PRIVACY PRESERVING IN ELECTRONIC VOTING

Ai Thao Nguyen Thi¹ and Tran Khanh Dang²

¹,² Faculty of Computer Science and Engineering, HCMC University of Technology
268 Ly Thuong Kiet Street, District 10, Ho Chi Minh City, Vietnam
Email: {thaonguyen, khanh}@cse.hcmut.edu.vn

Abstract

Electronic voting becomes a new tendency in many countries. It brings many benefits such as providing more convenience for voters to increase the potential voter turnout, reducing errors during counting phase, reducing costs and efforts, increasing democracy, etc. Benefits, however, always come along with challenges. The main concern is how to ensure the security of electronic voting systems, especially to preserve the voter privacy. In previous studies, researchers introduced a large number of requirements of electronic voting systems and provided the protocols satisfying as many requirements as possible. In this paper, we not only provide a comprehensive view of electronic voting systems but also analyze the privacy issues in the voting context. From these analyses, we introduce general architecture for privacy preserving in electronic voting, and then discuss possible techniques for addressing the related privacy issues.

Keywords: Cryptography, Electronic Voting, Internet Voting, Privacy, Secure Multiparty Computation

Introduction

In recent years, the Internet has demonstrated the exceeding growth rate which no previous technologies could obtain. Along with the rapid growth of internet using, many electronic commerce services have been brought into reality, replacing traditional ones which become more and more backward and time-wasting. Nowadays, people prefer the convenience of staying at home, buying things via Internet, making contract with partners online, or sometimes fulfilling governmental procedures without having to present at the administrative bodies like they used to. Being aware of those needs, the Governments worldwide have tried to digitalize their services as much as possible, turning them into electronic Government (e-Government) services.

Among e-Government services, electronic Voting (e-Voting) is one of the most important services for its characteristic and impact on the whole country. It is obvious that voting itself is the key of democracy. It is a regular activity of a nation in order for the citizens to raise their opinions on the political issues they are concerned about.

The traditional voting (also called voting by paper or signature voting [6]) is often held in public places under the supervision of Government officials to ensure the voting process fully meets the requirements set before election period. More concretely speaking, supervisors have to check whether the person who comes to the polling place is an eligible voter or not, they also have to guarantee that every valid voter casts his or her vote without any coercion or cheating. To complete this job on a large scale, it takes a lot of time, money and efforts, but nothing can assure there are no violations in this casting phase. In addition, when voters fulfill their duty in voting process, the Government officials have to count each paper ballot. The problem is that no matter how much efforts you make, human being cannot count a huge number of paper-ballots without a single mistake.

That is just one aspect of how problematic traditional voting could be. Another aspect relates to the other partner of voting process - voters. The drawbacks of traditional voting make voters feel inconvenient. What do you suppose to do if the Election Day is in your
working day and voting booth is far away from your office? In this case, you may feel that election is somewhat annoying, which results in your denying the right to vote. Voting thus is no longer the key of democracy. And that is exactly the reason electronic voting needs to be brought into reality.

At first, electronic voting system only included a machine which could automatically count the punch cards considered as ballots. You can learn more about the development of voting machine in [9]. However, that machine only helped election officials get the final result of voting process quickly, the registration and validation phase were still the same as traditional voting system. Voters had to go to the booths to cast their vote under the supervision of Government officials. These procedures thus consumed much resource. Therefore, the idea of Internet Voting in which voter only needs to send his or her vote to the election committee via internet appeared.

For its advantages, Internet voting starts drawing attention of not only Governments but also companies, organizations, etc. But as usual, great opportunities are accompanied by challenges. The first challenge that has to be addressed is security, especially privacy. How can we protect the content of votes against the vote-buying and the coercion? How can we conceal the identification of a voter from the curiosity of adversaries inside and outside the system? These are the questions that need to be answered before implementing the Internet voting system. Moreover, Internet voting is more vulnerable than other voting systems [7]. The second challenge is the practical aspect of Internet Voting system in which every paper is digitalized, leading to the fact that voters can no longer see where their votes are going to as they used to in traditional voting. Consequently, they may be suspicious about the final result of voting process.

The paper is organized as follows: in the next section we will provide the comprehensive view of electronic voting. Section 3 will discuss the privacy issues and other requirements of electronic voting. Section 4 will propose general architecture for privacy issues of e-voting system. Section 5 provides some techniques applied to the architecture above. And finally, conclusion and future work are included in section 6.

Overview of Electronic Voting

Definition of Electronic Voting

Electronic voting is simply the measure to capture the voter’s selections by digital data [7]. There are many voting forms all over the world depending on the purpose of voting such as: presidential election, election of a legislature, referendum … And each of the forms has its own way for voters to cast the ballot as well as for voting authorities to choose the winner.

- **One choice**: every voter has only one ballot to cast his or her vote. In every ballot, there is only one candidate chosen by voter. The winner is the one who gets the most credits, meaning he receives the greatest number of votes in the election.
- **Multi choices**: one voter can choose many candidates in his ballot at a time. In this model, there are three ways that voter can pick up their preferable candidates.
  - Non preferential voting: every choice is the same.
  - Ranked voting: a voter orders the candidate’s position in his ballot according to his preference.
  - Rated voting: a voter is allowed to compare candidates in quality and quantity. This means voters can give each of candidates a score showing how much they want this candidate to be a winner. In this kind of voting model, every assessment is absolutely independent. This model is more
flexible than previous ones, but it's so hard to calculate the final result that few scheme are designed to use it.

However, no matter what form the voting system belongs to, it always has four main phases for voters to complete their voting process.

- Registration: the citizen who has the right to take part in the voting process will register to the Government. If they are eligible, the election roll will be created and they will become voters.
- Authentication: before casting his vote, the voter has to prove himself to the system.
- Casting: the voters cast their vote.
- Tallying: All the votes will be counted at this phase. At the end of this phase, the final result will be announced to public.

Figure 1. Four main phases of voting process

Each of above phases can be implemented by physical or electronic procedures. If system is called electronic voting system, it has at least the tallying phase to be digitalized.

**Classification of Electronic Voting System**

From the historical point of view, voting system has been developing from centralized voting in a single Town Hall, to distributed polling place voting, and recently, to remote connectivity through the Internet, called Internet Voting. Briefly, electronic Voting is classified into two main types as we can see below.

*Polling place Voting*
Voters have to gather into some polling booths, and then cast their vote by the voting machines. Both voters and voting machines are also supervised by voting authorities. In this type of voting system, just two phases, casting and tallying, are digitalized; the others are still same as traditional voting system.

![Figure 2. Model for polling place voting system](image)

**Internet Voting**

There are no needs to waste any money or time to supervise the physical environment such as voting machines, polling-booths, ballots or voters… because the processes are done via the Internet. In this type of voting system, all of four phases are digitalized.

![Figure 3. Model for internet voting system](image)

**The privacy issues in voting system**

Privacy is the most important issue in electronic voting. Before finding out the way to preserve the privacy, we have to know what the privacy in voting context is. In [11], the privacy of voting system is defined that: “A secret ballot protocol is said to be private if the privacy of voters is preserved”. However, voter is not the only one component of voting system, another component that is not less important is the vote. Therefore, in this paper, we divide the concept of privacy of voting system into two aspects: privacy of vote and privacy of voter.

**Privacy of vote**

In this context, privacy means no one can know exactly what contents are in this ballot except voter. And if the voter speaks out what candidate he chosen, he still has no
evidences to prove what he said. There are two concepts involved in the privacy of vote: Perfect Privacy and Ballot Secrecy.

In [5], the authors introduced the concept of Perfect Privacy. It states that e-Voting system is called perfectly private if the information that an adversary maybe learn from a voter’s choice also can be learned if that voter chooses another candidate. In reality, voting process and system always have some leaks no matter how we prevent, so sometimes an adversary can obtain a little information from these sources. But if the voting system provides a perfect privacy, the information the adversary gets from the leaks or the corrupted voting authorities of the voting system cannot help him infer the true vote cast by the voter on Election Day.

Ballot secrecy was defined in [5] that: the e-Voting system is called perfect ballot secrecy if there is no way for anyone to prove to an adversary his or her real choice in the ballot even if the voter colludes with the adversaries. This property is also defined clearly in other way in [7][8]. According to these references, we can divide this property into two small requirements:

- Uncoercibility: the adversary cannot force any voters to vote for his own intention or to reveal their votes. This property is to avoid the attack of the outside threats. In this case, the coercker is the adversary.
- Receipt-freeness: with the receipt-freeness, no voter can prove the result of his voting process to the person wanting to buy the vote. This property is to avoid the attack of the inside threats. In this case, the voter himself is the adversary.

All these concepts above focus on the privacy of the individual votes. In [11], the authors introduced another requirement which preserves the privacy of a whole votes used to calculate the final result. That is fairness. It means that no one including voting authorities can get the intermediate result of the voting process before the final result is announced to public.

In summary, the voting system providing Privacy of Vote should keep the content of all votes unrevealed by the voting authorities who stay inside the system or even the voters themselves. More concrete speaking, the attackers outside the system cannot get the information even though they can collude with most of the voting authorities and coerce the voters.

**Privacy of voter**

In the definitions above, we can see that the voting system satisfying these requirements will protect the content of vote while the privacy of voter will emphasis the preserving privacy of voter’s identity. In [6], the authors introduced the anonymous requirement which has the closest meaning with the privacy of voter. It states that each individual vote cannot be linked to the voter who cast it.

However, if the voting system just ensures the anonymity of voter, some objects such as the voting authorities can read the plaintext of votes so they can know the intermediate result before the publication phase. This result can affect the final result of voting process because it influences on the voter’s psychology, therefore the final result can be deviated from the true desires of most voters.

Therefore, the voting system needs to achieve the privacy of vote and the privacy of voter to ensure the privacy of the whole system.

**Other requirements for voting system**

In [6], the authors provided a general view of electronic voting requirements, as followings
• Completeness: every voter who satisfies the voting policies of Government is always accepted to cast their vote.

• Integrity: It means that the votes cannot be modified, rejected, deleted, forgotten, or not be counted by any objects in the system (including the person who has a full access control). In other words, the votes will remain in original forms as they were cast by the voters in any cases. In [2], it is the important requirement because if a vote is modified, the final result can not reflect the true intention of voters though the computation processes make no mistakes.

• Robustness: this is an important requirement to take an electronic voting system into practice. In [11], the robustness means that no one (including authorities) can disturb or disrupt an election and the voting process of each voter is independent of others. Though an adversary tries to steal the private data of a voter to interfere with voter’s intention or to stop their own voting process in order to disturb others processes, the election still work till the last phase.

• Uniqueness: this requirement prevents a voter from casting more than one ballot. In [11], the uniqueness is defined that: “Every voter votes exactly one time”. Though, some voting protocols allow their clients to vote many times in Election Day, only their last vote is accepted by voting authorities.

• Verifiability: This is an essential requirement of electronic voting system. It helps the system take more credits of clients; therefore voters will do their voting duty more actively than they have to do in a voting system that they are not sure about the reliability of its output. According to [12], we have two types of verifiability:
  o Individual verifiability: In [7], this concept is also called atomic verifiability but the main meaning is the same. This requirement requires that every voter is able to check whether their vote is counted correctly or not. However, it seems very difficult to prove to voters that their vote is casted as intended while the receipt-freeness requirement is still satisfied. This is a remarkable problem which is discussed in [1].
  o Universal verifiability: It means that every voter who is interested in tally result can collect the information from voters and voting authorities to ensure that the final result is correctly computed from all the ballots casted by eligible voters.

• Convenience: if a voting system satisfies this requirement, it must help voters not only to cast their vote quickly in one session but also to complete their voting process with the special skills as least as possible.

• Efficiency: the computations used in the voting system labeled efficiency have to be done in the tolerable time, and any voter in such a system doesn’t need to wait for the others to complete their voting process in order to do his or her own.

• Mobility: it means the voters can cast their vote any time they feel comfortable and any place they are in Election Day. So they don’t need to pay any attention to the physical environment as they used to do with the traditional voting system. However, this requirement seems to collide with the uncoercibility because if they cast their vote in their private place, nothing can ensure they cast their vote without any coercion.

Architecture

As mentioned above, the voting process will take place in four main phases: Registration, Authentication, Casting and Tallying (Counting). Consequently, the architecture will have four main features representing the four phases. To ensure the security of e-voting system,
we add two more features undertaking the duties of preserving the privacy of voter’s identity and privacy of vote’s content as illustrated in Figure 4.

In this architecture, the voter will use his computer to access the Internet in order to perform voting process. At first, he sends request to Registration server which is responsible for asking Certificate Authority about voter’s personal information to check whether he satisfies voting policy or not. If the client is an eligible voter, Registration Server will send the certificate to voter. After that, voter sends this certificate to Privacy of Voter Server to transform his real ID into anonymous ID that will be used in later voting steps. To keep the identity of voter completely secret, no party in the system can acknowledge the link between the real ID and the anonymous one. It means that even the Privacy of Voter Server should not have any idea about the content of the real ID which it is going to blind. To satisfy this requirement, we can use blind signature technique introduced in the section below. Based on the anonymous ID and the corresponding blind signature, the voting authorities can check whether this voter is eligible or not without knowing his real identity.

![Figure 4: Architecture for privacy preserving in electronic voting](image)

To protect the content of the vote against being modified or revealed to anyone, we add the Privacy of Vote Server into the voting process. Privacy of Vote Server will blind the content of vote before voter starts his casting phase. We can use some of the typical cryptographic techniques introduced below to implement this feature. Another problem here is how to ensure the voting authorities do not decrypt the ciphertext of votes to modify the ballots or to gain the intermediate result. This problem can be solved by digital signature, hash function or threshold cryptography briefly introduced in [3]. Besides, in [4], we can find a simple method to avoid coerer and vote-buying which are important threats to privacy of vote. The simple method that does not require the complicated cryptographic techniques can be the recasting mechanism without sacrificing uniqueness requirement. This mechanism was first introduce in [3][4].

Finally, when the election period is over, the Tallying Server will perform. In this architecture, Public Board will announce the final result of the voting process. In some protocols, Public Board will update voter’s status for voter to verify that his vote has been casted as intended.
The cryptographic techniques

Privacy-preserving in electronic voting is an instance of SMC (secure multiparty computation) problem, where we want to find the outcome (final result) but we do not want to reveal the inputs (votes) to any party who jointly computes the final result. The idea of secure multiparty computation was first introduced by Yao in the paper [16]. To solve the SMC problem, many researchers created a protocol using cryptographic techniques. Therefore, in this paper, we will introduce three main cryptographic techniques to protect privacy in the voting context we discussed above.

Mix-net

The concept of mix-net was first introduced by Chaum [18]. A mix net is composed of a set of linked servers called mixes. Each of mixes only takes all the encrypted votes at one time, it will decrypt these votes by its private key, disorder and then send them to the next mix to do the same process. The process will be show as follows:

At first a voter will choose his most favorite candidate then show it on the ballot. After that, he will then encrypted it m times in succession by m public keys of m different mixes. At the tallying phase, all the votes are decrypted in reserve order of encrypting process by the private keys of each mix.

This technique can conceal the identities of voters and protect the content of votes before the final result is announced publicly. Besides that, a useful property of mix-nets is universal verifiability because the encrypted vote will become a plain-text after going through the mix-nets so it is easy to check the final result is correct or not. However, the main drawback of this technique is that it requires a large number of communication and computation costs. Moreover, if any linked servers of those in mix-net become an adversary, the final result will be wrong. That means the result of mix-net belongs to the operations of all linked servers. So far, there are no electronic mix net-based voting systems that are implemented especially on the large scale which requires a lot of mixes.

Homomorphic cryptography

Homomorphic cryptography is a type of encryption in which the result of an algebraic operation calculated by the plaintext is equivalent to the result of another algebraic operation calculated by the ciphertext. It is based on the homomorphism in abstract algebra. And this property has been used to create many secure voting protocols [10]. We denote $E(x)$ as the encryption of message $x$, the encryption is homomorphic if and only if by $E(x)$ and $E(y)$, we can calculate $E(x \oplus y)$ without decrypting neither $E(x)$ nor $E(y)$.

$$E(x) \otimes E(y) = E(x \oplus y)$$

In the expression above, the operation $\otimes$ is defined on the cipher space, while the operation $\oplus$ is defined on the message space. So if a voting system belongs to homomorphic encryption model, we don’t need to decrypt each of encrypted votes to find out the final result of voting process. Therefore, the plaintext of vote is never visible to anyone else except the voter himself. In other words, the protocol using this kind of technique can guarantee the privacy of individual votes while retaining the universal verifiability.

There are many protocols based on this technique. In this paper we just introduce three main protocols.
**Benaloh protocol**

This is one of the earliest models using the special property of homomorphic cryptography suggested by Benaloh in [15]. In this model, the voters only cast the Boolean vote such as “yes/no” for technical reason. They will divide their votes into n parts to share them between n voting authorities so that (t+1) out of these voting authorities can recover it. Then, each voting authority computes its encrypted share of the final result. Lastly, at least (t+1) voting authorities cooperate with each other to recover the final result.

The homomorphic encryption used in Benaloh protocol is

\[
E(x_1) \ast E(x_2) = E((x_1 \ast x_2) \mod r)
\]

We have the public key is the modulus \(m\) and the base \(g\) with a block size of \(r\), and the encryption of a message \(x\) is

\[
E(x) = (g^x \ast u^r) \mod m
\]

We can find out that

\[
E(x_1) \ast E(x_2) = (g^{x_1}u_1^{\ast x_1}) \ast (g^{x_2}u_2^{\ast x_2}) = g^{x_1+x_2}(u_1u_2)^r = E((x_1 + x_2) \mod r)
\]

**Cramer et al protocol**

The model of Benaloh is quite simple, but it has a high communication cost because each voter must cast his or her vote over \(n\) communication channels. Consequently, researchers found out another model which was optimal the communications between voters and voting authorities. In this model, the voters send their encrypted votes to a single combiner. Then, the combiner computes the encryption of final result by using the homomorphic property of cryptosystem. After that, the combiner forwards that result to the voting authorities and (t+1) out of them can recover the plain text of final result by using a threshold cryptosystem. More detailed information about threshold cryptosystem can be found in [13].

However, the encryption using in this model has a relatively high computational complexity when the number of candidates become large. If the number of voters is \(l\), the number of candidates is \(r\); the complexity is \(\Omega(l^{(r-1)/2})\) [7]. So researchers suggested another protocol which can decrease the computational complexity.

**Paillier protocol**

There are many protocols based on the cryptosystem of Paillier. This cryptosystem provides the more optimal decryption algorithm so it can decrease the complexity of computation process. The protocol of Paillier is the same as the protocol of Cramer model, but it uses another function to encrypt the private message.

We have the public key is the modulus \(m\) and the base \(g\) with a block size of \(r\), and the encryption of a message \(x\) is

\[
E(x) = (g^x \ast r^m) \mod m^2
\]

The homomorphic property is

\[
E(x_1) \ast E(x_2) = (g^{x_1}r_1^m) \ast (g^{x_2}r_2^m) = g^{x_1+x_2}(r_1r_2)^m = E((x_1 + x_2) \mod m)
\]

According to [8], the computational complexity of the protocol is linear in the size of Paillier modulus, and the number of candidates as well. That is

\[
\Theta(p \times |N|),
\]

\(p\): number of candidates

\(N\): the size of the modulus used in the Paillier cryptosystem
The protocols based on homomorphic technique satisfied most requirements of privacy in voting context. However, they are still inefficient for large scale elections because computational and communicational cost for the proof and verification of vote’s validity is quite large when there are a lot of candidates. In addition, homomorphic encryption protocols cannot be employed on multi-choices voting forms because of the characteristic of homomorphism.

**Blind signature**

The idea of blind signature was first introduced by Chaum [17]. It stemmed from the need that a person wants to verify the valid of his document but doesn’t want the verifier to know anything about its content.

We can imagine the process of blind signature as followings: the Owner of a document needs an Authority (signer) to sign on his document in order to legitimize. However, the owner doesn’t want the authority to gain any information from the data written in this document. Therefore, he deals with this problem by a trick that owner will put his document into an envelope (it is a tool to blind content of the document). He will also put a carbon-paper into it and then, sending it to the Authority. After receiving the envelope, the Authority will check the name of Owner at the outside; if the owner is valid user, the Authority will sign on the predefined area of the envelope without opening it, and sending it back to Owner. Finally, owner will have signature of authority without revealing any information written in his document.

A simple method to implement blind signature scheme is using the asymmetric cryptosystem RSA. We have some notation

- m: the document needs to be signed.
- d: the private key of authority (signer)
- e, N: the public key of authority
- s: the signature of m.
- r: the random number that the owner’s document chose, and satisfied $\gcd(r, N) = 1$.

The RSA blind signature scheme is implemented as followings

1. **Step 1:** the owner blinds his document by the blind factor and then sends it to the authority
   
   $$m' = mr^e \pmod{N}$$

2. **Step 2:** upon receiving $m'$, the authority will compute a blinded signature $s'$ by using his private key $d$, then send it back to the owner’s document.

   $$s' \equiv (m')^d \pmod{N}$$

3. **Step 3:** owner will un-blind to get the signature $s$ when he receives $s'$

   $$s \equiv s' \cdot r^{-1} \equiv (m')^d r^{-1} \equiv m^d r^{-1} \equiv m^d (\pmod{N})$$

However, this technique just helps to anonymize the identities of voters. In this case, the content of each vote will be threatened. If a voting protocol does not have any proper methods to protect the content of votes, the intermediate result of voting process will leak out; and it is easy for an adversary both inside and outside the system to modify or remove the vote to make to final result incorrect.

There are many electronic voting protocols applying the blind signature technique such as [3][4][6][14]. In Fujioka protocol [14], the voter writes down his choice in the ballot, encrypting it and sending to voting authority to get the blind signature. The voter un-blinds the signature and sends both the encryption and the signature to a voting authority through anonymous channel. After casting phase, voter will check his own encrypted vote then send the decryption key to voting authority anonymously. This protocol protects the content of ballots against announcing the intermediate result. However, it allows voter to
show his vote to the vote-buyers. That makes the uncoercibility and receipt-freeness requirements violated.

In summary, despite its weakness, there are many electronic voting protocols based on blind signature technique for its simple computational operators or small communicational cost. In order to fulfill the privacy of votes in the protocols based on blind signature, authors employ some supplementary techniques such as plaintext equality test in [19], recasting without scarifying uniqueness requirement in [4] …

Conclusion

In this paper, we have introduced the concept of electronic voting, and the reason electronic voting is becoming the new tendency in modern life. We have also provided the analysis about privacy in the voting context and requirements to implement an electronic voting system in the future. In addition, a brief introduction of general architecture along with potential solutions to privacy preserving problem in electronic voting and assessments of the ability to satisfy privacy requirements in voting context and the computational complexity of each solution are presented as well. In the future, we intend to put forth the detailed architecture that aims to address the problem of coercion and vote-buying, and then create a more efficient protocol for internet voting. This protocol not only has a smaller computational complexity but also remains a higher level of privacy.

References


