Seminario Teoria dei Codici - Elezioni elettroniche

Daniele Venzano (venza@iol.it)

10 settembre 2003

Sommario

Il seminario è diviso in due parti, nella prima verranno esposti alcuni protocolli di complessità crescente che utilizzano algoritmi crittografici per ottenere votazioni sempre più sicure, ma mai, come vedremo, completamente soddisfacenti. Esistono algoritmi che permettono votazioni che rispettano tutti i requisiti, ma sono alternativamente, estremamente complessi, o inapplicabili per un numero grande di votanti (a questa categoria appartengono i protocolli distribuiti).

1 Protocolli crittografici

Ci sono almeno sei requisiti che un buon sistema di votazioni deve possedere:

1. Solo gli aventi diritto possono votare
2. Nessuno può votare più di una volta
3. Nessuno può determinare per chi ha votato qualcun'altro
4. Nessuno può duplicare il voto di qualcun'altro
5. Nessuno può cambiare il voto di qualcun'altro
6. Ogni partecipante può verificare che il suo voto sia stato conteggiato

In oltre c'è un settimo requisito che può essere richiesto in alcune situazioni: 7. Ognuno sa chi ha votato e chi no.

Tutti i protocolli si basano pesantemente sulla crittografia a chiave pubblica e si assume che tutte le parti coinvolte posseggono almeno una coppia di chiavi pubblica/privata.

1.1 Primo protocollo

1. Ogni votante firma il voto con la sua chiave privata
2. Ogni votante cripta il proprio voto firmato con la chiave pubblica di una Central Tabulating Facility (CTF)
3. Ogni votante spedisce il proprio voto al CTF
4. Il CTF decripta i voti, controlla le firme, li conta e rende pubblici i risultati.

Questo protocollo soddisfa molte delle proprietà richieste. Il problema è che la firma è associata al voto, quindi il CTF sa chi ha votato per chi, e bisogna fidarsi completamente che non sfrutti questa conoscenza.

1.2 Secondo protocollo

Questo protocollo si divide in due parti, la prima è di preparazione al voto vero e proprio ed è assimilabile alla preparazione delle schede che si fa con le votazioni cartacee.

1. Ogni votante prepara 10 insiemi di messaggi, ognuno contenente un voto valido per ogni possibilità (se il voto richiede una risposta del tipo sì/no, ogni insieme conterrà due voti, uno per il sì e uno per il no). Ogni voto contiene anche un numero casuale abbastanza grande da evitare duplicati.

2. Il votante chiude ogni insieme di messaggi in una busta e manda tutte le buste al CTF. Questo passo si può ottenere anche per via digitale, ma è descritto così per semplicità.

3. Il CTF controlla che quel votante non gli abbia già spedito i voti precedentemente. Poi apre 9 delle dieci buste e controlla che i voti all'interno siano fatti correttamente. Infine firma i messaggi della decima busta (senza aprirla!) e li rispedisce al votante, aggiungendo il nome del votante nel suo database.

4. Il votante apre la busta e si ritrova con un insieme di voti convalidati dal CTF.

5. Il votante sceglie uno dei voti e lo cripta con la chiave pubblica del CTF.

6. Il votante spedisce il proprio voto.

7. Il CTF decripta i voti, controlla le firme, controlla il suo database per cercare numeri di seria duplicati, salva il nuovo numero nel database e aggiunge il voto al conteggio. Alla fine pubblica i risultati, compresi un elenco dei numeri di serie con associato il voto.

È molto difficile truccare questo sistema, se qualcuno cerca di mandare lo stesso voto due volte il CTF se ne accorge al passo 7 grazie ai numeri di serie. Se cerca di far convalidare più di un voto al passo 2, il CTF se ne accorge al passo 3. Non è possibile generare altri voti o intercettare e cambiare quelli di altri votanti perché la chiave privata del CTF non è conosciuta. Il passo 3 è necessario per assicurare che i voti siano unici, senza quel passo è possibile creare diversi insiemi di voti che siano diversi tranne che per il numero di serie (che deve essere unico per impedire che uno possa votare più volte). Il CTF non può determinare chi ha votato per chi perché non ha l'associazione tra i nomi e i numeri di serie. La pubblicazione dei voti con i numeri di serie permette di verificare se i voti sono stati conteggiati correttamente. Questo protocollo ha comunque dei problemi, se il passo 6 non è anonimo, il CTF può ricostruire chi ha votato per chi. Oppure il CTF può generare un gran numero di voti firmati correttamente e mandarli a se stesso. Infine se qualcuno scopre che il suo voto non è stato conteggiato correttamente, non alcun modo di dimostrarlo.

1.3 Terzo protocollo

La soluzione per cercare di limitare i problemi del protocollo 2 consiste nel dividere in due l'autorità centrale, una Central Legitimization Agency (CLA) che certifica i votanti e un CTF separato.

1. Ogni votante manda un messaggio al CLA per chiedere un numero di validazione

2. Il CLA risponde con un numero casuale e mantiene un elenco dei numeri con i nomi delle persone, in modo da verificare se qualcuno cerca di votare due volte.

3. Il CLA manda la lista dei numeri di validazione al CTF, ma tiene per sè la corrispondenza coi nomi delle persone.

4. Ogni votante crea un numero casuale e spedisce al CTF il voto, il suo numero casuale, il numero del CLA in un unico messaggio.

5. Il CTF controlla il numero di validazione sulla lista che ha ricevuto dal CLA. Se lo trova lo segna (per impedire che qualcuno possa votare due volte). Il CTF segna il numero casuale dell'utente nella lista di quelli che hanno votato per un certo candidato e aggiunge 1 al conteggio per quel candidato.

6. Dopo che tutti i voti sono stati ricevuti il CTF pubblica i risultati con gli elenchi dei numeri di identificazione.

Tutti i messaggi tra i votanti e le due autorità devono essere criptati e firmati per impedire che qualcuno possa impersonare altri o intercettare le trasmissioni. Ogni votante può guardare la lista di numeri di identificazione e verificare che il suo voto sia stato effettivamente conteggiato. Il CLA controlla che il CTF non abbia aggiunto dei voti confrontando quanti numeri di validazione ha assegnato con il numero di voti conteggiati. È possibile che qualcuno che non può votare tenti di indovinare un numero di validazione corretto; questo può essere reso estremamente improbabile se i numeri di validazione sono molti di più dei votatori (numeri di 100 cifre per milioni di votanti). Questo protocollo non offre protezione se il CLA o il CTF non si comportano correttamente, il CLA può certificare votanti che non hanno diritto o la stessa persona più volte. Quest'ultimo rischio può essere ridotto facendo pubblicare una lista dei votanti, se ci sono meno votanti che voti, vuol dire che c'è qualcosa che non va. Infine se il CTF e il CLA si mettono d'accordo possono correlare gli elenchi e scoprire che ha votato per chi.

2 Esempi nel mondo reale

Nessuna delle implementazioni attuali utilizza i protocolli visti qui sopra, anzi, l'applicazione di tecniche crittografiche è ancora molto scarsa, anche dove sarebbe richiesta per rendere sicuri i sistemi già in uso. Attualmente sistemi i voto elettronici sono usati in maniera estensiva (e pubblicizzata) solo in tre paesi: Stati Uniti, Brasile e India. Il sistema utilizzato dagli ultimi due è simile ed è considerato abbastanza sicuro. Il sistema brasiliano lascia anche una traccia stampata su carta durante le operazioni di voto che permette un riconteggio manuale in caso di problemi. Purtroppo le informazioni disponibili sui sistemi Indiano e Brasiliano non sono sufficienti per una descrizione completa del loro funzionamento.

2.1 Stati Uniti

Negli Stati Uniti ogni stato è libero di adottare il proprio sistema di voto e conteggio (cartaceo, meccanico o elettronico), sia l'azienda fornitrice, stabilendo propri parametri per la certificazione delle macchine di voto. Con le ultime elezioni presidenziali si sono evidenziati i gravi problemi di funzionamento dei sistemi meccanici in uso e molti stati hanno iniziato a rinnovare le macchine rifornendosi per l'80% da due aziende, considerate leader nel settore dei sistemi elettronici di voto. Queste due aziende sono Diebold e ES&S.

2.1.1 Funzionamento

I sistemi forniti da queste due aziende sono simili nel funzionamento, ogni seggio elettorale ha a disposizione un certo numero di terminali con cui il votante interagisce. Alla fine delle elezioni

ogni terminale spedisce a un server centrale (uno per contea) i suoi totali per il conteggio finale. In nessun caso il passaggio dei dati è autenticato o criptato in qualche modo.
Il votante si identifica con la macchina tramite una smart card (fornita sul momento o spedita per posta prima delle elezioni) o con un codice PIN datogli dagli osservatori nel seggio che verificano la sua identità.
Per votare alcuni sistemi hanno un lettore ottico, altri un touch screen, ma in nessun caso, nelle macchine fornite dalla Diebold o dalla ES&S, viene generato un voto cartaceo che possa essere verificato dall'elettore. A voto avvenuto la smartcard viene messa in uno stato 'cancellato' in modo da non lasciare la possibilità di utilizzarla per votare più di una volta.

2.1.2 Risultati dell'esame del sistema della Diebold

Nessuna delle due aziende che coprono la fetta più grande del mercato rilascia i sorgenti dei suoi sistemi, ma la prima ha lasciato diversi gigabyte di dati a disposizioni su un server FTP con accesso anonimo (non si è ancora capito se per incuria o volutamente). Il sito era continuamente aggiornato e conteneva manuali, sorgenti, pacchetti di installazione già pronti e altro materiale del sistema di voto Diebold. Il server FTP è stato chiuso il 29 gennaio 2003, ma in quel momento era già stato copiato su diversi altri siti in giro per il mondo.
Una parte dei file era protetta con il sistema di cifra fornito con PkZip, estremamente debole e per il quale esistono diverse utility di recupero delle chiavi di cifratura.
Un gruppo di esperti in sicurezza di sistemi informatici della John Hopkins University e della Rice University ha analizzato tutto il materiale non criptato (neanche quello protetto dal debole algoritmo di PkZip) cercando i possibili problemi che permettessero una manipolazione dei risultati del voto. I file protetti non sono stati considerati per prevenire possibili accuse di violazione del DMCA (Digital Millenium Copyright Act). Il sistema è stato trovato estremamente debole nei confronti di attacchi molto semplici e lascia addirittura una backdoor che permette di modificare i conteggi senza lasciare traccia dei cambiamenti.
I problemi sono stati rilevati non solo come errori di programmazione, ma anche nella progettazione dell'intero sistema. Innanzitutto le comunicazioni tra i terminali e il server centrale non sono criptate né autenticate e avvengono su canali molto insicuri (modem o wireless). Inoltre il sistema di smart card utilizzato per autenticare i votanti è abbastanza debole da permettere di votare più volte ad un singolo votante o di accedere al terminale in modalità di amministrazione.

2.1.3 Problemi riscontrati nel server centrale

Il sistema centrale di conteggio (uno per contea o collegio elettorale) è basato su un database di Microsoft Access (che gira su Windows o Windows CE) in cui sono contenute tutte le informazioni di conteggio e di autenticazione (ad es. password di amministrazione). Il database viene utilizzato tramite un software chiamato GEMS che visualizza una serie di report e non permette alcuna modifica. Quando i voti vengono ricevuti vengono registrati in una tabella, della quale vengono poi fatte due copie. La terza non si è ancora capito a cosa serva, mentre la seconda è quella utilizzata per visualizzare i totali per tutto il collegio (ma per visualizzare i conteggi dei singoli seggi viene usata la prima copia).
Anzichè utilizzare GEMS il database è utilizzabile direttamente tramite Access, visualizzandone la struttura e potendo modificare la seconda copia della tabella con i totali, siccome le tre tabelle sono completamente indipendenti le modifiche non vengono riportate nella tabella 1. Ad un esame i conteggi dei sigoli seggi saranno sempre corretti, mentre i totali potrebbero non corrispondere, senza che GEMS se ne accorga o segnali errore.

2.1.4 Problemi sui terminali

I problemi maggiori stanno nel sistema di smart card. Tutte le informazioni sono registrate su di esse in chiaro e allo stesso modo avvengono le comunicazione con l'host. È quindi possibile leggere le schede (e duplicarle) con un normale lettore commerciale. Due smart card speciali permettono rispettivamente di accedere alla modalità di amministrazione e di chiudere un elezione. Queste smart card sono rese speciali da due bit che vengono letti dalla macchina di voto.
Un modo in cui questo può essere sfruttato è un attacco DOS sui terminali, nel momento in cui viene inserita la tessera di fine elezioni la macchina non riceve più i voti degli elettori, se questo avvenisse durante la pausa pranzo, ad esempio, potrebbe scoraggiare parecchi votanti, e nel collegio giusto potrebbe causare un cambiamento significativo nei totali.

2.2 Brasile

Il sistema brasiliano produce per ogni voto una scheda stampata che può essere usata per un conteggio manuale.
Al termine delle votazioni ogni macchina produce diverse copie dei totali su carta più un dischetto. Il tutto sigillato in buste diverse viene consegnato al locale ufficio elettorale. Se il dischetto ha dei problemi le copie stampate possono essere usate inoltre un conteggio manuale è sempre possibile grazie alle schede stampate con i singoli voti.

2.3 India

Il sistema indiano è simile a quello brasiliano, ogni seggio ha diversi terminali più un server che raccoglie man mano tutti i voti e li conteggia. Al termine delle elezioni il server viene sigillato e portato al conteggio dove vengono scaricati e sommati i voti.

3 Bibliografia

Tutto il materiale della sezione 1 è stato preso da *Applied Cryptography* di Bruce Schneier (Capitolo 6 – Esoteric protocols)
La parte sui problemi del sistema Diebold è stata presa da due fonti differenti:

1. *http://www.avirubin.com/vote.pdf* per la parte dei problemi dei terminali
2. *http://www.truthout.org/docs_03/voting.shtml* per la parte dei problemi sul server backend. Questo articolo non contiene un'esposizione rigorosa come il precedente sui problemi dei terminali, quindi è possibile che le informazioni in esso contenuto siano inesatte.

La parte sul Brasile proviene da un post di Daniel Balparda de Carvalho sul numero del 15 gennaio 2001 di CRYPTO-GRAM.
La parte sull'India invece viene da un articolo di Slashdot (http://slashdot.org)