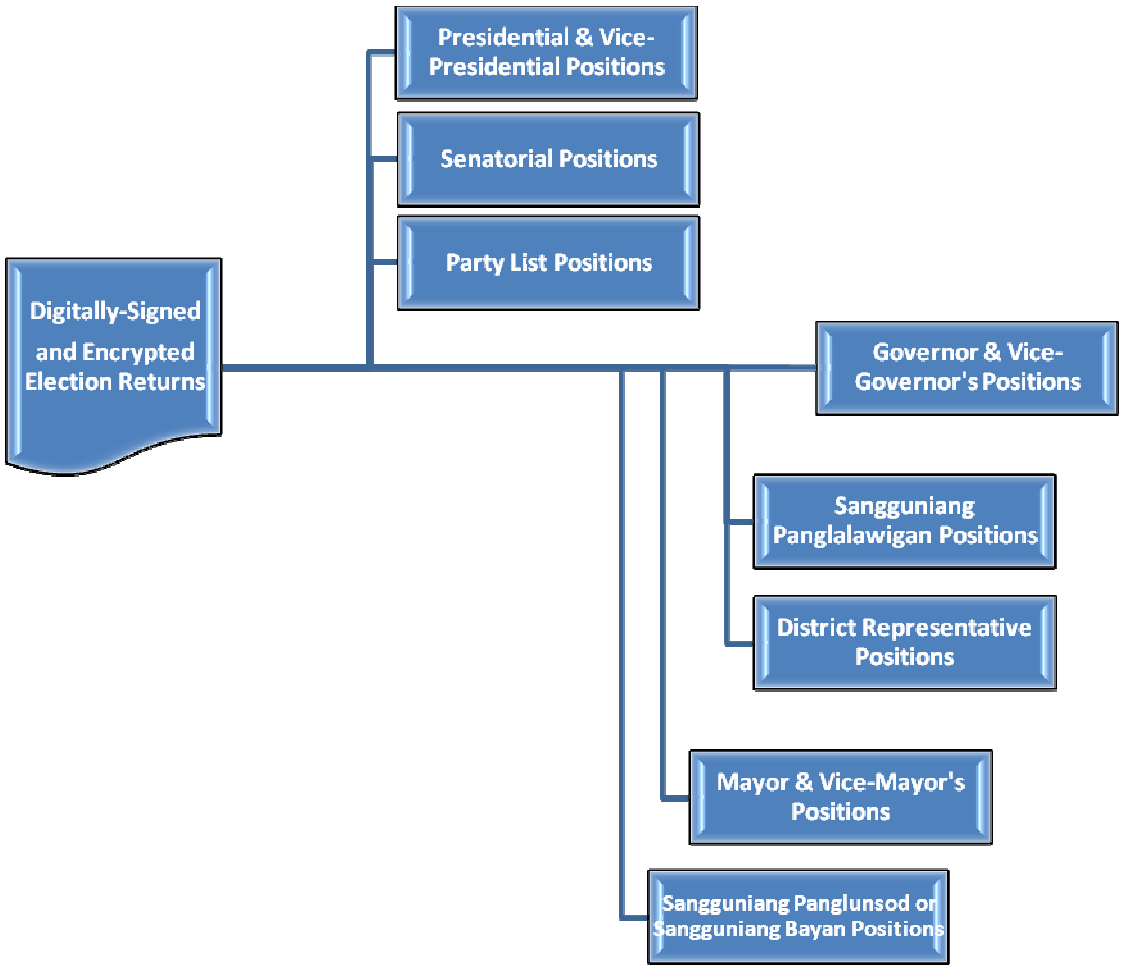


Board of Election Auditors and Board of Election Canvassers

There shall be a board of election auditors at each canvassing level that shall certify the results of the canvassing conducted by the board of election canvassers. Auditing the election results and the board of auditors shall be independent of the **COMELEC**.

There shall be a **National Board of Election Auditors (NBEA)** which shall audit, correct and certify the election results of the Presidential, Vice-Presidential, Senatorial and Party List Contests. It shall be composed of **15 members** with 1 chair and 2 vice-chairs. It shall be chaired by the **Chairman of the Commission on Audit.** The members shall be Filipino citizens of good moral standing coming from the civil societies, academe, and professional and other sectors. The **NBEA** shall be partitioned into three groups during presidential elections and into two groups during midterm elections. The Chairman of the NBEA shall chair the **National Board of Election Auditors** for the Presidential and Vice-Presidential Contest. The two Vice-Chairs of NBEA shall head each of the National Board of Election Auditors for the Senatorial Contest and for the Party List Contest.

Figure 11: Proposed Flow Chart for Canvassing the Votes, Certifying the Results and Proclaiming of the Winners



We propose the structure given in **Figure 11** for canvassing the votes, certifying the results, and proclaiming the winning candidates in the automated election system.

The **National Board of Election Auditors** shall organize and supervise the local board of election auditors.

The **board of election canvassers** shall canvass the votes and proclaim the winning candidates of the contests in its jurisdiction. The **board of election auditors** shall conduct an audit, correct if an error is found, and certify the correctness of the election results. The correction shall be based on the **accepted binary representations** since these represent the actual decisions of the voters. Proclamation of winners by the **board of election canvassers** shall be done only after the **board of election auditors** shall certify the correctness of the election results.

There shall be a **Special District Board of Election Canvassers** for canvassing the votes from the **ERs** for the district representatives representing **Taguig-Pateros**, **Maguindanao-Cotabato City**, and the like. Hence, a **Special District Board of Election Auditors** shall also be organized for these places.

The decision of the voters in a precinct is recorded in the Election Return (**ER**). It shall consist of the following:

1. list of candidates with their respective votes,
2. the sequence $V_{11}V_{12}\dots V_{1n_1}V_{21}V_{22}\dots V_{2n_2}\dots V_{k1}V_{k2}\dots V_{kn_k}N_1N_2\dots N_kE_1E_2\dots E_k$, whichis the decimal representation of the **ERs** as the decision of the voters in the precinct, and
3. the list of sequences of the form $v_{11}v_{12}\dots v_{1n_1}v_{21}v_{22}\dots v_{2n_2}\dots v_{k1}v_{k2}\dots v_{kn_k}n_1n_2\dots n_ke_1e_2\dots e_k$which is the accepted binary representation of an accomplished ballot.

The **ER** shall be digitally-signed and then encrypted by the assigned **Board of Election Inspectors**. The **ERs** in the precincts of highly-urbanized cities shall be transmitted to the **city election server** and to the **national election server**. The **ERs** in the municipalities and component cities of a province shall be transmitted to the **municipal/city election server**, to the **provincial election server**, and to the **national election server**.

As the **ERs** are received by the national election server, it shall be decrypted and the **digitally-signed ERs** shall be posted immediately in the website dedicated for posting the ERs. The **decrypted and digitally-signed ERs** shall be sent also to the servers of all the political parties and party-list organizations that participated in the election, and to the civil society groups that are accredited by the COMELEC.

If a PCOS machine fails to transmit its **ER,** then its transmission shall be done at the town or city canvassing center.

The **Municipal Board of Election Canvassers** shall canvass the votes in the **ERs** for the positions of **Mayor**, **Vice-Mayor** and **Sangguniang Bayan**. **Municipal Board of Election Auditors** shall audit, correct if there are errors, and certify the correctness of the election results before the proclamation of the winning candidates in the municipality. Note that if there are errors the correction shall be based on the accepted binary sequences each of which represents the true decision of the voters in the precinct.

The **City Board of Election Canvassers** shall canvass the votes from the **decrypted and digitally-signed ERs** for the positions of **CityMayor**, **CityVice-Mayor,** and **SangguniangPanglunsod**. If the city has **1** or more legislative districts, it shall canvass also the votes for the position of the **district representative or representatives**. If the city has 1 or more districts for **Sangguniang Panglalawigan**, it shall also canvass the votes of this contest. The **City Board of Election Auditors** shall audit, correct if there are errors, and certify the correctness of the election results before the proclamation of the winning candidates in the city.

The **Provincial Board of Election Canvassers** shall canvass the votes from the **decrypted and digitally-signed ERs** for the positions of **Governor**, **Vice-Governor**, **Sangguniang Panglalawigan** and the **District Representative or Representatives**. The **Provincial Board of Election Auditors** shall audit, correct if there errors, and certify the correctness of the election results before the proclamation of the winning candidates in the province.

The **Special District Board of Election Canvassers** shall canvass the votes from the **decrypted and digitally-signed ERs** for the position of district representative in each of the specially constituted

legislative district of Taguig-Pateros, Maguindanao-Cotabato City, etc. The **Special District Board of Election Auditors** shall audit, correct if there errors, and certify the correctness of the election results before the proclamation of the winning candidates in the special district.

The canvassing of votes from the **decrypted and digitally-signed ERs** for the positions of President and Vice-President shall be done by Congress. The correctness of the result shall be certified by the **National Board of Election Auditors** before the winners shall be proclaimed.

The COMELEC shall function as the **National Board of Canvassers (NBOC)** for the positions of senators and the party list representatives. It shall canvass the votes from the **decrypted and digitally-signed ERs.** The proclamation of winners shall be done after the National Board of Election Auditors for the Senatorial Contest and for the Party-List Contest shall certify the correctness of the election results.

Conclusion

In this paper, we state that the measure of accuracy of a PCOS machine for the Automated Election System shall be based on its determination of the absence or presence of a mark on a vote position.

Thus, the vote counting accuracy rating of the PCOS machine is bound above its vote position accuracy rating and is limited by its ballot configuration.

Also, the decision of a voter which is reflected in his or her accomplished ballot can be represented uniquely by a binary sequence. This representation can be expanded to include the nullified votes and the empty votes by suffixing two decimal sequences representing the nullified votes and the empty votes.

The length of the binary sequence is equal to the number of vote positions or candidates listed in the ballot. The length of the sequences representing the nullified votes and the empty votes is equal to the number of contests listed in the ballot.

If the voter verifies and accepts his or her votes as displayed in the monitor of the PCOS machine, the accepted binary representation is the actual decision of the voter. Hence, counting of votes can be done by adding the corresponding terms on these accepted binary representations.

The number of nullified votes and the number of empty votes in the precinct can also be determined by adding the corresponding terms of the decimal sequences.

The **Election Returns** of each precinct shall consist of the total number of votes in the precinct for each candidate, the decimal representation of the Election Returns, and the list of accepted binary representation including the representations for nullified and empty votes. These **ERs** are digitally-signed and then encrypted before transmitting them.

The board of canvassers at the local level shall canvass the votes from the **decrypted and digitally-signed ERs** for the positions at the local level.

The national board of canvassers shall canvass the votes from the **decrypted and digitally-signed ERs** directly.

Before the proclamation of winning candidates, the Board of Election Auditors shall conduct an audit on the **decrypted and digitally-signed ERs** make corrections if errors are found, and certify the correctness of the election results.

References:

1. http://en.wikipedia.org/wiki/Optical_mark_recognition#cite_note-12
2. Haag, S., Cummings, M., McCubbrey, D., Pinsonnault, A., Donovan, R. (2006). ***Management Information Systems for the Information Age*** (3rd ed.).Canada: McGraw-Hill Ryerson.
3. http://www.smartmatic.com/fileadmin/users/docs/SAESSAES1800_technicalsheet_v2.0.pdf
4. Batas Pambansa Bilang 881, ***Omnibus Election Code of the Philippines***.
5. http://www.comelec.gov.ph/modernization/2010_natl_local/technical_specifications.html#component1b
6. Republic Act 9369, an act amending R.A. 8436 entitled "AN ACT AUTHORIZING THE COMMISSION ON ELECTIONS TO USE AN AUTOMATED ELECTION SYSTEM IN THE MAY 11, 1998 NATIONAL OR LOCAL ELECTIONS AND IN SUBSEQUENT NATIONAL AND LOCAL ELECTORAL EXERCISES, TO ENCOURAGE TRANSPARENCY, CREDIBILITY, FAIRNESS AND ACCURACY OF ELECTIONS, AMENDING FOR THE PURPOSE BATAS PAMBANSA BLG. 881, AS AMEMDED, REPUBLIC ACT NO. 7166 AND OTHER RELATED ELECTIONS LAWS, PROVIDING FUNDS THEREFOR AND FOR OTHER PURPOSES", page 1, January 23, 2007.
7. "Principles and Best Practices for Post-Election Audits", www.electionaudits.org, page 6, September 2008