# Remote Poll Site Voting

# Project Part II

# Preliminary Design

*Submitted by:*
*Anjali Prakash*
*Vivek Haridas*
*Kathryn Hayes*
*Neda Khalili*

## Introduction

Our goal is to create a new voting system. In order to generate acceptance towards an electronic voting system, we will produce a voting system that allows both paper and electronic ballots. Remote poll site voting will be enabled with the help of a token that will allow a voter to visit any poll site and still vote on their own county ballot. This will be done with localized databases – we have omitted the need to use a centralized database, which creates a dependency upon a single point that could crash the entire system if that one point fails. The system, presented in a model in Figure 1, is as follows:

## Registration

The way a voter registers today will still be valid: whether by mail, the DMV, or personally walking themselves to the Town Hall. When registered, each voter will be issued a unique voter identification number. Also, each voter will be assigned to a default poll site in their county where they will be expected to go to vote, although not required. Each poll site will have a list of all their default voters.

The assumption is that most people will want to vote at a poll site close to where they live (the default poll site). However, for those who would like to, or need to, vote at another site, they will have to pre-register for a remote poll site. This pre-registration – or perhaps to make it more clear, re-registration – will have to happen in person, whether at a Town Hall or specific sites in the county where this can be done. When registering to vote at a remote poll site, the voter will be taken off their default poll site list and will instead be handed a token. This token encodes a unique voter identification number, the voter’s name and the default poll site name. Some form of identification must also accompany the voter in order to vote (social security card, driver’s license, state ID, passport, birth certificate). This depends on the policy of the state in which elections are being held.

A voter with a token will be able to vote at any poll site in the state, including their default site. The information on the token will enable the poll site to look up the voter’s information and produce the proper ballot for their county. The token will be taken from the voter when leaving the poll site to prevent multiple instances of voting at other poll sites.

A voter without a token can only vote in their default poll site. Provisional voting will be available in case of “complications” [see Provisional Voting].

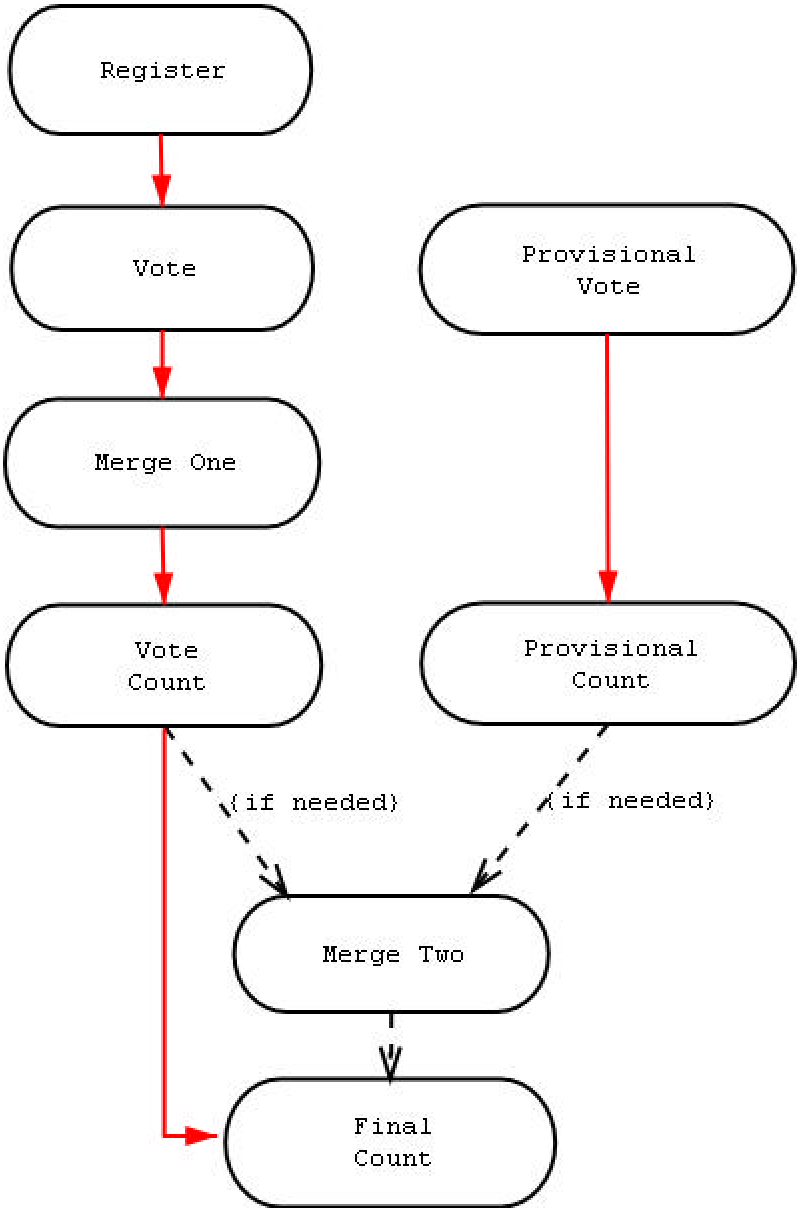

Figure 1: The Voting Process.

## Poll Sites

The poll sites will each have a CD of all the counties' ballots, for quick lookup for remote poll site voters. When a voter goes to a poll site to vote, they will be looked up in the local database for their voter ID. If it is their default site, the local machine will assign that voter to a voting booth (booth one, booth two, etc. – whatever is available next) and send the local ballot and the voter's ID to that booth (through a local connection). If the voter comes from another poll site, then the information on the token will be scanned for the voter ID and poll site ID, then sent to whatever voting booth the voter is assigned to (which includes the appropriate county ballot).

### The Voting System

The system we have chosen will incorporate both paper and electronic ballots. The voter will record their vote on a paper ballot, then an optical scanner will record the vote electronically. Both ballots will have "stamped" upon them a random number (originated by a random seed from the system) in order to connect the paper vote to the electronic vote. The paper ballot will then be dropped into a box to be used as the audit trail. The electronic vote will go on to be encrypted, and then it will be associated with the voter's ID number [see Encryption]. This is done in order to check for duplicate votes by the same voter (checked by the voter ID) [see Merging and Counting of Votes].

### Provisional Voting

If a voter arrives at a site without having their name on the poll site list and without a token (or with a faulty token), then provisional voting will take place. A temporary ID will be assigned to the voter (something that will not be confused with the unique voter IDs assigned, perhaps denoted by a negative sign to avoid a possible collision) at that poll site. The voter will still have to provide information and valid ID, for if the provisional votes are necessary in a final count, there will be a check for registration information (i.e. voter ID) from all the sites. If no registration information can be found, this provisional vote will be thrown out. In a way, this is *provisional* provisional voting; however, in this system, this is the only way to allow it.

The provisional votes will still be recorded the same way as regular voting – with paper and electronic ballots – but will be stored separately and only used in the final vote count if necessary. This means that they will be counted and merged only if the margin of victory in the election is less than the total number of provisional votes.

### Encryption

The encryption scheme being used is public key.

Each political party will have a public key that will be known to all the machines (distributed by CDs) and a private key that only one representative from the political party (for the entire state) will know. In addition, the Chief Justice (or another trusted individual) will have an additional public and private key pair. The public keys will encrypt the electronic ballot (and also the random number that links it to its paper ballot counterpart) while the collaboration of all the private keys will decrypt the ballot. The identity of the private keys will only be made known after the poll sites have shut down and the encrypted votes have been stripped of any links to the voter ID. Then, the Chief Justice will allow his private key to be known publicly, as will each member of the political party. The keys will be distributed to the counties that will be doing the voting [see System of Transfer of Votes]. They can be made public at this time because the danger of needing to encrypt the votes has passed at this point – it can no longer be traced to who made the vote.

Encryption is necessary only to hide the vote from its association to the voter ID. We need to link the voter ID to the encrypted vote in order to detect duplicate votes by the same voter. Once all the votes have been cast, the duplicate votes need to be sorted out [see Merging and Counting of Votes]. Once the duplicate votes have been sorted out, any trace of the voter ID to the encrypted vote needs to be removed. This is a method that will be further researched.

## Merging and Counting of Votes

In this implementation, votes need to be merged for the purpose of counting and to detect and remove duplicate votes. Votes from each polling site are tabulated and counted at a central location in each county [see System of Transfer of Votes]. Remote token votes are forwarded to their respective home counties for counting and merging.

Since the polling sites are decentralized, there is no way to check if a voter has voted multiple times at the time of voting. Because of this, it is necessary to merge the votes to detect duplicate voting.

A scheme is needed to detect multiple votes without revealing what the actual vote is. It would lessen the complexity of the problem if the identity of the voter who had duplicate votes would be revealed; however, more research is needed on this issue. The biggest issue is the vote: if it were to be revealed, it could lead to increased vote selling and extortion.

Each vote will have a unique random number and the combination of this number and the vote will be encrypted with multiple public keys.

What is needed is a function F that creates an association between the encrypted vote and the voter number to eliminate duplicates. When the votes are collected, duplicates can be identified without revealing the underlying vote.

The association needs to be dissolved before duplicate paper trail votes are detected and thrown out. After the association is dissolved, and the votes are decrypted with the private keys, the remainder is the votes and random numbers, which can be associated with the numbers printed on the paper ballots. In this way, both the electronic and the corresponding paper ballots can be thrown out.

## Storage of Votes

The decrypted votes should be stored on persistent storage for a long term. This is done for three purposes.

- The stored votes should be available for the counting. If a failure occurs before the counting, we can resort to manual counting of the votes (from the audit trail).

- They should be available after counting so that the accuracy of the optical scanners can be determined.
- They should be available in case the need for a recount arises. If a failure occurs before recounting is done, we can resort to manual counting of the votes (from the audit trail).

We should be able to detect if the electronic vote has been tampered with. This is especially important if the paper ballots are never checked – that is, a need for the audit trail does not arise. Consistent replication of data can be done to make sure that we have a complete copy of all the votes. Many checksum schemes exist to do consistency checks. Error detection and correction codes will be used to enforce consistency in replication.

To detect vote tampering we will use log-based storage. In such storage, data is never overwritten. An object (that contains data) can be modified, but the modification will be made on a different location on persistent storage (hard drives). Also, a log of the modification is kept in a separate log file. This makes it easier to detect if someone tampered with an electronic vote.

### System of Transfer of Votes

The voter and the public are the main components of the voting system. The data passes through the voters to the public through a small number of entities. This has to be done securely maintaining the integrity of the votes as well as the anonymity of the voters.

Figure 2 illustrates the working of this system.

The voters can vote from:

1. The default poll site
2. Any other poll site of the same county (after pre-registration)
3. Any poll site of the other counties (after pre-registration)

However, regardless of which poll site the voter votes, after the poll sites shut down, the votes will be sent to the voter's home county.

### Information Exchange Format

The information exchange system between the poll sites and the counties is based on public key cryptography. Key pairs are provided for each of the poll sites and the poll sites sign all data being sent from them.

The nature of data sent from every poll site is given as follows.
Consider counties A, B and C:

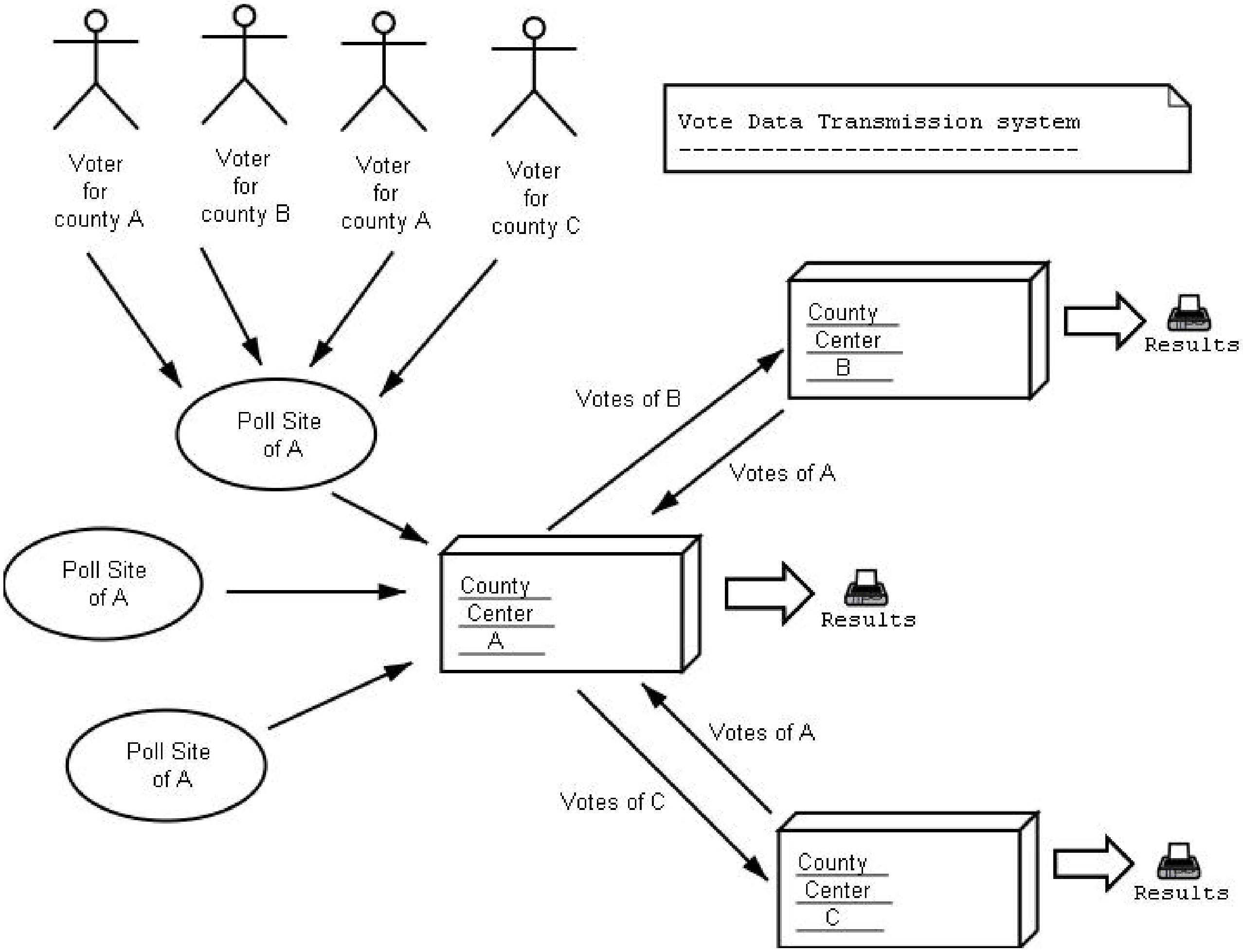

Figure 2: Transfer of votes

The format of data sent out of the poll sites is given below:

| | | |
|---|---|---|
| Data for county A signed by poll site | Data for county B signed by poll site | Data for county C signed by poll site |

Entire Data signed by poll site

The data stored for each of the foreign counties are signed separately at the poll site. This is to make sure of further transmission to the home county; any change in the individual data can be detected.

Further, the entire data is also signed to ensure the integrity of the information passing out of the polling station.

The data is passed out of the poll site either as magnetic mobile media or through the Internet. The data received at the home county is then checked for integrity.

The data meant for the foreign counties are then relayed to them. The signatures of the polling station ensure their integrity.

The data is then extracted at every home county and then used to perform the counting, the results of which are made public.

As the information relayed is already encrypted, there is no need to perform an extra encryption just for the transmission purpose.

## Finishing Thoughts

The system we have presented is still a work in progress. The heart of this system relies on a still unknown function F, which needs to break a link between the voter ID and the encrypted votes, without leaving any trace of such a link. However, we believe that once this function is realized (hopefully after some more research), we will be able to present a useful and functional electronic voting system.