Groupe de travail sur les machines à voter
Éléments de réflexion

Extrait des recherches de Chantal Enguehard - Laboratoire d'Informatique de Nantes Atlantique (LINA)
avec le support du European Computer and Communication Security Institute (ECSSI)- Bruxelles

30 novembre 2007

Les références bibliographiques n'ont pas été intégrées à ce document afin de ne pas le surcharger.

I - Rappel

1 - Définition du système de vote anonyme (extrait du modèle)

En France, les votes politiques doivent respecter plusieurs contraintes "sécuritaires"
- anonymat : chaque électeur vote en confidentialité
- unicité : chaque électeur dispose d'une voix
- non répudiation du vote : un électeur ayant voté ne peut soustraire son vote du système de vote
- confidentialité : il est impossible de relier un bulletin à l'électeur qui l'a exprimé
- sincérité : les résultats proclamés sont conformes aux votes des électeurs
- les votes ne sont pas révélés avant la clôture du scrutin

Lors d'un vote, les électeurs ont le choix entre plusieurs candidats ou plusieurs listes ou plusieurs réponses (cas des referendum).

C est l'ensemble des choix proposés. $C = \{c^i\}$

L'opération de vote se déroule dans le temps
- ouverture du bureau de vote : t_o
- fin du vote : t_n
- dans l'intervalle $]\ t_o\ ,\ t_n\ ]$, des votes sont enregistrés.

Le système de vote répondant aux critères ci-dessus reçoit des entrées **E** (les bulletins de vote) et produit un résultat **R** établi par la fonction f telle que $\mathbf{R} = f(\mathbf{E})$

Chaque **entrée** du système de vote est produite à un temps t, et exprime un choix c. Les entrées ne sont pas publiquement révélées.

Il s'agit d'une paire (t, c) avec $t_o < t \leq t_n$, c non défini a priori

L'ensemble des entrées est noté E $\mathbf{E} = \{ (t_i, c_i) , t_o < t_i \leq t_n \}$

Le système de vote possède plusieurs organes :

— une fonction z qui, appliquée à t, calcule z(t) et telle que la fonction z est injective et non monotone[1], la fonction z n'a pas de réciproque[2]. On note $z(t_i) = t'_i$

— une urne **U** définie comme une suite d'ensembles représentant ses états $\mathbf{(U_n)}$:

au temps $t = t_o$, l'urne est vide : $\mathbf{U_0} = \varnothing$. Le contenu de l'urne est public

L'insertion d'une entrée dans l'urne est définie par : $\mathbf{U_{i+1}} = \mathbf{U_i} \cup \{(t'_{i+1}, c_{i+1})\}$

au temps $t = t_j$, avec $t_o < t \leq t_n$, l'urne comprend les entrées qui ont été déposées entre les moments t_o et t_j (inclus). Son contenu n'est pas révélé.
$\mathbf{U_j} = \{(t'_j, c_j) / \forall i,j, \ t_i \neq t_j \ , t_o < t_j \leq t_n \}$ sans que leur valeur ne soit dévoilée.

au temps $t > t_n$, l'urne comprend les entrées qui ont été déposées entre les moments t_o et t_n (inclus). Son contenu est révélé.
$\mathbf{U_n} = \{(t'_j, c_j) / \forall i,j, \ t_i \neq t_j \ , t_o < t_j \leq t_n \}$

— un opérateur de validation de vote *valide* qui, pour chaque élément de l'urne, détermine sa valeur choisie dans l'ensemble $C \cup \{\text{blanc, nul}\}$

— un opérateur de comptage *comptage* qui s'applique, après la fermeture du bureau, sur les entrées stockées dans l'urne et qui révèle **R**, le résultat du vote. Le résultat est publiquement énoncé.

$$\mathbf{R = \{comptage \ (U_n) \}}$$

R, le résultat du système de vote, est calculé par l'aggrégation des éléments de l'urne.

$$R = \Big\{ \Big(c^k, \sum_{\substack{j=1,n \\ valide((t'_j, c_j)) = c^k}} \delta \ (t'_j, c_j) \Big) \text{ avec } c^k \in C \cup \{\text{blanc, nul}\} \text{ et } (t'_j, c_j) \in U_n \Big\}$$

Ce résultat peut aussi être vu comme la cardinalité de chaque partition de l'urne, l'urne étant partitionnée en fonction de la valeur du choix validé de ses éléments.

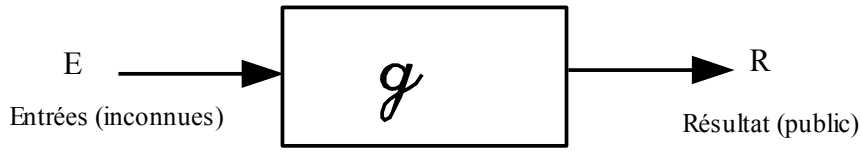
1 La relation d'ordre entre t n'est pas conservée pour z(t)

$\forall \ t_i, t_j \quad t_i < t_j \quad \not\Rightarrow z(t_i) < z(t_j)$

2 Il n'est pas possible de calculer t à partir de z(t).

II - Contrôle de la validité des résultats énoncés par une machine à voter

1 - Évaluation d'une machine à voter



$$R = g(E)$$

La machine à voter est un système de vote devant correspondre au système de vote tel qu'on l'a défini supra.

2 - Question

Peut-on s'assurer que le résultat obtenu du système représente, sans erreur, les votes des électeurs et leur aggrégation calculée selon le modèle du système de vote anonyme défini supra ?

3 - Présentation critique des différentes démonstrations utilisées pour affirmer la justesse des résultats

Il existe, en pratique, trois types de démonstrations en usage

1. Vérification par approximation :

   Il s'agit de mesurer l'écart entre les résultats fournis par le système évalué et des estimations statistiques.

2. Garanties apportées par les traitements :

   L'idée générale est de vérifier que le processus qui transforme les entrées pour produire un résultat est conforme à la définition du système de vote tel qu'on l'a défini supra.

3. Preuve de résultat :

   Il s'agit de prouver un lien de continuité ou d'identité entre les entrées et le résultat.

Si l'une de ces démonstrations peut être établie, il sera démontré qu'il existe un contrôle de la qualité des résultats fournis par le système évalué.

Si aucune de ces démonstrations ne peut être établie, il sera démontré qu'il n'existe aucun contrôle de la qualité des résultats fournis par le système évalué.

3.1 - Vérification par approximation

Il s'agit de mesurer l'écart entre les résultats fournis par le système évalué et des estimations statistiques.

Il existe deux types d'estimation statistiques qui seront examinées successivement :

- estimations statistiques concurrentes au vote : elles sont établies en même temps que les opérations électorales. Il 'agit des sondages de sorties d'urnes.

- estimations statistiques antérieures au vote : elles sont fondées sur des données recueillies avant les opérations électorales. Il s'agit des résultats aux précédentes élections et/ou des intentions de votes pondérées.

3.1.1 - Estimations statistiques concurrentes au vote

Les évaluations statistiques sont calculées sur un échantillon des électeurs auxquels il est demandé de révéler leur vote (sondage de sorties d'urnes).

Assertion

« *Les résultats sont justes puisqu'ils ne contredisent pas les sondages de sorties d'urnes.* »

Analyse de la validité de l'assertion

Les estimations statistiques dépendent de la qualité de l'échantillon sur lequel elles sont établies ainsi que du savoir-faire des personnes qui procèdent au sondage. Il existe plusieurs facteurs de biais :

- l'échantillon est de taille réduite et ne peut représenter certains segments de la population.

- le vote étant révélé, les sondés ont la possibilité de ne pas répondre à la question qui leur est posée.

- il n'existe aucune contrainte garantissant le fait que chaque personne sondée exprime son vote réel.

- les personnes procédant au sondage peuvent, consciemment ou inconsciemment, favoriser la représentativité de certaines catégories de la population au détriment d'autres catégories de la population.

Les sondages de sorties d'urnes peuvent être considérés comme le résultat d'un système de vote dont les modalités sont :
- votes révélés,
- échantillon partiel
- absence de contrainte validante
Ces modalités en font un système plus faible que le système de vote anonyme qui procède sur l'ensemble des électeurs.

Nous disposons donc de deux systèmes de vote, qui fournissent des résultats différents, et dont les modalités sont différentes : l'un procède par vote anonyme et l'autre par vote révélé ; l'un concerne tous les électeurs, l'autre n'opère que sur un échantillon.
Les deux système fournissent des résultats différents.

Nous procédons à un raisonnement par l'absurde pour déterminer lequel de ces deux systèmes de vote doit être considéré comme susceptible de fournir les résultats ayant la plus grande exactitude.

première possibilité : c'est le système de sorties d'urnes

La conséquence immédiate est qu'il devient inutile de déployer des élections concernant l'ensemble des électeurs, puisque, quels que soient les résultats de la consultation électorale, leur validité sera déterminée en fonction de leur conformité avec les sondages de sorties d'urnes.

deuxième possibilité : c'est le système de vote anonyme

Dans ce cas, les sondages de sorties ne peuvent être utilisés comme référence pour évaluer la validité des résultats du système de vote anonyme.

Par conséquent, les sondages de sorties ne peuvent servir de référence pour évaluer la sincérité d'un système de vote anonyme.

3.1.2 - Estimations statistiques antérieures au vote

Les estimations statistiques sont établies principalement par l'analyse d'élections précédentes et/ou des intentions de votes pondérées, intentions exprimées avant le déroulement des opérations électorales.

Assertion

> « Les résultats sont justes puisque la tendance du bureau de vote (observée lors des précédentes élections) est conforme à ce qui était prévu. »

Analyse de la validité de l'assertion

1. En démocratie, la tenue d'élections à intervalles réguliers consacre la possibilité d'avoir des alternances politiques. D'une élection à l'autre, les électeurs peuvent progresser dans leur culture et leur analyse politique, changer d'avis et exprimer, par leur vote, une opinion en rupture avec les opinions qui étaient les leurs dans le passé

2. Les accidents de prédictions sont toujours possibles. L'évolution des résultats électoraux est un phénomène dynamique complexe faisant intervenir de multiples paramètres. Il n'existe pas d'approche scientifique valide qui permette de prédire avec justesse les résultats de tels phénomènes en se fondant sur l'analyse du passé [1].

Les estimations statistiques fondées sur l'analyse du passé ne peuvent servir de référence pour évaluer la sincérité d'un système de vote anonyme.

3.2 - Garanties apportées par les traitements

Il s'agit de prouver que les traitements mis en oeuvre dans le système de vote à évaluer transforment les entrées du système de la même manière que le système de vote anonyme défini supra.

Il faut prouver que le système de vote est (1) immune de faute **et** (2) ne peut être altéré

Dispositif de vote traditionnel avec urne transparente et dépouillement public

(1) : le matériel de vote traditionnel est inerte et ne peut provoquer de faute.

(2) : du fait des propriétés physiques du plexiglas, du papier et de l'encre, et en l'absence de médiateur, le dispositif de vote traditionnel ne peut être altéré sans que le contrôleur et les parties présentes ne le voient (urne ouverte, bourrage d'urne, substitution d'urne, etc.) dans la mesure où la surveillance de l'urne et de l'opération de dépouillement sont effectuées avec l'attention nécessaire.

Les lois énoncées dans le code électoral fixent précisément le déroulement du dépôt des votes dans l'urne et du dépouillement de l'urne. Les améliorations qui ont été apportées à ce dispositif (la plus récente étant l'utilisation obligatoire d'urnes en plexiglas transparent) ont contribué à le sécuriser. Des ajustements sécuritaires peuvent encore être recherchés dans le but de parfaire ce dispositif de vote manuel et de le mettre à l'abri de fraudes manuelles.

Machine à voter électronique

(1a) : Tests

Le démarche de test pose comme hypothèse que le système ne sera aucunement altéré : le logiciel de vote doit être strictement identique (à la virgule près) à celui qui a été testé.

Des tests unitaires sont réalisés sur les composants puis des tests d'intégration sont menés sur la machine à voter.

Ces tests, s'ils peuvent valider un dispositif à un instant donné et pour un ensemble restreint d'interactions entre les utilisateurs et la machine, ne valident pas le dispositif pour l'ensemble des interactions possibles entre les utilisateurs et la machine car il n'est pas économiquement possible de tester toutes les interactions possibles .

Tester une machine à voter pour un jeu particulier d'interactions entre l'utilisateur et la machine ne prouve pas son bon fonctionnement dans tous les cas.

(1b) : Preuve formelle

Une preuve formelle est une démarche mathématique qui permet de tester virtuellement toutes les entrées possibles d'un programme et de vérifier que les sorties seront conformes à ce qui est attendu. Cette démarche pose comme hypothèse que le système ne sera aucunement altéré : le logiciel de vote doit être strictement identique (à la virgule près) à celui qui a été testé.

Dans le cas d'un système complexe comprenant de multiples composants eux-mêmes complexes (drivers, micro-code, etc.), la démarche de preuve doit être étendue à tous les composants et sous-systèmes.

(2) : Une machine à voter électronique peut être altérée sans qu'on puisse le voir

- le logiciel de vote, ou un autre programme présent sur la machine à voter, ne sont pas identiques à l'ensemble des logiciels initialement validés

- le logiciel de vote produit une erreur d'exécution, sans tomber en panne, à cause d'un défaut du matériel électronique (hardware).Cette erreur peut modifier des votes.

Effectuer une démarche de preuve formelle sur l'ensemble des programmes équipant un ordinateur de vote ne protège pas contre les erreurs d'exécution ou les problèmes matériels ou ... les modifications de configuration.

3.3 - Preuve de résultat

cas 1 : vote sans médiation

Cette preuve est valide si le vote se déroule sans médiation

C'est le cas du vote à main levée car le comptage peut être réalisé par plusieurs personnes à la fois mais, dans ce cas, le vote n'est pas anonyme.

cas 2 : vote avec médiation et matérialisation du vote

L'opération de vote est matérialisée dans un bulletin glissé dans une urne transparente par chaque électeur.

Il existe une continuité physique ininterrompu depuis le bulletin glissé dans l'urne par l'électeur jusqu'au bulletin compté lors du dépouillement sans possibilité de modification lors du stockage dans l'urne.

S'il y a une surveillance constante et rapprochée de l'urne par toutes les parties alors il est quasiment impossible de la substituer ou d'en modifier le contenu.

Durant l'opération de comptage effectuée publiquement, si la surveillance effectuée par le public et toutes les parties en présence s'exerce de manière efficace, on peut peut alors considérer que les entrées reçues ne peuvent être modifiées et que le comptage est juste.

cas 3 : vote avec médiation sans matérialisation du vote

On ne peut comparer les résultats énoncés par le système de vote avec les entrées qu'il a reçues car ces entrées sont inconnues du fait de l'anonymat. Et de plus, comme son nom l'indique, la dématérialisation exclut toute continuité physique.

Il a été évoqué des procédures de tests tendant à établir une preuve de résultat immédiatement avant chaque opération de vote réelle. Dans ce cas, les données entrées sont soigneusement notées et comparées avec le résultat fourni par la machine à voter.

Ce dispositif est inopérant car :

- il s'appuie sur l'assertion qu'une performance de passé réussie vaut garantie de résussite pour la performance à venir.

- il repose sur un jeu de données en entrée excessivement réduit (voir 3.2 (1a)).

La dématérialisation du vote interdit de vérifier la conformité des résultats énoncés avec les entrées reçues.

Les machines à voter avec trace papier, dans lesquelles les bulletins sont dématérialisés puis rematérialisés, n'apportent pas davantage de garanties (voir la note complémentaire points 1 et 2).

4 -En synthèse

On ne dispose d'aucun outil permettant de prouver que le résultat énoncé par la machine à voter correspond à l'aggrégation des votes qui ont été exprimés.

III - procédures informatiques visant à améliorer la confiance des citoyens

La confiance des citoyens dans le système électoral peut être considérablement améliorée par une meilleure transparence du système de vote.

Les procédures en vigueur (avec urne transparente et dépouillement public dans les bureaux de vote) permettent aux citoyens, scrutateurs, parties en présence et contrôleurs de suivre pas à pas la procédure de vote (garde de l'urne) et le dépouillement. Une étude de sécurité appropriée pourrait encore en améliorer l'exactitude. Les résultats de chaque bureau de vote sont établis avec transparence et n'ont donc pas besoin d'être vérifiés.

En revanche, les procédures en vigueur pour la totalisation des résultats des bureaux de vote ne sont pas publiques.

Les résultats de vote affichés au niveau national ne détaillent pas les résultats bureau par bureau : il n'est pas possible d'effectuer la vérification individuelle du bon report des résultats d'un bureau de vote.

Préconisation 1 : le site du ministère de l'intérieur pourrait afficher les résultats détaillés de chaque bureau de vote sur son site internet :

- nombre d'inscrits
- nombre de votants
- nombre d'émargements
- nombre de procurations
- nombre de bulletins blancs ou nuls
- total de voix obtenu par chaque candidat / liste / choix (en cas de referendum).

Ces informations sont publiques et peuvent accroître la confiance des électeurs dans le système électoral.
Cet affichage ne présente aucune difficulté technique ou éthique et peut être réalisé à faible coût.

Préconisation 2 : les procès-verbaux des bureaux de vote pourraient être digitalisés et mis en ligne pour la durée du contentieux électoral après anonymisation[3] des formulaires. Cette démarche s'inscrit dans le courant de la dématérialisation des actes administratifs et pourrait faciliter considérablement le travail du Conseil Constitutionnel.

La publication des procès-verbaux est susceptible d'apporter des informations quant à la réalité du terrain et aux possibilités d'accroître l'exactitude du système de vote. Ces informations peuvent contribuer à améliorer l'analyse de la sécurité du système de vote.

Cet affichage ne présente aucune difficulté technique ou éthique et peut être réalisé à faible coût.

3 Les identités ne sont pas effacées des documents originaux.

Note complémentaire

Contrôle de la validité des résultats énoncés par d'autres dispositifs électroniques

1 - Machine à voter imprimant les bulletins papier à la clôture du bureau

Les électeurs n'ont pas eu la possibilité de vérifier leur bulletin. Ce dispositif est équivalent aux machines à voter et est analysé en suivant les mêmes démonstrations (voir partie II.3).

On ne dispose d'aucun outil permettant de prouver que le résultat énoncé par une machine à voter imprimant les bulletins papier à la clôture du bureau correspond à l'aggrégation des votes qui ont été exprimés.

2 - Machine à voter avec bulletin papier vérifié par l'électeur

La partie II a démontré qu'en l'absence de matérialisation du vote, il est impossible de contrôler la justesse des résultats énoncés par une machine à voter.

Les machines à voter produisant un bulletin que se doit de vérifier chaque électeur procèdent à

- une matérialisation du bulletin (cette opération est susceptible d'erreur)

- un stockage des bulletins matérialisés dans une urne (approche mécanique sujette à panne)

Pour chaque machine, existent donc deux issues :

1 - Le résultat de l'urne est systématiquement et intégralement dépouillé

Le dispositif se ramène alors à un doublage du dispositif traditionnel de vote par un dispositif électronique.

2 - Le résultat de l'urne n'est pas systématiquement dépouillé

Le dispositif est alors équivalent à une machine à voter sans matérialisation des bulletins de vote et donc, sans aucun contrôle de l'exactitude des résultats.

Une vérification effectuée avec succès sur une machine à voter ne peut **en aucun cas** prouver le bon fonctionnement d'une autre machine à voter.

Le contrôle doit donc être effectué pour ***toutes*** les machines à voter.

3 - Comptage électronique des bulletins de vote

Il existe des urnes électroniques dans lesquelles chaque électeur introduit son bulletin. L'urne électronique fournit un résultat de vote à la clôture du bureau de vote.

Ce dispositif est équivalent à une machine à voter avec bulletin papier vérifié par l'électeur et est analysé en suivant les mêmes démonstrations (voir partie 2 de la note complémentaire).

Une vérification effectuée avec succès sur un dispositif de comptage électronique des bulletins ne peut **en aucun cas** prouver le bon fonctionnement d'un autre dispositif de comptage électronique des bulletins.

4 - Machine à voter avec vérification a posteriori par les électeurs

De nouveaux dispositifs de vote permettent aux électeurs de (1) vérifier, individuellement, si leur vote a été enregistré et (2) prétendent autoriser également la vérification de l'aggrégation des votes.

Les deux parties (1) et (2) doivent être valides pour que ces dispositifs soient prouvés.

(1) vérification individuelle de l'enregistrement de chaque vote

cas 1 : l'électeur dispose d'une preuve de son vote.

Cette preuve de vote le rend vulnérable aux tentatives de coercition (pressions, vente de votes).

Cette possibilité doit être écartée.

cas 2 : l'électeur ne dispose pas d'une preuve de son vote.

Si, lors de la vérification, l'électeur détecte une modification de son vote, il ne pourra en apporter la preuve, notamment en cas de règlement juridique du contentieux électoral.

Cette possibilité n'est pas opérationnelle et ouvre la porte aux allégations non fondées, elle doit donc être écartée.

La partie (1) n'étant pas valide, il est inutile d'examiner la partie (2).

La possibilité que chaque électeur vérifie individuellement son vote ne permet pas de valider les résultats du dispositif de vote.