THE INDIA FORUM
A JOURNAL-MAGAZINE ON CONTEMPORARY ISSUES

April 13, 2021

Vulnerabilities in the EVM-VVPAT Process

The Potential Threat to the Integrity of Elections

By: Kannan Gopinathan

The 'jugaad'-like manner in which the VVPAT-machine has been inserted into India's electronic voting process has endangered security; the process safeguards that were earlier designed to protect EVMs are now of little value.

If EVM-VVPATs are stand-alone machines not connected to any external device, as repeatedly claimed by the Election Commission of India (ECI), how does the Voter-Verified Paper Audit Trail (VVPAT) machine print the name and symbol of the chosen candidate? When and how are the names and symbols of the candidates uploaded on to the VVPAT? Does this affect the technical, physical and procedural security claims of our electronic voting process?

These were the questions that troubled me as a bureaucrat who was a District Election Officer/Returning Officer during the 2019 Lok Sabha elections and a technocrat with a background in Very Large-Scale Integration (VLSI) design of integrated circuits. As recently as in 2017, I had publicly defended the EVM-based election process as tamper-proof. This was based on my experience of being part of the 2013 Delhi Legislative Assembly elections as well as the 2014 Lok Sabha polls. But the introduction of VVPATs has introduced potential vulnerabilities into the process as I discuss in this article.

Before I present my arguments, I would like to make two categorical statements:

1. This article does not allege that any manipulation of EVMs has taken place. I am instead hoping to raise certain process concerns, which are relevant irrespective of the results of the ongoing state assembly elections.

2. This analysis takes EVM-VVPATs as black boxes as the ECI has refused to divulge the details of their design.

The EVM process before VVPATs were introduced relied on both technical and administrative (physical and process) safeguards to claim that the entire election process was tamper-proof. The ECI argued that the integrity of the election process did not depend solely on whether an EVM can be hacked in a purely technical sense, but on whether this could be done in a tightly secured physical environment by circumventing the process safeguards designed for various stages.

To elaborate, the ECI claims that the Control Unit (CU)-Ballot Unit (BU) combination is a rudimentary device with a pre-defined and unalterable function, that it does not interact with the outside world and that it is programmable only once. The EVM in a way was like a simple calculator—a stand-alone device that does not talk to any other device and carries out a certain pre-designed function and therefore cannot be manipulated. Though the ECI never made the design transparent nor gave any evidence to show that the EVM was a stand-alone one-time programmable device, the calculator analogy conveyed the message and convinced people.

> The process security aspects were considered to be the impenetrable armour of our electronic voting process.

When it was argued that given physical access even a calculator can be hacked, the counter-argument was that the EVMs were under physical security. If the EVMs were always kept under the protective custody of election officials, with CCTVs and armed police guarding strong rooms with double locks, how can you have physical access? And if you don’t have access, how can you hack an EVM? But physical security is defenceless against pre-programmed EVMs that are brought to the district and then activated based on a pre-decided number pattern. The ECI countered this argument by saying that this could not happen because of the process security aspects of our electronic voting process. We were relying on a combination of process checks, namely, (i) the candidate-agnostic nature of the EVMs, (ii) a two-stage randomisation of EVMs and (iii) three-stage mock polls in the presence of representatives of the political parties or candidates These process security aspects were considered to be the impenetrable armour of our electronic voting process.

But how impenetrable are these process checks—the foundation of our election security—and how does the introduction of VVPATs change the analysis? For that we need to delve in detail into these three aspects of the process checks.

Candidate-agnostic nature of the EVMs

The candidate-agnostic nature of the EVMs means that the voting machines are not electronically aware at any stage of the voting process of the names of the candidates, party or candidate symbols, or their sequence on the ballot paper pasted on the BU. The EVMs just register votes corresponding to the serial number of the candidates. To make the voter aware of the sequence of candidates, a ballot paper with the names and symbols of the candidates as published in Form 7A (the ECI's 'List of Contesting Candidates') is physically pasted on top of the BU, matching the corresponding buttons on it.

Polling officials carrying the EVMs for the Nagaland Assembly Election at a distribution centre in Dimapur (February 26, 2018) | Election Commission of India

While counting the votes, the same sequence as published in Form 7A is used to decipher the candidate names. Form 7A is finalised close to the elections, by which time the EVMs have been secured in the strong room after the first level check (FLC) is conducted in the presence of political party representatives. This prevents the possibility of any intelligent pre-programmed manipulation of the election process. If the EVM cannot electronically know, either at the time of programming or at the time of activation, as to who is Candidate 1 and who is Candidate 2, it cannot intelligently transfer votes in favour of one from the other. It can blindly do so from 1 to 2, but that would serve no purpose as there is no way to know who Candidate 1 is and who Candidate 2 is until the candidate list has been finalized, published in Form7A and pasted on the BU.

Two-stage randomisation

EVMs are randomly assigned to constituencies and then to polling stations. These two stages of randomisation, it was argued, blunted the possibilities of targeted EVM manipulation through prior programming before they were stored in district strong rooms/warehouses. This is because the information as to which EVM will go to which constituency and to which booth would not be known prior to randomisation.

> More than anything it was the three stages of mock polls that instilled trust and confidence in the political parties and the election candidates.

In the first randomisation process, which is done after the EVMs arrive at the district warehouses and the FLC is completed, the CUs and BUs are randomly assigned to assembly constituencies within the district. For assembly elections this was particularly effective as you couldn't do targeted constituency specific manipulation as you couldn't have known or influenced which EVM was going to be assigned to which constituency.

In the second randomisation process, CUs and BUs are clubbed and then these BU-CU units are randomly assigned to polling stations. The second randomisation is done immediately after the candidates are finalized and before the EVMs are commissioned. The commissioning of EVMs during the pre-VVPAT era used to refer to the setting of the number of candidates in the BU and CU, and pasting the ballot paper on the BU. Since only at this stage are EVMs allocated to specific booths, randomisation as a process check made targeted manipulation of a particular booth very difficult.

Three-stage mock polls

More than anything it was the three stages of mock polls that instilled trust and confidence in the political parties and the election candidates. For, here they could cast votes and verify the results themselves. A key principle in electronic voting is that any undetectable change in software should not lead to an undetectable change in the outcome. The mock poll is one way to verify this principle.

The mock poll lets the stakeholders cast their votes in randomly selected EVMs and verify for themselves if there has been any tampering. The first mock poll is done along with the FLC of the EVMs. “FLC of EVMs shall be completed, as far as possible, well before the issue of notification calling the election.”, says the ECI’s direction. Representative of all national and state-level recognised political parties are informed and encouraged to be part of this FLC process. Here, immediately after the engineers/staff of Bharat Electronics Ltd. (BEL) or Electronics Corporation of India Ltd. (ECIL) complete cleaning the machines, carry out visual inspections and full functionality checks of the EVMs, the mock poll is conducted.

> A key principle in electronic voting is that any undetectable change in software should not lead to an undetectable change in the outcome.

As part of the mock poll at this stage, 1 (one) vote is cast against each of the 16 candidate buttons in each EVM and the result is verified. More importantly, a mock poll is conducted with 1,200 votes cast in 1% of EVMs, 1,000 votes in 2% of EVMs and 500 votes in 2% of EVMs, and the results shown to the representatives of political parties. Thus, the representatives of political parties are allowed to randomly pick a total of 5% of EVMs to cast 500 or more votes in a mock poll and they can verify the results. The ECI further suggests that even in the rest of the machines, the number of votes polled during a mock poll at this stage should be to the satisfaction of the representatives of political parties. And that these representatives should be allowed to conduct the mock poll themselves if they so desire. If there are discrepancies at this stage, then the EVM is marked as defective and rejected.

EVMs for use in the 2014 General Elections at the Dumurjala distribution centre in Howrah (29 April, 2014) | ECI

The second mock poll is conducted immediately after the EVMs are commissioned. That is, after the candidates are finalised , the second randomisation and assignment of booths to EVMs has been done, and the ballot paper has been pasted on the EVMs. The commissioning of the EVMs, which is the preparation of the machines for voting in that particular constituency, is done in a highly secure environment and during the pre-VVPAT era no electronic device was allowed inside the commissioning room. Immediately after the ballot paper is fixed on the EVM, a mock poll is done in the presence of the candidates or their representatives, in which 1,000 votes are cast in 5% of the EVMs that are randomly selected. This mock poll is very seriously taken by all contesting candidates as it is the first time that they verify the EVMs with the candidate names and symbols pasted alongside the buttons on the BU.

Having been part of the election process during the pre-VVPAT era, I was also convinced about the effectiveness of the process and had defended it at the time..

The third mock poll is done on the day of the poll, when a minimum of 50 votes is cast in the presence of the polling agents and the results verified in each and every polling booth. Though this is the most extensive mock poll, it is arguably the weakest as well. Since it is known upfront that a fixed number of votes cast in the beginning of the polls in each polling station will be part of the mock poll, theoretically a hack can easily bypass the first few votes, thereby preventing detection of foul play.

These safeguards in combination with other aspects of the election process like participation of the representatives of parties/candidates during the opening and closure of strong rooms, secured transportation of EVMs etc., were used to argue that there was a fool-proof system of electronic voting. Along with this, the forceful assertion of the ECI that the EVM is a rudimentary one-time programmable stand-alone device not connected to any external device helped instil confidence in both the electorate and the political parties.

Having been part of the election process during the pre-VVPAT era, I was also convinced about the effectiveness of the process and had defended it at the time.

Enter the VVPAT

The implementation of the current VVPAT design as a patch work on the existing EVMs has compromised various security features of the pre-VVPAT era and has shaken my confidence in the electronic voting process.

Being the Returning Officer for one of the parliamentary constituencies, I had to attend the training of Returning Officers, conducted by the Election Commission in the India International Institute of Democracy & Election Management (IIIDEM), as part of the preparation for the 2019 Lok Sabha elections. This was the first time that the VVPAT was going to be used nationwide in a Lok Sabha election. We were told that the VVPAT would print the name and symbol of the chosen candidate as and when the voter pressed the button in the BU as an additional transparency measure for voters.

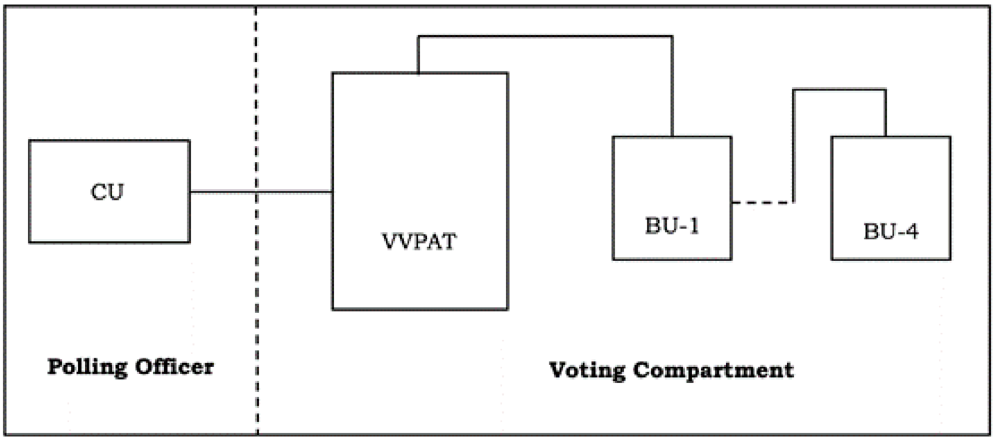
[T]he current VVPAT design as a patch work on the existing EVMs has compromised various security features of the pre-VVPAT era and has shaken my confidence in the electronic voting process.

This seemed to go against the candidate agnostic nature of EVMs and I was immediately curious as to how the VVPAT could know which candidate's name and symbol it is to print when a voter casts her votes on the BU. I was told that the BEL/ECIL staff would upload the sequence, names and symbols of candidates to the VVPAT as part of the commissioning process. This raised a barrage of questions immediately: How is this data uploaded? The EVM-VVPAT has to be connected to an external device for the data upload. How can we say then that it is a stand-alone device? Is this connection wired or wireless? What is the communication protocol? Who makes the software for this? What device is it connected to? Where is this information about the names and symbols stored? Does it mean it has programmable memory? Will it not then seriously weaken the ECI's own arguments?

Of course, my questions were not answered at the time.

This introduction of VVPATs and the associated changes in the voting process should have ideally led to a complete re-evaluation of all administrative and technical safeguards. But that did not happen. Let us try to investigate and understand the issues. Since the ECI has refused to reveal any design/code/schematic-related information on the EVM or VVPAT, we will be doing a black-box analysis with publicly available information in the Election Commission's own documents.

The EVM-VVPAT has to be connected to an external device for the data upload. How can we say then that it is a stand-alone device?



The Ballot Unit-VVPAT-Control Unit Design

The first and foremost question related to the VVPAT is its location in this new EVM-VVPAT design. Where does it fit into the erstwhile CU-BU system that was designed not to interact with any other device, as repeatedly claimed by the Technical Expert Committee (TEC)?

As currently implemented, the VVPAT sits right in between the BU and CU and is connected to both (Figure 1). This essentially means the vote flow is from the BU to the VVPAT and then from the VVPAT to the CU. The vote is now being cast in the CU by the VVPAT and not directly by the BU . Such a compromised design was possibly adopted to ensure the use of the existing BUs and CUs—some kind of a *jugaad*. But this introduced a serious vulnerability into the integrity of the election process. There is now a programmable device in between the voter casting a vote and the CU recording that vote, thus potentially introducing a vulnerability. A key principle of the election process is "Cast as intended, recorded as cast and counted as recorded". Any vulnerability in the VVPAT is now a vulnerability in the entire electronic voting process. This design flaw, in my opinion, is serious enough that we should discard the current design for this reason alone.

> Any vulnerability in the VVPAT is now a vulnerability in the entire electronic voting process.

Further analysing the technical safeguards, it is clear that the EVM as a BU-CU-VVPAT combination is no more a stand-alone machine, for we now need to upload symbols, names and the sequence of the candidates to the VVPAT. That means the VVPAT has to be connected to an external device.

According to the BEL document available in the public domain, and from my own experience overseeing the 2019 Lok Sabha elections, symbol loading is a three- step process.

First, the VVPAT sheet has to be created on a laptop through a symbol loading application that is either downloaded from the ECI server or copied from another local device. The next stage is to upload this VVPAT sheet to a device called the Symbol Loading Unit (SLU) through a 9-pin serial cable. And the final stage is connecting these SLUs to the VVPAT through a 9-pin serial cable for uploading the symbols to the VVPAT. The details of the symbol loading application and the details of the SLU are not available in the public domain. We also need to be aware of the fact that we have to connect each VVPAT to an external device for every constituency before every election, and this has to be done after the candidate sequence is finalized and the names of the candidates and the symbols are mapped to the BU buttons.

> The new EVM, i.e., the BU-VVPAT-CU combination, is neither a stand-alone nor rudimentary device. It is no more a simple stand-alone calculator.

Since the VVPAT has to be connected to the SLU before every election as the candidates' names and symbols will vary from constituency to constituency and election to election, this would also mean there has to be a programmable memory in the VVPAT where this information of the candidates can be uploaded. The VVPAT also has a printer unit that prints the VVPAT slip. There is also a light that gets activated only when the voter votes, a sensor that senses the falling of the paper slip into the box, etc. All of these require drivers, in addition to the micro-controller that controls all this. This effectively means that the VVPAT is not a rudimentary device as claimed so far about the CU, and thus rendering effectively useless in the post-VVPAT scenario the entire technical safeguard argument of the pre-VVPAT EVM process. The new EVM, i.e., the BU-VVPAT-CU combination, is neither a stand-alone nor rudimentary device, as one could claim earlier with the BU-CU design. It is no more a simple stand-alone calculator.

Coming to the administrative safeguard aspects of the process, the main safeguard earlier was obviously the physical security of the EVMs. But now we are giving access to the EVM itself. Physical access is provided through a serial connector to an external device, which is connected to a laptop, which, in turn, uses a symbol loading application either from the internet or otherwise to generate a VVPAT sheet. The laptops and SLUs that are brought to the commissioning room by the technical staff of BEL or ECIL are not under the custody of the District Election Officers prior to the commissioning of the EVMs. Further, with attacks like MITM (man-in-the-middle), even the technical staff need not be aware of what has gone into the laptop or what has been transferred to the SLU or the VVPAT. This renders invalid the earlier argument that any manipulation can happen only with the connivance of multiple election officials. It is entirely possible that neither the District Election Officers/Returning Officers nor the technical staff uploading the VVPAT sheet are aware of what is getting transferred between the devices.

> Once you provide physical access of the EVM to an external device, it does not matter whether there is armed security or there are CCTVs guarding the strong room for the rest of the time.

To put it in other words, there is no way the district election officials can know or verify what has been transferred between the VVPAT and SLU/Laptop. Once you provide physical access of the EVM to an external device, it does not matter whether there is armed security or there are CCTVs guarding the strong room for the rest of the time. The new process thus surrenders the erstwhile physical security argument.

Let us now see if the pre-VVPAT era administrative safeguards or process checks will be able to prevent or detect any tampering.

The first safeguard of candidate anonymity is no more valid as the VVPAT is now loaded with information of the names of candidates and symbols. The EVM-VVPAT as one unit is now well aware as to which vote is being cast to which party. That is the only way it can print that VVPAT slip.

The second process safeguard was the two-stage randomisation process for allocation of EVMs to constituencies and assigning them to individual polling stations. Here we have now introduced two vulnerabilities.

> [N]ow as soon as randomisation is done, anyone who has access to the EMS can remotely know and track which EVM is going to which constituency and to which polling station.

The first is not linked to the introduction of VVPATs but is the result of another thoughtless technological intervention in the election process. Randomisation used to be a localised process. Unconnected to the internet and done locally in each district. But with the introduction of the EVM Management System (EMS), both stages of randomisation are now done online through a centrally hosted software system with the ECI. What this means is that now as soon as randomisation is done, anyone who has access to the EMS can remotely know and track which EVM is going to which constituency and to which polling station.

The second vulnerability is that both the randomisations are done prior to the commissioning of EVMs. This implies that when the external device is connected to each VVPAT, it is already known as to which specific booth that particular VVPAT is going. This opens up the possibility of targeted polling station specific manipulation.

This makes the two stages of randomisation now useless as a process safeguard.

Polling officers standing in queue to collect EVMs at Tasi Nangial Academy School, Gangtok for the 3rd phase of the 2009 General Election in Sikkim (29 April, 2009) | ECI

The third safeguard and the most crucial in my view when the EVM was a stand-alone machine, was the three-stage mock polls done in the presence of representatives of the political parties or candidates. Now the first mock poll associated with the FLC is no more a safeguard as the external device is connected much after the FLC is done. It is akin to doing a full-scan first and then connecting the computer to the internet. You cannot rely on the previous scan anymore. Only those mock polls that are conducted after the EVMs are commissioned (during which time the VVPAT is connected to the external device) can claim to have any sanctity as a process check.

> [T]he robustness of the post-VVPAT electronic voting process can no longer be defended by making the same technical and administrative safeguard arguments as in the pre-VVPAT design flow.

As described earlier, we have two such mock polls after the EVMs are commissioned. One that is done immediately after commissioning and the other on the day of the poll. In the one that is done immediately after the EVM is commissioned, we cast 1,000 votes on randomly selected 5% of the EVMs. But here, the date, time and session information available with the EVM-VVPAT introduces a possible way to bypass the mock poll. The ECI itself claims that the date and time stamping of every key press has been done from the M2 (second) version of the EVM onwards, and that the VVPAT records the session information that is also printed on the VVPAT slip. As the day of actual poll is known at the time of connecting the VVPAT to the external device, it makes it easy, at least theoretically, to bypass the mock poll on any day other than the polling day with a simple date check. Any mock poll done other than on the polling day will therefore not be able to detect a potential manipulation.

The third mock poll, which is done on polling day, can again be easily overcome, as argued earlier, with a count check that will trigger the manipulation algorithm only after a certain number of votes are cast,

Whether or not this information is accessible programmatically, or whether or not changes can be made to the program through the serial connection etc., can only be authoritatively established or rejected if the ECI makes the design transparent and allows public scrutiny. Till such time we ought to be worried about the possible implications of connecting an external device to the EVM-VPVPAT (a) after the candidate list is published, (b) after both stages of randomisation are done, and (c) after the election date has been announced. We are thus surrendering all three process safeguards of the pre-VVPAT design. This is why the robustness of the post-VVPAT electronic voting process can no longer be defended by making the same technical and administrative safeguard arguments as in the pre-VVPAT design flow.

> [A] fundamental requirement for a VVPAT to be verifiable is that the voter should have full agency to cancel a vote if not satisfied…

All this does not mean I am arguing against the introduction of the VVPAT or arguing for a return to the pre-VVPAT EVM design. An EVM-only solution could not have guaranteed any of “the cast-as-intended, recorded-as-cast and counted-as-recorded” principles of a correct voting process. The introduction of the VVPAT was indeed a step in the right direction to ensure transparency, verifiability and correctness of elections. Unfortunately, the ECI cut corners, did *jugaads*, while adopting the VVPAT and this has meant that while we have surrendered the process safeguard aspects of the pre-VVPAT design, we have not benefited from a true VVPAT-based voting process either.

‘Verifiable’ and ‘Audit’ are the key words in the VVPAT. The VVPAT is supposed to be a machine that generates a paper trail that is individually verifiable at the time of voting and is publicly auditable at the time of counting. Prof Subhashis Banerjee of IIT Delhi, from whose public sessions on electronic voting I learned the principles of electronic voting, states that a fundamental requirement for a VVPAT to be verifiable is that the voter should have full agency to cancel a vote if not satisfied; and that the process to cancel must be simple and should not require the voter to interact with anybody. This implies that voter agency is important in the process of verification. Without it, a VVPAT can be manipulated in multiple ways to print and store slips without active verification by the voter. This is a fundamental flaw in our VVPAT design, for we rely on passive verification by the voter of the VVPAT slip. And not only does our VVPAT design deny an opportunity to the voter to actively verify the slip, but our legal framework also discourages her from complaining in the case of an observed discrepancy.

> [W]e have not only cut corners in the design of the VVPAT, we have also resorted to a ‘*jugaad’* auditing process.

Further, the test vote process that is available to the voter on complaint is both narrow in scope and discouraging in nature. As and when the voter complains that the VVPAT has printed and deposited a wrong slip, she is allowed a test vote in the presence of election officials. And if the paper slip gets printed correctly during this test vote, the voter can be fined/prosecuted. This test vote is highly restrictive as it essentially only checks if the very next vote has been manipulated the same way as the one which was cast. It rather naively assumes that any manipulation would be continuous and consecutive. If there is a discrepancy noted in one vote, it should necessarily be replicated in the next one. Thus, the test vote as a verifiability mechanism allows us to verify only if the very next vote is similarly manipulated or not. This in no way aids either the election officials or the voter herself to verify if the vote that she cast was indeed correctly cast as intended. The penalising legal framework reflects the ECI’s mistrust of the voter and blind trust in the machine, as the voter is punished for doubting the machine.

Finally, we have not only cut corners in the design of the VVPAT, we have also resorted to a *jugaad* auditing process. An election to a constituency is an independent event, be it a parliamentary or assembly constituency. The outcome in one constituency does not impact the outcome in another, and each outcome can be independently manipulated. Hence the decision to arrive at a sampling size for auditing should meet the statistical requirement irrespective of the election being a by-poll to a single constituency or a general election to the entire Lok Sabha.

An EVM distribution centre in Visakhapatnam (May 6, 2014) | ECI

To determine the sample size for the audit (deciding on the number of polling stations where the CU results are to be tallied with VVPAT slips) we should have taken each constituency as an independent population and should then have arrived at the sample size based on the required confidence level and the population size (the number of polling stations in that constituency). Instead, we initially adopted a sample size of one polling station per assembly constituency and later increased this to five polling stations per assembly constituency with no publicly available statistical backing for these decisions. K. Ashok Vardhan Shetty, a former officer of the Indian Administrative Service ,has shown in this article, how our auditing strategy fails to conform to fundamental sampling principles, how it leads to very high margins of error and how it defeats the very purpose of the introduction of the VVPAT.

> The ECI should ask itself as to what the voters would prefer: the correct candidate declared elected a day later or a wrong candidate declared elected a day earlier?

We have to be clear that with the process checks compromised and with the EVM-VVPAT being connected to an external device during elections, our only way to ensure the correctness of an election is through true verification at the time of polling and with adequate auditing at the time of counting. Without active verification of the paper slip by the voter, no amount of auditing of the paper trail can make us confident about the correctness of the election. But even if we accept that the passive verification that we have now is valid, it is unfathomable as to why the ECI would refuse to have an auditing strategy that meets statistical requirements.

The correctness of the election is so much more important than the speed at which it is conducted. And the Election Commission understands this very well. That is why it spreads the polling over multiple phases even stretching it to more than a month to ensure a safe and secure environment for voting. But what use is all that effort if it refuses to apply the same prudence during the counting of votes, with an untenable argument that there will be administrative inconvenience with associated delays. The ECI should ask itself what the voters would prefer: the correct candidate declared elected a day later or a wrong candidate declared elected a day earlier?

In addition to the above concerns, the introduction of the VVPAT has also created confusion as to what constitutes a ‘vote’. There are two items that express the will of the people now. One, the voter verified (albeit passively verified) paper slip stored physically in the VVPAT box and, two, the unverified but recorded vote stored electronically in the CU. Which one is the true representation of the will of the voter? What happens when there is a discrepancy between the two? These are again questions that ECI need to answer after careful consideration. For, if the VVPAT slip takes priority, then it will get defined as the ‘vote' and the natural demand would be to count every such ‘vote’. And if the CU count is the ‘vote’, then what is the purpose of a voter verified paper trail?

While I am tempted to give suggestions on the way forward, I believe that acting on ad-hoc suggestions of ‘let us do this and let us do that’ is what has landed us in the mess that we are in today.

Electronic voting is a fairly developed field. Its advantages and disadvantages are known. The Election Commission of India should arrive at a solution after due consultation with all the stakeholders and develop a transparent process that involves a peer review for vetting the design.

“Chasing two rabbits and catching none”, is not where the voting process of the largest democracy in the world should be.